Phenotypic and Genetic Characterization of Domestication and Yield Component Traits in the Perennial Grain Crop Intermediate Wheatgrass

A DISSERTATION SUBMITTED TO THE FACULTY OF THE UNIVERSITY OF MINNESOTA BY

Kayla Rae Altendorf

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

James A. Anderson
Advisor

April 2020

Acknowledgements

I am grateful to my advisor, Dr. James Anderson, who has been my best advocate and has provided calm, patient, and unwavering support to me throughout both my M.S. and Ph.D. degrees. Jim, you allowed me the freedom to direct my own projects but were there for me with an open door whenever I needed your expertise and advice. I also want to thank my former advisor Dr. Donald Wyse for agreeing to meet with me on that fateful day in 2012 to discuss graduate opportunities in plant breeding, a field I knew little about at the time. I am grateful to both Jim and Don for taking a chance on me and providing me what has been the opportunity of a lifetime, one that has altered the trajectory of my life, and opened doors for my future that I never thought possible. I am grateful to my committee member and mentor Dr. Lee DeHaan. Lee, your insight played a critical role in the success of the projects described herein and I learned a tremendous amount from you along the way. You kept me on my toes in the field, teaching me to always look for ways to improve a process by increasing efficiency and quality. I am fortunate to have had you as a close collaborator. I am also indebted to Dr. Steve Larson of the USDA-ARS in Logan, UT for his patience and dedication in teaching me to use and understand methods for genetic mapping and map creation in outcrossing species and for his significant contributions to this work. I also want to thank my collaborator, Dr. Xiaofei Zhang, who helped design and implement this project and set me up with a solid foundation upon which to build. I also want to acknowledge Dr. Jared Crain, Dr. Jeffrey Neyhart, and Dr. Kevin Dorn for their guidance, coding resources, and contributions. I am incredibly grateful to the many individuals who provided technical support, notably Brett Heim and Marty Christians, as well as students from the MAST program at the University of Minnesota and the 2017 and 2018 interns at the Land Institute. I am grateful for my funding support from the Perennial Agriculture Project in conjunction with the Malone Family Land Preservation Foundation and The Land Institute. Finally, I want to thank my husband, Dr. Garett C. Heineck, for his contributions to Chapter 1, for his encouragement, and for being there for me throughout the highest highs and lowest lows of graduate school. Garett, it has been both an honor and adventure to go through this together and I am incredibly grateful I have you by my side. I look forward to our

future and the many problems we will solve together. Finally, I want to thank my parents for teaching me about dedication, hard work and self-confidence. Mom, thanks for always checking in on me, for your numerous offers to help, and for never missing a seminar or ceremony.

Dedication

This doctoral dissertation, the first to be completed in the history of my extended family, is dedicated to my late grandparents who were all farmers in Minnesota: Jerry E. and Joanne M. (Patnode) Pettit and Raymond M. and Marian E. (Eull) Altendorf.

Abstract

Chapter 1

Perennial plants provide extensive environmental services, and increasing their prevalence on the agricultural landscape is one way to improve sustainability. Direct domestication of intermediate wheatgrass (*Thinopyrum intermedium*) as a perennial grain crop is underway, and selection has focused primarily on improving seed size and yield. Breeders are limited by the lack of understanding of yield and its relationship with component traits in this species. We characterized a large population (n = 1168) of IWG spaced plants in St. Paul, MN and Salina, KS in 2017 and 2018 for a series of 13 yield component traits. Family by environment by year interaction was highly significant ($P < 0.001$) for all traits, indicating that family performance and rankings were variable across environments. In year two in St. Paul, yield plant^{-1} and reproductive tiller numbers nearly doubled, while all other yield components, including yield spike^{-1}, thousand grain weight, spikelets spike^{-1} and florets spikelet^{-1} significantly decreased. Bivariate correlation analyses between component traits revealed consistent trends across environments and highlighted positive associations of seven traits with grain yield. Structural equation modeling (SEM) was conducted using an initial path model that more clearly delineated relationships into direct and indirect effects. When yield is measured on a yield spike^{-1} basis, floret site utilization was the primary contributor to yield, followed by spikelets spike^{-1}, florets spikelet^{-1} and thousand grain weight. When measured on a plant^{-1} basis, reproductive tiller number was the most significant contributor to yield in all environments. Considering the limited indirect effects of biomass and maturity component traits, and the potential for strong indirect selection for tillering when selecting for yield plant^{-1}, we suggest a two-step selection procedure on yield spike^{-1} and floret site utilization and explored alternative methods for collecting data on this labor-intensive trait. Future work should test the predictive ability of yield spike^{-1} and floret site utilization in IWG spaced plants and seed production swards.

Chapter 2

Intermediate wheatgrass (IWG) is an outcrossing, cool season grass species currently undergoing direct domestication as a perennial grain crop for human consumption. Selection targets for this crop are numerous, and breeding may be made more efficient by improving knowledge of the underlying genetic control for traits of interest. Nested association mapping (NAM) has proven useful in dissecting the genetic control of important agronomic traits in a wide range of crop species. This work introduces an intermediate wheatgrass NAM developed by crossing ten phenotypically divergent donor parents to a low-shattering common parent in a reciprocal manner, yielding 1,168 F_1 progeny from 10 families. Using genotyping by sequencing, we identified 8,003 SNP markers and developed a population-specific consensus genetic map with 3,144 markers across 21 linkage groups. Using both genome wide association mapping (GWAS) and linkage mapping both combined across and within families, we characterize the genetic control of flowering time measured in two different ways across two locations and two years. We detected 35 QTL in GWAS and 20 in linkage mapping, demonstrating the complex genetic control of flowering time that is variable across years and locations as well as within families. The IWG NAM population was effective at detecting previously identified QTL, as well as a new QTL that aligns closely to the well-characterized flowering time orthogene from barley, *Ppd-H1*. Results demonstrated the utility of the NAM for understanding of the genetic control of flowering time and should be applied to additional traits of interest in IWG.

Chapter 3

Perennial grain crops have the potential to improve agricultural sustainability, but few existing species produce sufficient grain yield to make this an economically viable option. The perennial forage species intermediate wheatgrass (*Thinopyrum intermedium;* IWG) has shown promise in undergoing direct domestication as a perennial grain crop using phenotypic and genomic selection. However, decades of selection at the current pace will be required to achieve yields on par with annual small grain crops. Marker aided selection could accelerate progress if important genomic regions associated with

domestication were identified. Here we utilize the IWG Nested Association Mapping (NAM) population, with 1,168 F_1 progeny across ten families to dissect the genetic control of brittle rachis, floret shattering, free-threshing ability, tillering, floret site utilization, plant height, and thousand grain weight. We used genome wide association analysis (GWAS) with 8,003 SNP markers and linkage mapping using within-family and combined analyses using a custom genetic map, with a robust phenotypic dataset collected from four unique year by location combinations. A total of 32 QTL in GWAS and 18 in linkage mapping were detected in at least two environments, and most large effect QTL were in common across the two analysis methods. We reveal that the genetic control of domestication traits in IWG is complex, with significant QTL across multiple chromosomes, sometimes within and across homoeologous groups and effects that vary across families. However, in many cases these QTL align closely with putative orthogenes for known domestication traits in related species and may serve as precise targets of selection and directions for further study to advance the domestication of IWG.

Table of Contents

List of Tables

List of Figures

Chapter 1

Floret Site Utilization and Reproductive Tiller Number are Major Drivers of Yield in Intermediate Wheatgrass Spaced Plants

Introduction

Intermediate wheatgrass [*Thinopyrum intermedium;* (Host) Barkworth & D.R. Dewey; IWG hereafter] is a cool-season grass that is currently undergoing direct domestication as a perennial grain crop for human consumption (DeHaan et al., 2014). IWG is an outcrossing, primarily self-incompatible allohexaploid (2n = 6x = 42) native to Europe and Asia, and introduced to western North America in 1932 (Dewey 1962; Ogle et al, 2011). It is a rhizomatous grass with a spike inflorescence featuring multiple sessile spikelets, one at each rachis node. Within each spikelet is a series of florets, each including a lemma, palea, and a caryopsis with varying levels of fertility (Larson et al., 2019). In 1983, researchers at the Rodale Research Center (Kutztown, PA) began surveying a panel of nearly 100 perennial grass species for their domestication potential, and in 1988 initiated germplasm evaluation of IWG (Wagoner, 1990; Shauer, 1988). IWG was chosen due to its relatively large seed size, amenability to mechanical harvest, nutritional similarity to annual wheat, and vigorous perennial growth (Wagoner, 1990). After two cycles of recurrent selection for fertility and seed size, the germplasm was transferred to The Land Institute (Salina, KS; TLI hereafter) where phenotype-based recurrent selection began in 2003 (Cox et al, 2010). Breeding programs also were established in 2011 at the University of Manitoba (DeHaan et al., 2014) and the University of Minnesota (Zhang et al., 2016) within what is now known as the Forever Green Initiative.

Perennial agroecosystems are known to provide more environmental services compared with annual cropping systems, including reduced soil erosion and nutrient leaching as well as increased carbon sequestration and water use efficiency (reviewed in Glover et al., 2010 & Pimentel et al., 2012). These benefits are facilitated by the presence of year-round living cover and extensive root systems that are able to access water and nutrients more efficiently throughout the soil profile. IWG specifically is known to more efficiently utilize available soil water and applied N, and is predicted to have less NO3-N leaching compared to annual crops like wheat and maize (Culman et al., 2013; Jungers et al., 2019). IWG also has shown high rates of evapotranspiration, which suggests that most precipitation in this system is either used by the plant or recharges groundwater, and

not lost to surface runoff (de Oliveira et al., 2018). IWG also has higher net carbon uptake compared with annual crops (de Oliveira et al., 2018), and soils under IWG display increased rates of carbon mineralization (Culman et al., 2013). These responses provide evidence for substantial carbon sequestration potential, though the extent to which this occurs may depend in part on soil type (Sprunger et al., 2017).

While perennials can provide extensive ecosystem services, they generally allocate fewer resources to seed production compared to selected annuals (Vico et al., 2016). This limitation is largely a result of the environment in which they evolved - one that favors longevity - and can be overcome by intense selection (DeHaan et al., 2004). For example, after two cycles (TLI Cycle 2) of recurrent selection for grain yield spike^{-1} and seed mass, production swards yielded 100-108% (depending on seeding date) more than the forage cultivar 'Rush' (961-893 kg ha^{-1} vs 456-443 kg ha^{-1}) (Jungers et al., 2017). After 5 cycles (TLI Cycle 5) grain yield spike^{-1} was predicted to have increased by 143% and seed mass by 60% in a spaced plant setting (DeHaan et al., 2018). Despite these improvements in seed mass and yield, seed yield in production swards of IWG are known to decline with age, starting with the second year (Jungers et al., 2017). For example, in Minnesota environments, yield of TLI Cycle 2 germplasm decreased in the second through fourth years of production by 80% in unfertilized plots and 65% in fertilized plots (Tautges et al., 2018). The exact cause for yield decline is not known, but sod-bounding (Culman et al., 2013), decreases in plant density (Jungers et al., 2017), suppression of reproductive tiller initiation (Frahm et al., 2018) and reductions in harvest index (Tautges et al., 2018) have been suggested.

Due to the heterogeneous nature of intermediate wheatgrass, individual genets (genetically unique individual plants) are evaluated in spaced plant nurseries, typically on 1-m centers, as is customary in forage breeding. Individuals that yield well in this environment, either on a yield spike^{-1} (derived from a subsample of 3-10 spikes) basis, and in some cases, yield plant^{-1} (all spikes in a spaced plant) basis, are selected as parents for crossing blocks for both population improvement and cultivar development. Interestingly, in this less competitive environment, yield is known to nearly double when measured on a yield plant^{-1} basis in the second year (Zhang et al., 2016; Cattani et al.,

2016), shining light on a decades-old dilemma relating to the low predictive ability of performance in spaced-plants to seed production environments both for forage and seed yield. In the recent history of breeding IWG for grain yield, there have been no formal published reports of the evaluation of the same IWG germplasm in both spaced and seed production sward environments, but breeders are aware of the challenges associated with this breeding methodology, and may choose to evaluate half-sib families of parental materials in sward environments as part of the selection process.

Yield is a highly complex trait controlled by many genes and influenced by many component traits. For example, in the case of IWG, traits like reproductive tiller number, spikelets spike^{-1}, florets spikelet^{-1}, thousand grain weight and floret site utilization, all play a role in determining yield. IWG breeders have a relatively limited understanding of which yield component traits are the most important contributors to yield. From bivariate correlation analyses with yield plant^{-1} in IWG, one study showed strong positive correlations with biomass plant^{-1}, height, and head weight; albeit no correlation with heading date (Zhang et al., 2016). In another analysis of component traits and yield plant^{-1}, spikelets spike^{-1} and head weight had the strongest correlations with yield (Bajgain et al., 2019a). In a study using a path analysis approach of four component traits, biomass and heads plant^{-1} were the most important for yield plant^{-1}, followed by seed mass (Cattani et al., 2018). Another study dissected the genetic control of seed size and seed dimensions in IWG using QTL mapping (Zhang et al., 2017). These analyses have provided useful information regarding the dynamics of yield in IWG as measured as yield plant^{-1} in a spaced plant setting, and in some cases also have identified QTL as well as suggestions for utilizing them as fixed effects in genomic selection (e.g. Zhang et al., 2017; Bajgain et al., 2019a). However, the relative importance of these component traits on yield has yet to be demonstrated, and such an analysis, followed by validation, would provide insight into a trait that may be a candidate for direct or indirect selection.

One approach to identifying important component traits is to assess trait correlations with seed yield; however, this approach ignores the complex causal relationships that exist among traits, and the confounding indirect effects that one trait may have on another. Path analysis, developed by Sewall Wright in the 1920's, is a

method for assessing the strength and relative importance of relationships between multiple variables and has the ability to parse out the direct effect and indirect effects that one trait has on another (Wright, 1921). This approach was made popular in the assessment of yield component traits after the publication of a seminal paper on crested wheatgrass (Dewey & Lu, 1959) and has since been used to assess yield in barley (Dofing & Knight, 1992), perennial ryegrass (Abel et al., 2017); meadow fescue (Fang et al., 2004); wheat (Donaldson et al., 2001); and durum wheat (Garcia del Moral et al., 2003), among others. One limitation of path analysis is that it assumes the a-priori defined relationships between variables are correct (Lamb et al., 2011). Structural equation modeling (SEM) is a modern form of path analysis, and is especially fitting for the analysis of yield components as it allows the data to best define the relationships between variables (Lamb et al., 2011). In SEM, an initial model is defined by the investigator *a priori* using theoretical knowledge, the literature, and expected relationships and is then adjusted to fit the data based on modification indices, which are unresolved covariances in the covariance-variance matrix (Grace & Keeley, 2006). In several cases, path analysis and SEM have shown the ability to uncover trends that would not have been observed in simple univariate or bivariate analyses (e.g. Garcia del Moral et al., 2003; Grace, 2006). For example, in a study of yield components in durum wheat, path analysis revealed the importance of spikes per square meter, and kernels spike^{-1} for yield in certain environments, whereas correlation analyses suggested otherwise. In this case, path analysis was insightful for providing traits as selection targets for increasing yield (Garcia del Moral et al., 2003). As such, considering the limited information on yield components in IWG, utilizing both correlation and SEM approaches has the potential to reveal important trends in the data not seen using simple correlation analyses.

The objectives of this research were twofold: 1) To evaluate a population of 1,168 IWG genets in a spaced-plant environment for grain yield and component traits in 4 environments at St. Paul, MN and Salina, KS in 2017 and 2018; and 2) To assess the relationships among component traits and yield spike^{-1} and yield plant^{-1} using correlation analyses and SEM. An improved understanding of the dynamics of yield and its component traits in a spaced plant environment may increase the efficiency and precision

of selection in breeding programs by identifying important selection targets. Furthermore, the population used in this study was developed for QTL mapping, and this analysis will provide direction into which component traits should be subjected to further analysis based on their relative contributions to yield.

Materials and Methods

Germplasm

The germplasm used in this study was developed as a Nested Association Mapping (NAM) population (Yu et al., 2008). In brief, ten phenotypically diverse parents were selected from UMN Cycle 2 (derived from TLI Cycle 3), and were crossed to one common parent in a reciprocal fashion to form 10 F1 families with approximately 117 individuals each (total n = 1168). The progeny, along with at least 6 replicates of each parent, were propagated into four clones each into 5 x 5 cm Jiffy Pots (Plantation Products, Norton, Massachusetts) and transplanted in a spaced plant design on 1-m centers in a RCBD with two blocks in Salina, KS at TLI using a jab-type planter (Johnny's Selected Seeds, Winslow, Maine) and St. Paul, MN at the University of Minnesota Agricultural Experiment Station (STP; hereafter) using a mechanical transplanter in fall 2016. A two-plant border was established around the population to limit edge effects. Plots were hand weeded and cultivated with a multivator (Ford Distributing, Marysville, OH) as necessary to control weeds. A pre-emergent herbicide, Dual II Magnum (S-metolachlor, Syngenta US), was applied at STP in April 2017 and May 2018 at a rate of 1.75 L ha^{-1}. Herbicides were not used at TLI. At both locations, plots were mowed to 15 cm height after harvest and fertilized with urea (56.0 kg ha^{-1} at STP; 78.5 kg ha^{-1} at TLI) in fall 2017.

Weather and Precipitation Data

Weather data for St. Paul was obtained from the National Oceanic Atmospheric Administration (https://www.noaa.gov/) from the St. Paul Agricultural Experiment Station (Station ID: USC00218450). Weather data for Salina, KS was obtained from the TLI Weather Station for 2017 and 2018, and data for ten year averages were obtained

from the Salina Municipal Airport (Station ID: USW00003919). Ten year averages, from 2006 to 2016, were calculated for daily average temperature [(daily min + daily max)] / 2 and precipitation accumulation, and the results were plotted alongside values for 2017 and 2018 (Supplementary Figure 1.1).

Evaluation of Maturity, Flag Leaf Area and Height

Due to variability in spike length among genets, it is difficult to visually estimate the proportion of the spike that has emerged from the boot using just one growth staging method (eg. Feekes or Zadoks). To address this, when spikes were approximately 50% emerged on average, the length of the spike emerged from the boot was measured in centimeters on one representative spike per plant. In cases of high within-plant variability (e.g. ~5 cm), the average of two was recorded. This number was divided by the final spike length calculated after harvest to determine percent emerged. In IWG, early spike emergence is not always coupled with early anthesis. Thus, Feekes maturity was also recorded for each plant when anthers were showing (approx. 1700 h) and when approximately 50% of the plants were in anthesis. Stages were coded as an ordinal variable from 1-10 for data analysis (ranging from boot stage [1] to kernels watery ripe [10]). When flag leaves were fully expanded after anthesis, the width (mm) and length (cm) were measured on one representative flag leaf per plant using a transparent ruler. In cases of high within-plant variability (e.g. ~5 cm length; ~5 mm width), an average of two leaves was used. These numbers were multiplied to approximate flag leaf area (cm^2). Plant height was recorded in cm at maturity when the plants reached maximum height.

Harvest and Reproductive Tiller Ct.

Harvest occurred in two stages: 1) ten spike sample; and 2) reproductive tiller counts and remaining spikes harvest. First, ten of the larger, most uniform spikes were cut ~5 cm below the peduncle from each spaced plant. This sample is referred to as the ten-spike sample. IWG spaced plants have exceptional within-plant variability, especially in the first year; thus, uniform spikes were chosen not for selection, but to limit experimental error (DeHaan et al., 2018). Ten-spike samples were placed in 61 cm

unwaxed paper bread bags and were stored upright with peduncle down in cardboard boxes to avoid shattering or breakage during transport and storage. Second, the remaining spikes from each spaced plant were gathered and clipped ~10 cm below the peduncle using a large garden shears. The underside of the cut bundle was photographed in the field using a Canon digital SLR camera and stems were later counted on the computer by clicking with the "multi-point" function in ImageJ (Supplementary Figure 1.2; Schneider et al., 2012). Considering the height of the cut, we assumed all previously cut stems were included, and that all white circles in the image represented seed-bearing tillers. Previous observations revealed that vegetative (non-seed-bearing) tillers do not reach such a tall height in spaced plant environments (data not shown). Due to storage and transportation limitations, remaining spikes were only bagged and threshed to collect yield data in one block per environment. In 2018, remaining spike samples were first threshed in the field using a Small Vogel Plot Thresher (ALMACO, Nevada, Iowa) to reduce the volume of the sample. All samples were dried for one week at 35 °C. Remaining spike yields were not harvested for TLI in 2018 due to the lack of seed resulting from the extreme drought (Supplementary Figure 1.1).

Head weight, Spikelets Spike^{-1}, Florets Spikelet^{-1}, Spikelet Density and Stem Diameter

The total number of spikes in a ten-spike sample was verified and recorded to limit calculation errors. In reality this number can vary due to counting errors in the field and small plants with fewer than ten harvestable spikes. After trimming attached stems to 1 cm length, all spikes within a sample were weighed to calculate head weight spike^{-1}. Three representative spikes were chosen from the ten-spike sample and used to determine spikelets spike^{-1}, florets spikelet^{-1} and stem diameter. The number of spikelets was counted on three spikes and averaged. To estimate the number of florets spikelet^{-1}, the florets were counted in three spikelets on one representative spike, one from the bottom, middle and top ⅓ of one spike and the average was calculated. To address the variability in the size of spikelets and florets, a spikelet was defined as having two glumes and at least one complete floret. A floret was defined as having a lemma and palea, and as florets decreased in size towards the apical end of the rachilla axis, those that were less

than half the size of the adjacent floret were considered rudimentary and not counted. Spike length was measured by aligning three spikes end to end along a ruler and dividing the total length by three. Spikelet density was calculated by dividing the average spikelets spike^{-1} by spike length. Stem diameter was calculated by aligning three spikes side by side and measuring the total diameter at 1 cm below the peduncle using a digital calipers. This number was then divided by three.

Yield

All samples were placed back into a 35 °C dryer for at least 24 h before threshing to maintain consistent moisture content, and were threshed using a Wintersteiger LD350 (Reid, Austria) with the following settings: air flow: 3.5; RPM: 1000, 660, 380; concave: 0140-133 2x6 mm. Samples were sieved twice with a 12/64 round sieve (SEEDBURO, Des Plaines, IL). Grain yield (attached hulls included) was recorded. To account for variation in threshability across samples, percent clean (dehulled) seed was rated by comparing the samples to known standards in petri plates. Yield spike^{-1} was calculated by dividing the ten spike yield by the number of spikes in the ten-spike sample. Yield plant^{-1} was obtained by summing ten-spike yield and remaining spikes yield. To account for differences in the weight of seeds with and without hulls, yield was adjusted accordingly using the following equation, which accounts for the weight of hulled seed being 30% hull (lemma and palea; DeHaan et al., 2018):

$$\text{Adjusted yield} = [\text{Yield} * (1 - \% \text{ Clean seed} * 0.70)] + (\text{Yield} * \% \text{ Clean seed})$$

Thousand Grain Weight and Floret Site Utilization

A subsample from the ten spike sample consisting of 50-100 unbroken, de-hulled seeds was counted and weighed on a Marvin Seed Analyzer (GTA Sensorik GmbH, Neubrandenburg, Germany) to determine thousand grain weight (TGW hereafter). The full sample (hulled and dehulled) was then imaged to determine total seed count. Floret site utilization (FSU hereafter) was obtained using the following equation using data from the ten spike sample:

FSU = Total seed count / (Florets spikelet^{-1} * Spikelets spike^{-1} * Spikes in the ten spike sample)

This equation yields FSU in the economical sense, as opposed to biological. The distinction is that economical is what is left after harvesting and cleaning, whereas biological represents the percent filled or partially filled florets (Elgersma 1985).

Data Analysis

All data analyses were conducted in the program R v3.5.2 (R Core Team, 2018). To test for interactions of families by environment, an initial linear mixed effects model was run for each trait which included family, location, and year as fixed effects, and genet, the interactions of location by rep, and location by rep by family by genet, as random effects. This is essentially a split-plot in time design considering data collection in a single environment over two years was a repeated measure (Gomez and Gomez 1984). Results showed highly significant interactions between family and location, family and year, and location and year (Supplementary Table 1.1). These interactions, differences in plant age between 2017 and 2018, and major differences in climate across the two locations, warranted the separate analysis of all environments (year by location combinations; n = 4). Linear models were fit within each environment for each response variable with fixed terms for genet nested within family and block (n = 2). Estimated marginal means were calculated for each genet and trait using 'emmeans' (Lenth et al., 2019). Broad-sense heritability was calculated using a completely random model on a genet mean basis using the following formula:

$$H = \sigma_g^2 \,/\, \left(\sigma_g^2 \,+\, \sigma_e^2 \,/\, r\right)$$

Where σ_g^2 is the variance of genet (progeny only; parents were excluded), σ_e^2 is the error variance and r is the number of reps or blocks (n = 2). To calculate narrow sense heritability, a separate linear model was fit with family and rep as random effects.

Narrow sense heritability is additive genetic variance divided by phenotypic variance and was calculated for half-sib families sharing a common parent using the following equation (Falconer & Mackay, 1996):

$$h^2 = (4 * \sigma_F^2) / [(4 * \sigma_F^2) + (\sigma_e^2)]$$

Where σ_F^2 is the variance of family and σ_e^2 is the error variance. Pearson product moment correlations were calculated between all variables using 'Hmisc' (Harrell & Dupont, 2019). Emmeans for each genet within each environment were then standardized to have a mean of zero and a standard deviation of 1 (Hill et al., 2017; Li et al., 2019). SEM using these values was done in 'lavaan' (Rosseel et al., 2018), and figures were constructed in 'ggplot2' (Wickham, 2016).

Development of the Initial Model

Structural equation modeling requires an initial path diagram that hypothesizes causal relationships among variables in a dataset and is based on knowledge, experience and the literature (Grace, 2006). Variables can be latent (not measured per se, but approximated using measured variables; a concept, for example) and/or observed. Observed variables are preferred in the natural sciences (Lamb et al., 2011) and were the only form included in this analysis. The initial model was designed with the assumption and understanding that yield components develop sequentially, and that their effects on one another are unidirectional (Dofing and Knight 1992), meaning that later developing components do not influence earlier developing components (Abel et al., 2017).

Considering this is the first extensive investigation of yield components in IWG spaced plants, the initial model was constructed based on the available literature on IWG, and yield components literature on other forage and grain crops. In the model, yield plant^{-1} is shown as a function of the number of reproductive tillers and the yield spike^{-1} (Elgersma, 1990c; Studer et al., 2008; Figure 1.1). Yield spike^{-1} is a function of grain size (as represented by TGW), the number of spikelets spike^{-1} and florets spikelet^{-1} and their fertility (FSU) (Elgersma, 1990c, Philipp et al., 2018). In accordance with the unidirectional influence of yield components, since spikelets spike^{-1} and florets spikelet^{-1}

are determined during spike differentiation, which is the earliest occurring of all the yield components tested (Bonnett, 1936), they have no paths leading to them.

Here we include spikelet density, or the number of spikelets per cm length of spike. Spike length can be highly variable in IWG, has been shown to be positively correlated with yield in IWG (Bajgain et al., 2019a), and adjusting for the number of spikelets allows for the assessment of spike length on yield without confounding with opportunities to produce seeds. The spike itself has been shown to be an important source of water-soluble carbohydrates to developing seeds in perennial ryegrass (Trethewey & Rolston, 2009). In wheat, the dry weight of the spike at anthesis is indicative of a long stem elongation phase, which is often associated with an increase in seed set (Gonzales et al, 2011; Pedro et al., 2012; Steinfort et al., 2017). The stem has been shown to be important in perennial ryegrass for translocation to grains during seed fill (Colvill and Marshall 1984). Flag leaves, both for their ability to assimilate and translocate carbon to developing seeds, have in some cases been shown to have a positive influence on seed yield (e.g. in meadow fescue; Fang et al., 2004; and in wheat; Blake et al., 2007), but this has not been the case in ryegrass (Trethewey & Rolston, 2008; Armstead et al., 2009). Plant height and biomass were shown to be positively correlated with plant yield in IWG (Zhang et al., 2016), which contradicts data from modern wheat improvement, where increases in harvest index and GA-sensitive dwarfing genes that reduce height have played a major role in increasing grain yield (Calderini et al., 1994; Rebetzke & Richards, 2000). We sought to parse out these relationships by testing height and while we did not measure total biomass, proxies such as reproductive tiller number, stem diameter, flag leaf area, and reproductive tiller count may provide insight.

Heading date and flowering time are important for adaptation of cultivars to different environments and play a role in optimizing the timing of anthesis and determining yield. Early flowering can reduce spikelet number, but it is also a way to reduce temperature stress post-anthesis (Guo et al., 2018; Liu et al., 2014). There is some evidence in wheat that early flowering is not related to fertility, but findings in ryegrass suggest it may play a role (Guo et al., 2016; Armstead et al., 2008). The relationships as

described and shown in Figure 1.1 were tested in this population for each of the four environments.

Model Fitting

The assessment of model fit in SEM is typically conducted using a Chisq-test, where an insignificant p-value (alpha = 0.05) indicates an acceptable model (Grace, 2006). However, models with large sample sizes (greater than 150) tend to exhibit significant p-values even when the absolute deviations between the fitted model and the covariances of the dataset are small (Grace, 2006). In this case, where sample sizes were exceptionally high (ranging from 552-1162), it is recommended to consider multiple alternative indices (Grace, 2006). First, the Root Mean Square Error of Approximation (RMSEA), which is adjusted for sample size, is sensitive to the number of parameters in a model, and is associated with a 90% confidence interval (Grace, 2006; Hooper et al., 2008). Recommended cutoffs for RMSEA vary, but it is generally regarded that a value below 0.07 is considered a good fit, with a 90% confidence interval ranging between 0 and 0.08 (Hooper et al., 2008). The Comparative Fit Index (CFI) is another commonly reported measure of model fit because it is of those least affected by sample size (Fan et al., 1999). A CFI above 0.95 generally indicates a good model fit (Hu and Bentler, 1999).

Modification indices were evaluated when the model was shown to have a poor fit with the dataset according to the fit indices described above. Paths with the largest "mi" value (estimate of model improvement) were considered for addition to the model if they logically fit the criteria for sequential development and influence of yield components. For example, TGW cannot predict florets spike^{-1}, because florets spike^{-1} is determined before TGW. Suggested modifications that did not meet this criteria were ignored. Once fit indices reached close to their recommended values, additional paths, especially those with low mi were not added in an effort to avoid overfitting (Grace, 2006).

Assessing Alternative Measures of Yield Components and Reducing Sample Size for Spikelets Spike^{-1} and Florets Spikelet^{-1}

A variety of proxies were calculated and sub-sampling was conducted to evaluate alternative, time-saving methods for collecting data on some of the more intensive yield component traits. All proxies and sub-samplings were correlated to the original, raw data (not the 'emmeans') using Pearson correlation analyses. The original dataset values are estimates and subsamples themselves, but they are the best available for determining whether sample sizes could be reduced. Fertility index was calculated by dividing the weight of seed yield spike^{-1} by head weight spike^{-1}. Seed count was estimated by dividing yield spike^{-1} by the average weight of a seed determined from the TGW estimates. Spikelets spike^{-1} was initially counted on three spikes per spaced plant. The dataset was subsampled 100 times to randomly select the spikelet counts from two (from which the average was taken) and one spike. Florets spike^{-1} was initially determined as the average count of three spikelets (one from the bottom, middle and top one third). On average across environments, florets spikelet^{-1} was highest in the middle floret, and spikelets on the bottom and top had 8.6 and 12.2% fewer florets. An estimated floret spikelet^{-1} was calculated using data from only the middle spikelet using the following equation:

$$Florets_{middle} + (Florets_{middle} * 1 - 0.086) + (Florets_{middle} * 1 - 0.122) / 3$$

Finally, all estimated values were used to calculate an estimated floret site utilization using the equation described previously.

Results

Temperature and Precipitation

Temperature variations and precipitation accumulation across environments are shown in Supplementary Figure 1.1. Notable deviations from the ten year average include especially high rainfall in 2017 at TLI, followed by a severe drought in 2018 that began towards the end of 2017 and lasted through anthesis and harvest in 2018, with a maximum deficit of 21.7 cm accumulated precipitation from the ten-year average just after anthesis. This drought was accompanied by higher than average temperatures beginning before and lasting through anthesis. St. Paul experienced normal temperatures

and precipitation accumulation in 2017 and 2018, with the exception of a relatively cool spring in 2018.

Location and Year Effects on Overall Performance

Initial combined analysis of variance revealed significant variation across environments and significant genet by environment and family by environment interactions ($\alpha = 0.001$) within each location for each trait (Supplementary Table 1.1). The presence of significant interactions provide further support for the analysis of environments on an individual basis, results for which are presented in Supplementary Table 1.2. Tukey's HSD calculated within the 'emmeans' package separated all environments into different groups for all traits with the exception of reproductive tiller number and yield plant^{-1} (Figure 1.2; values in Supplementary Table 1.3).

All references to yield herein are grain yield and not forage. On average, in 2017 across locations, spaced plants yielded similarly on a yield plant^{-1} basis, but plants at TLI exhibited 23% lower yield spike^{-1} , lower TGW, but achieved similar yields plant^{-1} to St. Paul grown plants via a higher number of reproductive tillers and more spikelets spike^{-1} and florets spikelet^{-1} (Figure 1.2). Spaced plants at TLI had larger flag leaves and stems and more spikelets spike^{-1}, florets spikelet^{-1}, and lower spikelet density (more space between spikelets), but overall were shorter in height.

In 2018 at STP, on average, yield plant^{-1} nearly doubled (90% increase) along with reproductive tiller counts (126% increase) but yield spike^{-1} decreased by 23%. Over 92% of individuals exhibited a decrease in yield spike^{-1} in the second year. Plants generally decreased in size with regards to flag leaf area, stem diameter, thousand grain weight, and had fewer florets spikelet^{-1} and spikelets spike^{-1}, resulting in a 27% increase in FSU. During the drought in 2018 at TLI, all yield components decreased dramatically. Yield spike^{-1} decreased 97%, as evidenced by low FSU (average 3%) and when seeds were present they weighed 30% less on average than the previous year. Flag leaves, stems, height and reproductive tiller counts were all substantially lower, and the spike emergence percent showed the largest range at a given time compared with all other environments (0 - 1.9% emerged). At harvest in 2018, 27% of spaced plants were either

dead or deemed unharvestable due to lack of spikes, and 73% of those that were harvested yielded less than 0.025 g spike^{-1}, resulting in an excess of missing data for that environment.

Bivariate Correlations

Trends in correlations between traits were consistent across environments. All bivariate correlations described here are significant at the $\alpha = 0.001$ level. Reproductive tiller ct. showed the strongest positive correlation with yield plant^{-1} in all environments (avg. $r = 0.66$), followed by yield spike^{-1} (0.58), height (0.40), spikelets spike^{-1} (0.37), and FSU (0.37) (Supplementary Tables 1.4 – 1.7). FSU showed the strongest correlation with yield spike^{-1} (avg. $r = 0.66$), followed by spikelets spike^{-1} (avg. $r = 0.38$). Flag leaf area, height, stem diameter and florets spikelet^{-1} were more strongly correlated with yield spike^{-1} at STP compared with TLI (Supplementary Tables 1.4 – 1.7). Significant positive correlations existed between biomass traits such as height, tillering, flag leaf area, and stem diameter. These persisted in all environments, though flag leaves were more independent of other biomass traits at TLI. Spike emergence percent and Feekes maturity were positively correlated in all environments (avg. $r = 0.64$).

Heritability

Average across environments, all broad sense heritabilities were high with a narrow range (H = 0.48 - 0.82) and suggesting high repeatability across genets within environments (Figure 1.2; Supplementary Table 1.8). Narrow sense heritabilities were more variable ($h^2 = 0.18$ - 0.60), with spikelet density, maturity, height and TGW exhibiting the highest values (avg. h^2 ranging from 0.89 - 0.93). Yield spike^{-1}, flag leaf area, FSU, and yield plant^{-1} had low average narrow sense heritabilities (avg. h^2 ranging from 0.18 - 0.20).

Structural Equation Modeling

The initial model required modification in all environments except TLI 2018 where it had an acceptable fit (Table 1.1). For STP 2017 and 2018, modification indices

suggested a direct path from florets spikelet^{-1} to FSU, resulting in adequate fit indices. For TLI 2017, two paths were added, one from florets spikelet^{-1} to FSU and another from height to YPS. Variance explained (r^2) by the model for the endogenous variables is shown in Table 1.2.

The direct and indirect effects of each yield component trait on yield plant^{-1} and yield spike^{-1} are shown in Figure 1.3 (exact values can be found in Supplementary Tables 1.9-1.12). Overall trends were consistent across environments, with minor differences between effect sizes. For yield spike^{-1}, FSU was consistently the most positively contributing factor in all environments with $\sigma_{direct} = 0.70$ at STP in 2017 and 2018, and 0.84, 0.90 at TLI. Following FSU was spikelets spike^{-1}, florets spikelet^{-1} and lastly TGW. Florets spikelet^{-1} had a small negative indirect effect via a reduction in FSU at STP 2017, 2018 and TLI 2017 ($\sigma_{indirect} = -0.18, -0.32, -0.22$). For maturity traits, emergence percent was only significant at STP 2018 with a small effect ($\sigma_{indirect} = 0.13$), and Feekes flowering stage only at STP 2018 with a very minor effect ($\sigma_{indirect} = 0.08$) and at TLI 2018, where it had a larger effect ($\sigma_{indirect} = 0.33$) mainly via its positive effect on FSU (Supplementary Tables 1.9 – 1.12). For biomass traits, spikelet density had a consistent, negative effect in all environments except TLI 2018. This effect was fairly substantial relative to other indirect effects ($\sigma_{indirect} = -0.15, -0.22, -0.16$). Stem diameter was only significant in STP environments with increasing importance in the second year ($\sigma_{indirect} =$ 0.10 and 0.22). Flag leaf area followed a similar trend in STP ($\sigma_{indirect} = 0.10$ and 0.17), but had a negative influence at TLI in 2018 ($\sigma_{indirect} = -0.15$). Reproductive tiller number had a significant but small indirect effect ($\sigma_{indirect} = 0.09$) on yield spike^{-1} only at STP 2017. There was a significant but small direct effect of reproductive tillers on TGW and FSU ($\sigma_{direct} = 0.07$ and 0.08, respectively) at STP 2017. During the drought at TLI 2018, reproductive tillers had a significant negative effect on TGW ($\sigma_{direct} = -0.13$), but its overall indirect effect on yield spike^{-1} was not significant.

When yield was measured on a yield plant^{-1} basis, reproductive tiller number consistently had the greatest direct effect ($\sigma_{direct} = 0.59$, 0.63, 0.57), followed by yield spike^{-1}. The effects of spikelets spike^{-1}, florets spikelet^{-1} and TGW were less for yield

plant^{-1} than they were with yield spike^{-1}. Maturity and biomass traits follow a similar trend, though effect sizes decrease in every case.

Discussion

Gains from selection have been realized for increasing yield in intermediate wheatgrass over the last decade using recurrent and genomic selection methods, primarily focusing on yield spike^{-1} and TGW (DeHaan et al., 2018; Zhang et al., 2017). Despite these improvements, IWG faces several challenges. For one, IWG grain yields and seed sizes remain relatively low. For example, yields of IWG are reported to range between 625 to 1110 kg ha^{-1} and seed size averages 6.8 - 9.3 mg seed^{-1}, depending on the environment (Tautges et al., 2018; Bajgain et al., 2019a). Furthermore, yields of IWG are also known to decline after the first year of production, which puts the long-term profitability of the system at risk. These challenges highlight the need for continued breeding and selection in IWG to improve sustained grain yields. Yield is a highly complex trait, and understanding the underlying relationships among component traits is critical to selecting parental material for population improvement (Grafius, 1956). Furthermore, identifying component traits with high heritability and with strong relationships with yield can better inform selection by combining both direct and indirect selection methods (Falconer and Mackay, 1996).

Performance Across Years and Environments

Several mechanisms have been suggested in the literature to explain yield decline in IWG, but the issue has not yet been resolved. Considering the severe drought at TLI in 2018 and the relative consistency of the growing conditions in STP 2017 and 2018, STP provides the best opportunity to examine changes in performance due in part to plant age. Interestingly, in this spaced plant environment, where plants are on 1-m centers and cultivation between plants is necessary to maintain plant identity, yield plant^{-1} increased by 90% in the second year, a finding that is supported by other studies (Zhang et al., 2016; Cattani et al., 2018; Bajgain et al., 2019a). This increase was also accompanied by near doubling of reproductive tiller number, and a decrease in almost all other yield

component and biomass traits (yield spike^{-1}, thousand grain weight, spikelets spike^{-1} and florets spikelet^{-1}, flag leaf area and stem diameter). Larson et al. (2019) also reports the performance of IWG in spaced plants over two and three years in two locations. While they do not report yield plant^{-1}, similar decreases over time in yield spike^{-1}, seed mass, seed dimensions, florets spikelet^{-1} and stem width were observed, though whether this decrease occurred between the second or third year was variable depending on the trait. Variations in temperature can affect the regulation and length of developmental phases in wheat and ryegrass, and in turn affect yield component traits, specifically spikelets spike^{-1} (Rawson and Richards, 1993; Ryle 1965). Thus the effects of plant age and environment cannot be unequivocally resolved in this case. However, evidence suggests that future investigations of yield decline in IWG should consider reductions in component traits, specifically on yield on a spike^{-1} basis.

Decreases in yield spike^{-1} in the second year, in concert with a near doubling of reproductive tillers, suggests that increases in yield plant^{-1} were driven primarily by reproductive tillers. The SEM model supports this finding, wherein reproductive tiller count was the primary driver of yield measured in this way in all environments (Figure 1.3; Supplementary Tables 1.9 – 1.12). Cattani et al. (2018) also demonstrated the importance of the contribution of reproductive tillers per unit area, along with plant biomass, to IWG yield in spaced plants. Concerns about sod-bounding and reproductive tiller suppression in seed production sward environments of IWG may suggest that breeders should select against plants that tiller aggressively, and in favor of plants that instead maximize yield spike^{-1} with fewer reproductive tillers. SEM results suggest that if breeders are selecting on yield plant^{-1}, reproductive tiller number would be the primary trait under indirect selection. Yield plant^{-1} in IWG has shown low stability and predictive ability across years (Cattani, 2016), though it can be improved to some degree by adjusting for the area occupied by a spaced plant (Cattani et al., 2018). Furthermore, there is substantial evidence in the forage breeding literature that selection for forage and grain yield plant^{-1} in a spaced plant environment has low predictive ability in swards (e.g. Lazenby and Rogers, 1964; Asay and Johnson 1997; Waldron et al., 2008); however, reports of whether this discrepancy occurs in IWG have not been published and requires

further study. Furthermore, in the SEM analysis, the indirect effects of critical components of seed yield, such as FSU, opportunities to produce seeds and TGW, diminish when considered on a yield plant^{-1} basis. This poses a risk in that selection on yield plant^{-1} is more likely to miss individuals that exhibit high values for these traits and may limit gain from selection in future cycles. As an example, in a study of selection methods for yield in wheat, yield spike^{-1} or yield spike^{-1} and yield, showed less variability across years than selection for yield alone (Alonso et al., 2018). Interestingly, the indirect effect of reproductive tiller number on yield spike^{-1} via TGW and FSU, was only significant in STP 2017, and the effect while positive, was negligible. The bivariate correlation analysis supports this finding, with an average r across environments of 0.20. This suggests there is limited evidence for a tradeoff between excessive reproductive tillering and yield spike^{-1}, and highlights the importance of determining whether excessive tillering is advantageous for yield in a sward environment and if the tendency of a genotype to tiller is stable across growing environments.

Relationships Between Yield Component Traits

Simple bivariate correlations highlighted the importance of many yield component traits. For example, when yield was measured on a per spike^{-1} basis, FSU, spikelets spike^{-1}, florets spikelet^{-1}, flag leaf area, height, stem diameter, Feekes flowering time and TGW all had a significant, positive correlation coefficient over $r = 0.30$ in either STP or TLI on average across years, and almost all correlations between traits were significant (Supplementary Tables 1.4 – 1.7). While FSU clearly holds the strongest correlation with yield spike^{-1}; in some environments, other traits closely follow. Thus, the results suggest that all the aforementioned traits may contribute to yield, and perhaps that IWG breeders should continue to measure and consider all in selection decisions in their efforts to increase yield. Correlations measured in this way, while informative, fail to parse out direct and indirect effects, which in some cases may inflate a correlation coefficient and overemphasize its importance, or highlight a relationship that lacks a logical interpretation for breeding. As an example, height is positively correlated with yield in IWG, but height does not directly contribute to yield. The more likely case is that

taller plants tend to have more opportunities to produce seeds (longer spikes may equal more spikelets), or more biomass to support larger seeds, or more seeds. The SEM model, which was constructed *a-priori* using logic and research of yield components on IWG and other grass species, more accurately defines the unidirectional pathways in which traits may influence yield (Figure 1.1), and effectively distinguishes between direct and indirect effects. Results from this approach more clearly highlighted FSU and opportunities to produce seeds (spikelets spike^{-1} and florets spikelet^{-1}) and TGW as the main drivers for yield spike^{-1}, and, depending on the environment, diminishes the effects of biomass traits (flag leaf area, stem diameter and height) and maturity traits (spike emergence and Feekes flowering time) whose indirect effects even become insignificant in some environments. In the case of height, using SEM, it becomes clear that its effect on yield per spike^{-1} is relatively small, and in cases where it remains important, it is primarily associated with an increase in TGW which in turn affects yield (Figure 1.3; Supplementary Tables 1.9 – 1.12).

When considering yield spike^{-1}, results from the SEM analysis indicate that FSU, opportunities to produce seeds and TGW are the main contributors and may serve as potential selection targets. Due to the variability in indirect effects of biomass and maturity across environments, no single trait stands out as a potential selection target for indirect selection on yield spike^{-1}. Stem diameter and flag leaf area appear to offer a marginal benefit in both years at STP, but not at TLI. Though their effects were small, they were positive. Since these traits were both easy to measure, they may be worthwhile to continue monitoring in the breeding program. While spike emergence and flowering time were correlated (avg. $r = 0.64$), their effect on yield appears to differ. Early spike emergence and flowering time offered a small benefit at STP in 2018, but these effects were insignificant in 2017 (Figure 1.3; Supplementary Table 1.9 – 1.12). Flowering time at TLI in 2018 during the drought was beneficial in that earlier flowering plants (but not necessarily early emerging) tended to have higher FSU (Supplementary Table 1.9 – 1.12). Early flowering has been a target of selection for Kansas environments as a way to decrease post-anthesis heat and drought stress (Lee DeHaan, personal communication). The results support this approach, as the indirect effect of flowering time in 2018 even

exceeded direct effects of opportunities to produce seeds and TGW (Figure 1.3). Adequate accumulated precipitation or first year growth habit likely masked this trend in 2017 (Supplementary Figure 1.1).

The consistent and large influence of FSU on yield in IWG is not surprising, as FSU has been referred to as a limiting factor for yield in IWG, ryegrass, as have similar traits in wheat, like grains spike^{-1} and spike fertility (Knowles, 1977; Elgersma and Śnieżko, 1988; Philipp et al., 2018). It is common for perennial species to utilize only a percentage of their opportunities to produce seeds (Wiens, 1984). FSU is a complex trait and can be influenced by a myriad of things including but not limited to: pollen availability (both spatial and temporal), resource competition, self-incompatibility, seed abortion, seed shattering, and buildup of mutational load due to outcrossing (Armstead et al., 2008; Charlesworth 1989; Charlesworth and Charlesworth 1987). In this study, the average FSU in an environment ranged from 37-60% (excluding TLI 2018 which was 3%). Larson et al. (2019) reported even lower levels of FSU in IWG (referred to as seeds per floret) at 12 - 22%, depending on the environment. FSU rates in perennial ryegrass are also highly variable, and can range from 8-73% (Elgersma and Sniezko 1988). In a QTL mapping study of seed set in perennial ryegrass, two genomic regions were identified - one as a recessive mutation that results in near-sterility, and a second that is partially recessive and results in a less dramatic reduction (Armstead et al., 2008). In IWG, a recent GWAS aligned seed size, fertility and seed-yield related QTLs with a self-incompatibility (Z) candidate gene (Larson et al., 2019). Furthermore, the relatively low r^2 value in the SEM models for FSU would indicate that the biomass and maturity variables included explain only a small amount of the variation for FSU and that perhaps genotype or environmental variation play a larger role (Table 1.2). Broad sense heritability estimates for FSU was moderate (avg H = 0.60), and narrow sense was low (avg h^2 = 0.18), suggesting that while this trait is repeatable across genets, the additive variance for this trait is relatively low. FSU is a complex measurement which relies on assumptions that small samples taken for spikelet and floret counts are representative of the entire sample. This may explain why in some cases FSU can exceed 1 in cases where

the spike chosen for counting may have been poorly representative of the sample (Figure 1.2).

It is not clear whether the negative influence that floret number has on FSU, as shown in all environments except TLI 2018, is due to some biological relationship, or whether it is simply a function of how floret site utilization is calculated. Interestingly, there is a positive, but relatively weak correlation (avg $r = 0.29$; data not shown) between florets spikelet^{-1} and seeds spikelet^{-1}, suggesting that more florets does not always result in more seeds (or at least economical seeds). SEM results would indicate, however, that florets spikelet^{-1} has a large direct effect on yield, but if one was to consider the total effect of florets (summing both direct and indirect), its influence on seed yield would decrease to less than that of spikelets spike^{-1} and TGW (Figure 1.3).

Opportunities to produce seeds are plentiful in IWG, but relatively few are utilized. In the case of wheat, yield gains throughout improvement have been made via increased grain number and yield spike^{-1} and spikelet^{-1} (Philipp et al., 2018). Most often in wheat, which also has an indeterminate floret meristem, florets that survive and are fully developed by the time of heading are typically fertile (Sreenivasulu and Schnurbusch 2012), whereas in IWG, this is not the case. In ryegrass, 90% of unused florets were the result of ovule degeneration occurring within ten days after anthesis (Elgersma and Snieżko, 1988; Boelt and Studer, 2010). A similar study in IWG may provide necessary insight into the cause of floret infertility.

Spikelet density had a consistent, negative influence on yield across all environments except TLI 2018 (Figure 1.3), primarily via a large negative effect on TGW (avg. $\sigma_{direct} = -0.29$) and less so on FSU (avg. $\sigma_{direct} = -0.09$; Supplementary Tables 1.9 – 1.12). This may indicate that the effect is more likely a resource allocation issue and perhaps not related to pollen obstruction, which may arise due from having a compact spike.

The importance of FSU on yield spike^{-1} would suggest that it should be a major target of selection in IWG. In a study of selection for yield in wheat, a two-step selection process including selection on individuals with high yield and then on the genotypes with highest spike fertility resulted in greater gains compared with selection on yield alone

(Alonso et al., 2018). In the present study, yield spike^{-1} exhibited a high broad sense heritability (avg. H = 0.75), but had the second to lowest narrow sense heritability (h^2 = 0.19) of all traits, suggesting that results were repeatable across blocks within locations, but that the contribution of additive variance is limited. Combining this trait with FSU in a two-step selection method as was demonstrated by Alonso et al. (2018) might be a promising way to identify high yielding individuals that also efficiently utilize their floret sites. However, due to the low heritability of both yield spike^{-1} and FSU, gains from selection may be limited or slow. Breeders may also choose to select for genotypes that maximize yield spike^{-1} and FSU, while limiting the number of reproductive tillers and specifically the change in reproductive tiller number from year 1 to year 2. Furthermore, it will be important to monitor these traits over years and in seed production sward environments to address the issue of yield persistence in IWG.

Phenotyping Optimization

The methods used to collect FSU data were tedious and labor-intensive. Assuming the estimate for FSU that was obtained for this analysis was "correct", or the best available, we tested fertility index, a commonly used proxy for FSU, and subsampling from the raw, unadjusted dataset to determine whether this important trait could be measured with fewer labor inputs. Fertility index, or the ratio of grain yield spike^{-1} to head weight spike^{-1}, involves weighing before and after threshing (eliminating the need for any counting), was a moderate predictor of FSU (r = 0.54 - 0.96, depending on the environment; Supplementary Figure 1.3). Fertility index appears to plateau at ~60 %, and genets with this ratio exhibit a high degree of variability for FSU, suggesting that selection on this trait alone may miss individuals with high FSU. This is because fertility index is driven by three principal variables: headweight, FSU, and seed weight. Because FSU was measured in the economic sense and grain weight was variable, it is very likely that selecting on fertility index would not select the best individuals for FSU; however, it would likely screen out poor performing individuals. Fertility index is also driven by compactness of spike. Therefore, holding seed number and weight constant,

two spikes may have drastically different fertility indices based on the size and compactness of seed bearing structures.

Considering FSU's unique importance for yield, we then tested to see whether this datapoint could be approximated using a smaller sample size for spikelets spike^{-1} and florets spikelet^{-1}. We suggest that reducing to two spikes for spikelets spike^{-1} was a strong predictor of spikelets spike (r = 0.95 - 0.96, depending on the environment) and one may less adequately capture the variation with a single spike (r = 0.85 - 0.88; Supplementary Figures 1.4 & 1.5). For florets spikelet^{-1}, we observed consistently fewer in the bottom (~8% fewer) and top (~12% fewer) thirds of the spike relative to the middle (Supplementary Figure 1.6). Counting florets within just a single spikelet (middle) and adjusting for this decrease was a good predictor of florets spikelet^{-1} (r = 0.83 - 0.93) (Supplementary Figure 1.7). Interestingly, variability for spikelet spike^{-1} in a sample and in florets spikelet^{-1} in the different sections of the spike both decreased in the second year at both locations, suggesting that lowering sample sizes may be more appropriate in the second year, though this may require validation in different environments.

As an alternative to image analysis or a seed counter to count seeds, one can divide the grain yield spike^{-1} by the TGW. This proxy was highly predictive (r = 0.96 - 0.99) of the actual count (data not shown). Note that grain weight was adjusted for threshability (see Methods), which may be especially important in a cases where genotypes exhibit low threshability ratings and much of the yield mass is attributed to hulls. While it tended to underestimate the seed count (91% of the time), 93% of samples were within 20% of their observed values. Using this proxy, and a reduced sample size for spikelets spike^{-1} (n = 2) and florets spikelet^{-1} (n = 1), we predicted FSU with adequate accuracy (r = 0.88 - 0.98) (Supplementary Figure 1.8). This method would be a significant time savings for anyone seeking to measure FSU in intermediate wheatgrass spaced plants. Finally, we also tested whether yield plant^{-1} could be predicted by multiplying yield spike^{-1} and reproductive tiller number. This was an accurate predictor (r = 0.86 - 0.91) and would be a significant time and storage savings, but the predictive ability decreased as yield increased (Supplementary Figure 1.9). It is important to remember that while this population exhibited a high amount of phenotypic variation, the

individuals were all half- or full-sibs, and the same accuracy may not be observed in a more diverse population or in different environments.

Tables and Figures

Table 1.1. Fit indices of the modified structural equation models for each location and year.

Model	n[a]	RMSEA (CI)[b]	CFI[c]	Chisq-Test[d]
STP 2017	1008	0.072 (0.060 - 0.084)	0.98	$\chi^2 = 130$; df =21, $P = 0.000$
STP 2018	1135	0.059 (0.048 - 0.071)	0.98	$\chi^2 = 105$; df = 21, $P = 0.000$
TLI 2017	1162	0.069 (0.058 - 0.081)	0.98	$\chi^2 = 125$; df = 19, $P = 0.000$
TLI 2018	553	0.050 (0.027 - 0.074)	0.99	$\chi^2 = 29$; df = 12, $P = 0.004$

[a] n, number of complete observations used in the model
[b] RMSEA, root mean square of approximation followed by a 90% confidence interval in parentheses
[c] CFI, comparative fit index
[d] Chisq-Test statistic, followed by degrees of freedom and p-value

Table 1.2. Variance explained by the modified structural equation models for the endogenous variables for each location and year.

Variable	STP 2017	STP 2018	TLI 2017	TLI 2018
Thousand Grain Weight	31.2	33.2	26.1	18.6
Floret Site Utilization	14.0	29.5	11.9	17.7
Yield Spike^{-1}	92.9	88.4	90.6	91.2
Yield Plant^{-1}	70.2	60.6	64.6	-

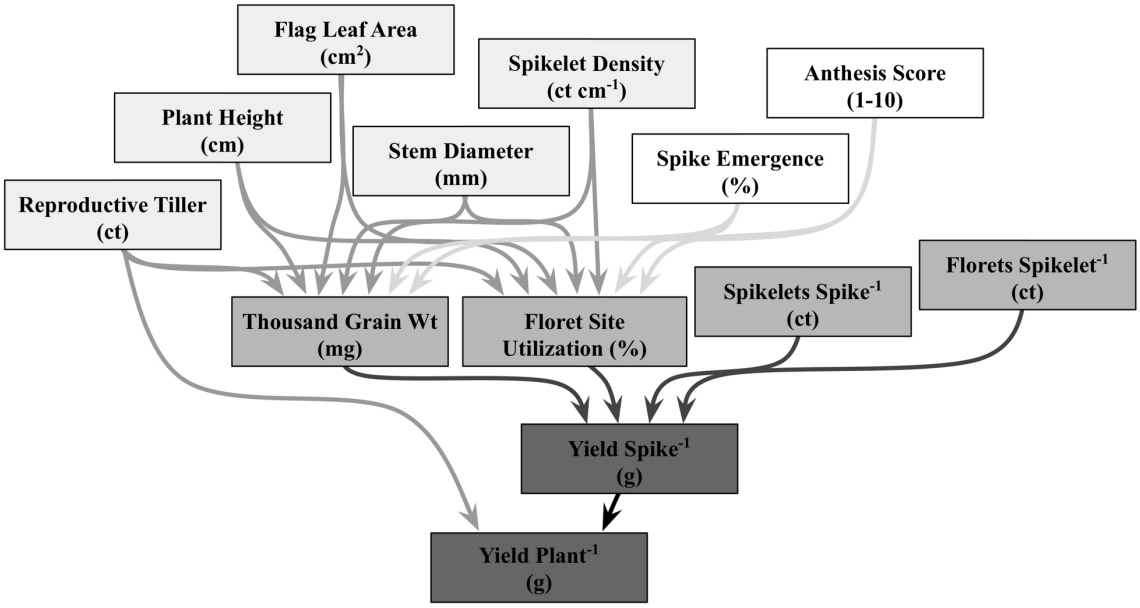


Figure 1.1. Initial model for the relationship between yield and component traits in intermediate wheatgrass spaced plants. Groups of traits are colored in different shades of gray, including: biomass traits (light gray), maturity traits (white), opportunities to produce seeds, seed size and fertility (medium gray), and final measures of yield (dark gray).

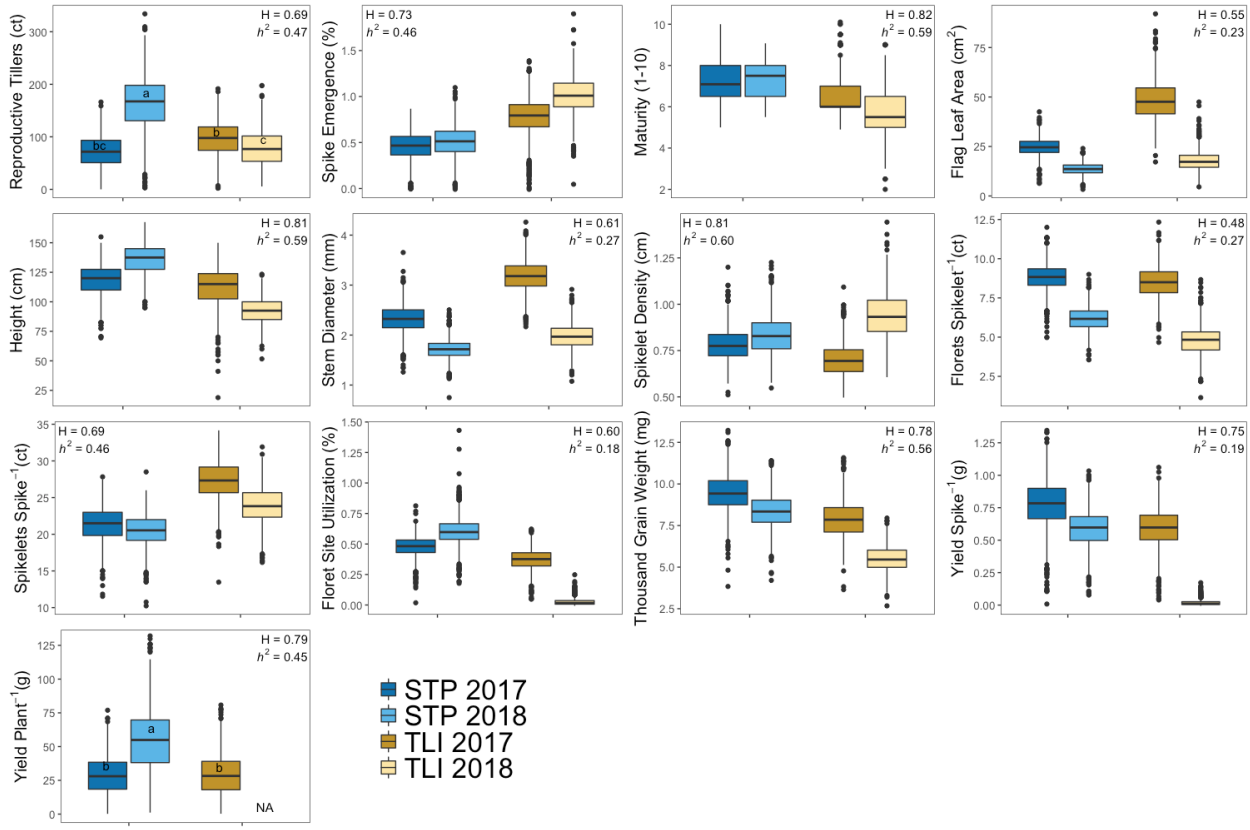


Figure 1.2. Boxplots of estimated marginal means for each genet in each environment for each trait. Blue is St. Paul, MN (STP) in 2017 and 2018, yellow is Salina, KS at The Land Institute (TLI) in 2017 and 2018. All environments were significantly different from each other according to Tukey's HSD except reproductive tillers and yield plant^{-1} where means separations are shown and environments with different letters are significantly different ($\alpha = 0.05$). Average heritability estimates, both broad (H) and narrow sense (h^2), across all four environments are shown in the top corner.

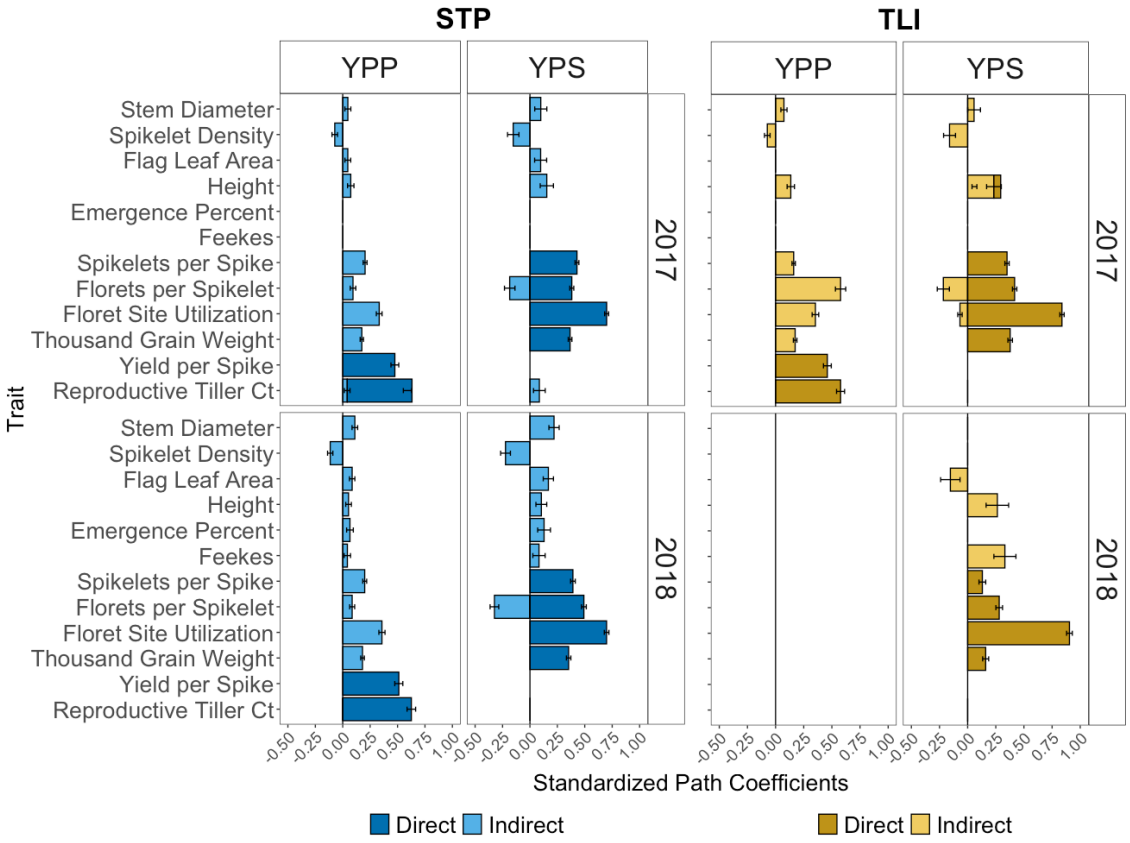


Figure 1.3. Significant ($\alpha = 0.05$) direct and indirect standardized path coefficients for each yield component trait as they contribute to yield per plant (YPP) and yield per spike (YPS) in intermediate wheatgrass spaced plants in St. Paul, MN (STP) and Salina, KS (TLI) in 2017 and 2018. Insignificant paths are excluded. The traits Yield per Spike and Reproductive Tiller Ct are not applicable to YPS as no paths connect these traits. Error bars indicate 95% confidence intervals.

Supplementary Materials

Supplementary Table 1.1. Analysis of variance tables for each trait across all four environments included to demonstrate significant family by location by year interactions as justification for the analysis of all environments individually.

Trait

Reproductive Tillers (ct)

Source	Sum Sq[a]	Mean Sq[b]	NumDF[c]	DenDF[d]	F value	Pr(>F)
famID	1163854.36	129317.15	9	4517	127.60	1.13E-214
loc	6707.19	6707.19	1	2	6.62	0.123909669
year	3342022.30	3342022.30	1	4574	3297.75	0
famID:loc	178768.75	19863.19	9	4484	19.60	1.36E-32
famID:year	124315.63	13812.85	9	4573	13.63	7.79E-22
loc:year	6730699.83	6730699.83	1	4574	6641.54	0
famID:loc:year	411620.14	45735.57	9	4573	45.13	2.47E-78

Spike Emergence (%)

Source	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
famID	18.97	2.11	9	4486	100.87	2.74E-172
loc	5.10	5.10	1	2	244.10	0.00405584
year	46.66	46.66	1	4478	2233.17	0
famID:loc	6.15	0.68	9	4455	32.70	3.60E-56
famID:year	3.50	0.39	9	4478	18.62	8.04E-31
loc:year	20.20	20.20	1	4478	966.61	2.51E-192
famID:loc:year	1.71	0.19	9	4478	9.09	9.78E-14

Feekes (1-10)

Source	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
famID	682.59	75.84	9	4543	148.66	6.29E-247
loc	89.10	89.10	1	2	174.64	0.00562237
year	247.97	247.97	1	4609	486.02	1.75E-102
famID:loc	406.64	45.18	9	4509	88.56	2.66E-152
famID:year	140.61	15.62	9	4608	30.62	1.67E-52
loc:year	591.27	591.27	1	4609	1158.90	8.59E-227
famID:loc:year	42.73	4.75	9	4608	9.31	3.98E-14

Flag Leaf Area (cm2)

Source	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
famID	16819.68	1868.85	9	4562	51.65	8.88E-90
loc	9292.29	9292.29	1	2	256.83	0.00386574
year	998942.22	998942.22	1	4645	27609.75	0
famID:loc	5910.58	656.73	9	4530	18.15	5.55E-30
famID:year	8225.65	913.96	9	4644	25.26	7.48E-43
loc:year	218209.57	218209.57	1	4645	6031.09	0
famID:loc:year	3449.29	383.25	9	4644	10.59	2.08E-16

Height (cm)

Source	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
famID	156098.97	17344.33	9	4398	216.28	0
loc	38790.05	38790.05	1	2	483.71	0.00205538
year	3628.46	3628.46	1	4464	45.25	1.96E-11
famID:loc	2472.92	274.77	9	4365	3.43	0.0003263
famID:year	28340.21	3148.91	9	4463	39.27	6.93E-68
loc:year	825193.25	825193.25	1	4464	10290.10	0

famID:loc:year	10489.33	1165.48	9	4463	14.53	1.88E-23

Stem Diameter (mm)

Source	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
famID	29.82	3.31	9	4451	58.68	7.60E-102
loc	31.87	31.87	1	2	564.28	0.00176607
year	1830.73	1830.73	1	4461	32418.64	0
famID:loc	3.38	0.38	9	4419	6.66	1.61E-09
famID:year	3.17	0.35	9	4461	6.24	8.32E-09
loc:year	204.92	204.92	1	4461	3628.80	0
famID:loc:year	2.18	0.24	9	4461	4.29	1.45E-05

Spikelet Density (cm)

Source	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
famID	11.29	1.25	9	4537	230.90	0
loc	0.00	0.00	1	2	0.55	0.53489006
year	45.67	45.67	1	4533	8405.97	0
famID:loc	0.56	0.06	9	4512	11.40	7.80E-18
famID:year	0.75	0.08	9	4533	15.34	6.58E-25
loc:year	18.69	18.69	1	4533	3440.63	0
famID:loc:year	0.41	0.05	9	4533	8.35	1.89E-12

Florets Spikelet-1 (ct)

Source	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
famID	540.61	60.07	9	4493	71.50	8.43E-124
loc	144.87	144.87	1	2	172.44	0.00570697
year	22281.41	22281.41	1	4535	26521.40	0
famID:loc	12.55	1.39	9	4461	1.66	0.09294428
famID:year	120.29	13.37	9	4535	15.91	6.19E-26
loc:year	597.91	597.91	1	4535	711.68	9.04E-146
famID:loc:year	24.48	2.72	9	4535	3.24	0.00063113

Spikelets Spike-1 (ct)

Source	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
famID	4852.36	539.15	9	4461	134.34	9.93E-225
loc	6603.10	6603.10	1	2	1645.24	0.00058303
year	9958.46	9958.46	1	4470	2481.27	0
famID:loc	429.21	47.69	9	4426	11.88	1.07E-18
famID:year	529.57	58.84	9	4470	14.66	1.11E-23
loc:year	3880.91	3880.91	1	4470	966.98	2.32E-192
famID:loc:year	371.65	41.29	9	4470	10.29	7.33E-16

Floret Site Utilization (%)

Source	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
famID	1.24	0.14	9	4438	18.88	2.70E-31
loc	5.78	5.78	1	2	792.01	0.00123421
year	24.77	24.77	1	4316	3393.05	0
famID:loc	0.30	0.03	9	4405	4.64	3.91E-06
famID:year	0.58	0.06	9	4319	8.85	2.61E-13
loc:year	115.96	115.96	1	4316	15883.60	0
famID:loc:year	0.27	0.03	9	4319	4.08	3.15E-05

Thousand Grain Weight (mg)

Source	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
famID	637.99	70.89	9	5412	133.97	9.85E-229
loc	394.12	394.12	1	2	744.83	0.00097479
year	3218.17	3218.17	1	3966	6081.92	0
famID:loc	21.50	2.39	9	5378	4.51	6.16E-06
famID:year	41.74	4.64	9	3908	8.77	3.72E-13
loc:year	384.69	384.69	1	3966	727.02	3.55E-147
famID:loc:year	31.47	3.50	9	3908	6.61	2.01E-09

	Source	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
Yield Spike-1 (g)	famID	2.53	0.28	9	4409	23.88	2.65E-40
	loc	3.30	3.30	1	2	280.56	0.0035029
	year	288.07	288.07	1	4042	24495.30	0
	famID:loc	1.13	0.13	9	4384	10.68	1.49E-16
	famID:year	0.89	0.10	9	4046	8.43	1.44E-12
	loc:year	76.49	76.49	1	4042	6503.86	0
	famID:loc:year	0.38	0.04	9	4046	3.62	0.00016199

	Source	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
Yield Plant-1 (g)	famID	65466.25	7274.03	9	2425	58.51	4.32E-97
	loc	127.11	127.11	1	2	1.02	0.43179989
	year	356465.45	356465.45	1	1239	2867.30	0
	famID:loc	1524.78	169.42	9	3068	1.36	0.19934488
	famID:year	11091.77	1232.42	9	1242	9.91	8.56E-15
	loc:year	-	-	-	-	-	-
	famID:loc:year	-	-	-	-	-	-

[a] Sums of squares
[b] Mean squares
[c] Numerator degrees of freedom
[d] Denominator degrees of freedom

Supplementary Table 1.2. Analysis of variance results from the mixed effects linear model within each unique environment for all traits.

Reproductive Tillers

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	2604.8	289.4	216.0	3.00E-239
rep	1	234.1	234.1	174.7	3.24E-37
famID:plantID3	1190	6802.4	5.7	4.3	3.76E-124
Residuals	1141	1529.1	1.3	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	3033.3	337.0	130.7	6.96E-168
rep	1	1062.4	1062.4	411.9	2.97E-78
famID:plantID3	1188	8162.3	6.9	2.7	1.06E-59
Residuals	1123	2896.8	2.6	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	1789.4	198.8	133.5	1.25E-170
rep	1	122.0	122.0	81.9	6.21E-19
famID:plantID3	1193	5465.5	4.6	3.1	4.39E-77
Residuals	1125	1675.9	1.5	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	987.7	109.7	27.2	9.20E-43
rep	1	215.9	215.9	53.6	4.84E-13
famID:plantID3	1188	7700.6	6.5	1.6	1.02E-15
Residuals	1087	4380.8	4.0	-	-

Spike Emergence

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	12.4	1.4	97.2	2.95E-133
rep	1	0.0	0.0	3.5	0.061401547
famID:plantID3	1171	42.0	0.0	2.5	3.09E-53
Residuals	1103	15.6	0.0	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	19.1	2.1	155.3	6.97E-190
rep	1	0.2	0.2	12.0	0.000540912
famID:plantID3	1186	46.7	0.0	2.9	1.77E-68
Residuals	1117	15.3	0.0	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	10.9	1.2	64.7	2.05E-95
rep	1	0.2	0.2	13.2	0.000296361
famID:plantID3	1191	87.3	0.1	3.9	2.92E-109

Residuals	1111	20.8	0.0	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	18.6	2.1	92.4	9.69E-124
rep	1	4.4	4.4	196.4	9.51E-41
famID:plantID3	1182	63.3	0.1	2.4	1.03E-43
Residuals	962	21.5	0.0	-	-

Feekes

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	652.8	72.5	216.9	3.41E-238
rep	1	19.3	19.3	57.8	6.07E-14
famID:plantID3	1181	1250.8	1.1	3.2	2.55E-80
Residuals	1119	374.1	0.3	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	765.1	85.0	273.5	2.10E-278
rep	1	12.6	12.6	40.5	2.89E-10
famID:plantID3	1193	1501.3	1.3	4.0	7.47E-117
Residuals	1149	357.1	0.3	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	460.9	51.2	138.5	3.84E-175
rep	1	22.7	22.7	61.4	1.07E-14
famID:plantID3	1192	1808.0	1.5	4.1	9.13E-117
Residuals	1121	414.4	0.4	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	1159.8	128.9	256.0	4.84E-265
rep	1	17.7	17.7	35.2	3.95E-09
famID:plantID3	1192	2409.2	2.0	4.0	7.52E-114
Residuals	1125	566.2	0.5	-	-

Flag Leaf Area

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	3166.2	351.8	19.8	4.72E-31
rep	1	2686.0	2686.0	151.1	1.17E-32
famID:plantID3	1184	41560.2	35.1	2.0	2.39E-30
Residuals	1119	19896.4	17.8	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	1918.5	213.2	27.0	1.13E-42
rep	1	984.6	984.6	124.9	1.38E-27
famID:plantID3	1192	18956.9	15.9	2.0	1.24E-32
Residuals	1145	9029.0	7.9	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)

famID	9	34704.4	3856.0	64.4	2.31E-95
rep	1	1669.1	1669.1	27.9	1.54E-07
famID:plantID3	1192	177457.6	148.9	2.5	2.34E-52
Residuals	1125	67316.5	59.8	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	3833.3	425.9	17.7	9.88E-28
rep	1	1407.2	1407.2	58.6	4.12E-14
famID:plantID3	1191	48882.2	41.0	1.7	7.77E-20
Residuals	1125	27009.5	24.0	-	-

Height

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	143619.0	15957.7	483.9	0
rep	1	846.3	846.3	25.7	4.75E-07
famID:plantID3	1187	247084.4	208.2	6.3	9.03E-188
Residuals	1122	36997.6	33.0	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	101036.6	11226.3	201.3	2.43E-227
rep	1	0.5	0.5	0.0	0.924025919
famID:plantID3	1183	221174.6	187.0	3.4	3.90E-88
Residuals	1126	62805.5	55.8	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	156498.6	17388.7	212.8	4.33E-236
rep	1	3040.0	3040.0	37.2	1.46E-09
famID:plantID3	1194	414697.9	347.3	4.2	7.55E-123
Residuals	1130	92349.1	81.7	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	50017.5	5557.5	92.2	1.75E-128
rep	1	25367.1	25367.1	421.1	9.79E-80
famID:plantID3	1191	180342.5	151.4	2.5	1.99E-53
Residuals	1125	67776.0	60.2	-	-

Stem Diameter

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	21.9	2.4	47.3	1.93E-72
rep	1	0.1	0.1	2.7	0.09775956
famID:plantID3	1173	149.0	0.1	2.5	3.06E-51
Residuals	1111	57.0	0.1	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	10.8	1.2	40.5	7.86E-63
rep	1	0.8	0.8	27.4	2.00E-07
famID:plantID3	1186	71.8	0.1	2.0	3.94E-33

Residuals	1114	33.0	0.0	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	17.1	1.9	28.6	5.64E-45
rep	1	0.0	0.0	0.1	0.754082019
famID:plantID3	1191	194.9	0.2	2.5	3.91E-51
Residuals	1114	73.9	0.1	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	15.2	1.7	37.0	1.65E-56
rep	1	3.1	3.1	67.1	8.15E-16
famID:plantID3	1182	107.3	0.1	2.0	2.11E-28
Residuals	958	43.6	0.0	-	-

Spikelet Density

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	6.3	0.7	329.3	3.41E-305
rep	1	0.0	0.0	2.8	0.093950741
famID:plantID3	1171	10.4	0.0	4.1	2.89E-116
Residuals	1103	2.4	0.0	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	10.0	1.1	255.0	3.32E-264
rep	1	0.2	0.2	55.9	1.53E-13
famID:plantID3	1188	15.2	0.0	2.9	3.24E-71
Residuals	1123	4.9	0.0	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	5.6	0.6	245.6	2.67E-257
rep	1	0.0	0.0	6.0	0.014862135
famID:plantID3	1191	12.3	0.0	4.1	1.50E-116
Residuals	1114	2.8	0.0	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	6.0	0.7	88.2	2.76E-119
rep	1	1.4	1.4	188.3	2.83E-39
famID:plantID3	1182	27.2	0.0	3.1	1.18E-68
Residuals	961	7.2	0.0	-	-

Florets per Spikelet

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	247.5	27.5	37.2	5.73E-58
rep	1	1.0	1.0	1.4	0.243007464
famID:plantID3	1172	1650.4	1.4	1.9	2.55E-27
Residuals	1111	821.0	0.7	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)

famID	9	111.1	12.3	18.2	2.29E-28
rep	1	25.6	25.6	37.7	1.15E-09
famID:plantID3	1186	1092.1	0.9	1.4	1.69E-07
Residuals	1089	739.8	0.7	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	483.2	53.7	66.3	1.63E-97
rep	1	1.8	1.8	2.2	0.138270838
famID:plantID3	1191	1814.9	1.5	1.9	1.60E-26
Residuals	1116	904.3	0.8	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	90.2	10.0	12.8	1.81E-19
rep	1	37.7	37.7	48.3	6.68E-12
famID:plantID3	1182	1562.8	1.3	1.7	1.32E-17
Residuals	961	749.6	0.8	-	-

Spikelets per Spike

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	2355.3	261.7	90.6	2.97E-126
rep	1	4.5	4.5	1.6	0.210626572
famID:plantID3	1172	9174.8	7.8	2.7	5.69E-61
Residuals	1111	3209.5	2.9	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	2367.4	263.0	80.7	2.84E-115
rep	1	136.9	136.9	42.0	1.35E-10
famID:plantID3	1188	8019.3	6.8	2.1	1.52E-34
Residuals	1123	3658.3	3.3	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	3422.1	380.2	99.2	8.17E-136
rep	1	0.0	0.0	0.0	0.988510548
famID:plantID3	1191	12682.2	10.6	2.8	2.75E-64
Residuals	1115	4273.8	3.8	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	2560.7	284.5	93.1	2.04E-124
rep	1	21.3	21.3	7.0	0.008460519
famID:plantID3	1182	9256.3	7.8	2.6	5.28E-50
Residuals	961	2936.1	3.1	-	-

Floret Site Utilization

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	1.3	0.1	31.4	3.83E-49
rep	1	0.2	0.2	44.9	3.34E-11

famID:plantID3	1172	14.8	0.0	2.7	2.13E-58
Residuals	1095	5.2	0.0	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	1.2	0.1	10.0	7.24E-15
rep	1	0.5	0.5	36.7	1.93E-09
famID:plantID3	1186	28.2	0.0	1.7	2.46E-20
Residuals	1077	14.8	0.0	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	1.0	0.1	26.2	2.75E-41
rep	1	0.0	0.0	9.6	0.001982893
famID:plantID3	1190	16.1	0.0	3.3	2.63E-85
Residuals	1103	4.5	0.0	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	0.1	0.0	21.8	8.79E-33
rep	1	0.1	0.1	115.3	4.90E-25
famID:plantID3	1084	1.3	0.0	2.0	8.73E-23
Residuals	716	0.4	0.0	-	-

Thousand Grain Weight

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	796.3	88.5	185.3	6.68E-214
rep	1	0.2	0.2	0.5	0.488344611
famID:plantID3	1173	2321.9	2.0	4.1	8.04E-117
Residuals	1108	529.1	0.5	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	714.6	79.4	190.0	2.27E-219
rep	1	32.5	32.5	77.8	4.23E-18
famID:plantID3	1188	1753.1	1.5	3.5	7.63E-96
Residuals	1134	473.8	0.4	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	834.5	92.7	190.4	3.18E-218
rep	1	12.2	12.2	25.0	6.49E-07
famID:plantID3	1190	2049.5	1.7	3.5	7.24E-95
Residuals	1114	542.6	0.5	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	104.7	11.6	35.6	2.19E-36
rep	1	2.2	2.2	6.9	0.009450345
famID:plantID3	545	358.9	0.7	2.0	1.67E-08
Residuals	191	62.4	0.3	-	-

Yield per Spike

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	5.9	0.7	61.9	1.08E-91
rep	1	0.9	0.9	86.5	7.28E-20
famID:plantID3	1173	75.3	0.1	6.0	1.55E-176
Residuals	1096	11.6	0.0	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	4.7	0.5	61.6	1.36E-91
rep	1	1.5	1.5	172.6	9.31E-37
famID:plantID3	1187	43.5	0.0	4.3	3.01E-124
Residuals	1113	9.4	0.0	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	1.1	0.1	18.6	4.68E-29
rep	1	0.2	0.2	34.5	5.51E-09
famID:plantID3	1190	50.0	0.0	6.3	4.32E-185
Residuals	1103	7.4	0.0	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	0.1	0.0	16.7	9.01E-25
rep	1	0.0	0.0	88.3	1.09E-19
famID:plantID3	1026	0.7	0.0	1.6	5.89E-11
Residuals	610	0.2	0.0	-	-

Yield per Plant

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	39196.8	4355.2	71.5	2.23E-24
famID:plantID3	1000	166883.6	166.9	2.7	2.44E-05
Residuals	47	2864.8	61.0	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	126614.1	14068.2	61.9	1.08E-36
famID:plantID3	1128	539989.8	478.7	2.1	3.48E-06
Residuals	99	22501.8	227.3	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	29087.9	3232.0	901.1	0.025848201
famID:plantID3	1152	217023.8	188.4	52.5	0.109722278
Residuals	1	3.6	3.6	-	-

Supplementary Table 1.3. Trait means across environments. Same data as Figure 1.2 except actual values are shown here for reference.

Location	**STP**								**TLI**							
Year	**2017**				**2018**				**2017**				**2018**			
Trait (units)	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**
Reproductive Tillers (ct)	0	166	72	30.14	3	334	163	53.41	2	191	96	32.24	5	197	79	33.86
Spike Emergence (%)	0	0.87	0.46	0.16	0	1.1	0.51	0.17	0	1.39	0.77	0.21	0.05	1.9	1.01	0.2
Maturity (1-10)	5	10	7.2	0.91	5.5	9.1	7.4	0.98	4.9	10.1	6.6	0.99	2	9	5.8	1.24
Flag Leaf Area (cm)	6	43	25	4.54	3	24	14	3.03	17	92	48	9.68	5	48	18	4.8
Height (cm)	69	155	118	13.16	95	168	135	11.93	19	150	112	15.84	52	123	92	10.1
Stem Diameter (mm)	1.26	3.65	2.32	0.28	0.75	2.51	1.72	0.19	2.17	4.26	3.18	0.31	1.07	2.91	1.97	0.24
Spikelet Density (cm)	0.51	1.2	0.78	0.09	0.55	1.23	0.83	0.1	0.5	1.09	0.7	0.09	0.61	1.44	0.94	0.13
Florets Spikelet-1 (ct)	5	12	8.8	0.92	3.6	9	6.2	0.74	4.7	12.3	8.5	1	1.1	8.7	4.8	0.91
Spikelets Spike-1 (ct)	11.5	27.8	21.4	2.28	10.2	28.5	20.5	2.16	13.5	34.2	27.3	2.68	16.2	31.9	23.9	2.39
Floret Site Utilization (%)	0.02	0.81	0.48	0.09	0.18	1.43	0.6	0.12	0.05	0.62	0.37	0.09	0	0.25	0.03	0.03
Thousand Grain Weight (mg)	3.8	13.2	9.5	1.18	4.2	11.4	8.4	1.04	3.6	11.6	7.9	1.12	2.7	7.9	5.5	0.8
Yield Spike-1 (g)	0.01	1.35	0.77	0.19	0.08	1.03	0.59	0.15	0.04	1.06	0.59	0.15	0	0.17	0.02	0.02
Yield Plant-1 (g)	0.2	76.9	28.9	14.06	1.1	132	54.9	23.32	0.3	80.8	29	14.56	-	-	-	-

Supplementary Table 1.4.

Pearson correlation coefficients for yield component traits of intermediate wheatgrass spaced plants in St. Paul, MN in 2017.

	Reproductive Tillers (ct)	Spike Emergence (%)	Maturity (1-10)	Flag Leaf Area (cm)	Height (cm)	Stem Diameter (mm)	Spikelet Density (cm)	Florets Spikelet-1 (ct)	Spikelets Spike-1 (ct)	Floret Site Utilization (%)	Thousand Grain Weight (mg)	Yield Spike-1 (g)	Yield Plant-1 (g)
Reproductive Tillers (ct)	1	-	-	-	-	-	-	-	-	-	-	-	-
Spike Emergence (%)	0.17***	1	-	-	-	-	-	-	-	-	-	-	-
Maturity (1-10)	0.22***	0.65***	1	-	-	-	-	-	-	-	-	-	-
Flag Leaf Area (cm)	0.33***	0.15***	0.11***	1	-	-	-	-	-	-	-	-	-
Height (cm)	0.46***	-0.05.	0.17***	0.26***	1	-	-	-	-	-	-	-	-
Stem Diameter (mm)	0.2***	0.15***	0.13***	0.48***	0.1***	1	-	-	-	-	-	-	-
Spikelet Density (cm)	-0.04	0.07*	0.12***	-0.04	-0.45***	-0.1***	1	-	-	-	-	-	-
Florets Spikelet-1 (ct)	-0.05	0.1***	-0.04	0.33***	-0.17***	0.45***	0.07*	1	-	-	-	-	-
Spikelets Spike-1 (ct)	0.37***	0.17***	0.3***	0.4***	0.13***	0.5***	0.42***	0.27***	1	-	-	-	-
Floret Site Utilization (%)	0.2***	0.09**	0.14***	0.11***	0.26***	0.05.	-0.14***	-0.22***	0.05.	1	-	-	-
Thousand Grain Weight (mg)	0.24***	-0.03	-0.01	0.24***	0.42***	0.29***	-0.44***	0.01	-0.03	0.05.	1	-	-
Yield Spike-1 (g)	0.36***	0.14***	0.21***	0.46***	0.36***	0.52***	-0.06*	0.34***	0.54***	0.64***	0.39***	1	-
Yield Plant-1 (g)	0.71***	0.13***	0.22***	0.35***	0.37***	0.32***	0	0.08**	0.47***	0.41***	0.24***	0.65***	1

. Indicates significance at the 0.1 level
* indicates significance at the 0.05 level
** indicates significance at the 0.01 level
*** indicates significance at the 0.001 level

Supplementary Table 1.5. Pearson correlation coefficients for yield component traits of intermediate wheatgrass spaced plants in St. Paul, MN in 2018.

	Reproductive Tillers (ct)	Spike Emergence (%)	Maturity (1-10)	Flag Leaf Area (cm)	Height (cm)	Stem Diameter (mm)	Spikelet Density (cm)	Florets Spikelet-1 (ct)	Spikelets Spike-1 (ct)	Floret Site Utilization (%)	Thousand Grain Weight (mg)	Yield Spike-1 (g)	Yield Plant-1 (g)
Reproductive Tillers (ct)	1	-	-	-	-	-	-	-	-	-	-	-	-
Spike Emergence (%)	-0.18***	1	-	-	-	-	-	-	-	-	-	-	-
Maturity (1-10)	-0.02	0.71***	1	-	-	-	-	-	-	-	-	-	-
Flag Leaf Area (cm)	0.16***	0	0.04	1	-	-	-	-	-	-	-	-	-
Height (cm)	0.4***	-0.2***	-0.01	0.37***	1	-	-	-	-	-	-	-	-
Stem Diameter (mm)	-0.02	-0.09**	-0.07*	0.46***	0.27***	1	-	-	-	-	-	-	-
Spikelet Density (cm)	0.09**	0.31***	0.28***	-0.08**	-0.32***	-0.16***	1	-	-	-	-	-	-
Florets Spikelet-1 (ct)	-0.15***	0.13***	0.05.	0.28***	0.01	0.36***	0.01	1	-	-	-	-	-
Spikelets Spike-1 (ct)	0.24***	0.02	0.08**	0.27***	0.12***	0.35***	0.54***	0.21***	1	-	-	-	-
Floret Site Utilization (%)	0.06*	0.13***	0.18***	0.18***	0.18***	0.14***	-0.16***	-0.3***	-0.06*	1	-	-	-
Thousand Grain Weight (mg)	0.11***	-0.16***	-0.15***	0.31***	0.43***	0.39***	-0.39***	0.14***	-0.06.	0.06*	1	-	-
Yield Spike-1 (g)	0.09**	0.13***	0.18***	0.51***	0.35***	0.55***	-0.01	0.41***	0.43***	0.54***	0.43***	1	-
Yield Plant-1 (g)	0.6***	-0.02	0.11***	0.32***	0.37***	0.19***	0.12***	0.06*	0.35***	0.3***	0.2***	0.51***	1

. Indicates significance at the 0.1 level
* indicates significance at the 0.05 level
** indicates significance at the 0.01 level
*** indicates significance at the 0.001 level

Supplementary Table 1.6.

Pearson correlation coefficients for yield component traits of intermediate wheatgrass spaced plants in Salina, KS at the Land Institute in 2017.

	Reproductive Tillers (ct)	Spike Emergence (%)	Maturity (1-10)	Flag Leaf Area (cm)	Height (cm)	Stem Diameter (mm)	Spikelet Density (cm)	Florets Spikelet-1 (ct)	Spikelets Spike-1 (ct)	Floret Site Utilization (%)	Thousand Grain Weight (mg)	Yield Spike-1 (g)	Yield Plant-1 (g)
Reproductive Tillers (ct)	1	-	-	-	-	-	-	-	-	-	-	-	-
Spike Emergence (%)	0.28***	1	-	-	-	-	-	-	-	-	-	-	-
Maturity (1-10)	0.16***	0.67***	1	-	-	-	-	-	-	-	-	-	-
Flag Leaf Area (cm)	0.04	0.16***	0.15***	1	-	-	-	-	-	-	-	-	-
Height (cm)	0.55***	0.17***	0.06*	-0.02	1	-	-	-	-	-	-	-	-
Stem Diameter (mm)	-0.08**	0.2***	0.23***	0.37***	0.02	1	-	-	-	-	-	-	-
Spikelet Density (cm)	-0.05.	0.08**	0.1***	-0.1***	-0.29***	-0.09**	1	-	-	-	-	-	-
Florets Spikelet-1 (ct)	-0.19***	0.25***	0.28***	0.27***	-0.27***	0.36***	0.07*	1	-	-	-	-	-
Spikelets Spike-1 (ct)	0.2***	0.15***	0.25***	0.16***	0.13***	0.35***	0.52***	0.17***	1	-	-	-	-
Floret Site Utilization (%)	0.15***	0.04	0.06.	-0.08**	0.24***	-0.02	-0.09**	-0.25***	0.02	1	-	-	-
Thousand Grain Weight (mg)	0.2***	0.05.	-0.07*	0.04	0.37***	0.03	-0.39***	-0.18***	-0.19***	-0.08**	1	-	-
Yield Spike-1 (g)	0.22***	0.19***	0.22***	0.12***	0.34***	0.28***	-0.03	0.18***	0.37***	0.72***	0.2***	1	-
Yield Plant-1 (g)	0.67***	0.3***	0.24***	0.05.	0.48***	0.08**	0.03	-0.04	0.31***	0.42***	0.11***	0.58***	1

. Indicates significance at the 0.1 level
* indicates significance at the 0.05 level
** indicates significance at the 0.01 level
*** indicates significance at the 0.001 level

Supplementary Table 1.7.

Pearson correlation coefficients for yield component traits of intermediate wheatgrass spaced plants in Salina, KS at the Land Institute in 2018.

	Reproductive Tillers (ct)	Spike Emergence (%)	Maturity (1-10)	Flag Leaf Area (cm)	Height (cm)	Stem Diameter (mm)	Spikelet Density (cm)	Florets Spikelet-1 (ct)	Spikelets Spike-1 (ct)	Floret Site Utilization (%)	Thousand Grain Weight (mg)	Yield Spike-1 (g)	Yield Plant-1 (g)
Reproductive Tillers (ct)	1	-	-	-	-	-	-	-	-	-	-	-	-
Spike Emergence (%)	0.2***	1	-	-	-	-	-	-	-	-	-	-	-
Maturity (1-10)	0.11***	0.62***	1	-	-	-	-	-	-	-	-	-	-
Flag Leaf Area (cm)	-0.15***	0.04	0.25***	1	-	-	-	-	-	-	-	-	-
Height (cm)	0.5***	0.13***	0.15***	0.11***	1	-	-	-	-	-	-	-	-
Stem Diameter (mm)	-0.26***	-0.09**	0.25***	0.51***	0.11***	1	-	-	-	-	-	-	-
Spikelet Density (cm)	0.13***	0.28***	0.04	-0.24***	-0.27***	-0.34***	1	-	-	-	-	-	-
Florets Spikelet-1 (ct)	-0.16***	0.11***	0.3***	0.37***	0.09**	0.52***	-0.31***	1	-	-	-	-	-
Spikelets Spike-1 (ct)	0.1***	-0.07*	0.18***	0.15***	0.23***	0.34***	0.39***	0.09**	1	-	-	-	-
Floret Site Utilization (%)	0.19***	0.3***	0.43***	0	0.22***	0.08**	0.07*	0.06*	0.16***	1	-	-	-
Thousand Grain Weight (mg)	-0.03	-0.04	0.06	0.15***	0.33***	0.17***	-0.3***	0.09*	0.04	0.08.	1	-	-
Yield Spike-1 (g)	0.14***	0.27***	0.45***	0.07*	0.25***	0.19***	0.03	0.22***	0.22***	0.93***	0.25***	1	-
Yield Plant-1 (g)	-	-	-	-	-	-	-	-	-	-	-	-	1

. Indicates significance at the 0.1 level
* indicates significance at the 0.05 level
** indicates significance at the 0.01 level
*** indicates significance at the 0.001 level

Supplementary Table 1.8. Heritability estimates for each trait in each environment. Mean values are shown in Figure 1.2.

	STP						TLI			
	2017		2018		2017		2018		Average	
Trait	**Ha**	h^{2b}	**H**	h^2	**H**	h^2	**H**	h^2	**H**	h^2
Reproductive Tillers (ct)	0.83	0.58	0.73	0.55	0.76	0.52	0.45	0.25	0.69	0.47
Spike Emergence (%)	0.70	0.47	0.76	0.57	0.78	0.30	0.70	0.48	0.73	0.46
Maturity (1-10)	0.80	0.64	0.84	0.63	0.81	0.47	0.84	0.62	0.82	0.59
Flag Leaf Area (cm)	0.52	0.17	0.55	0.22	0.67	0.37	0.46	0.17	0.55	0.23
Height (cm)	0.90	0.69	0.80	0.62	0.83	0.58	0.69	0.46	0.81	0.59
Stem Diameter (mm)	0.65	0.30	0.58	0.30	0.63	0.20	0.58	0.29	0.61	0.27
Spikelet Density (cm)	0.85	0.68	0.80	0.68	0.84	0.61	0.75	0.42	0.81	0.60
Florets Spikelet-1 (ct)	0.54	0.30	0.33	0.21	0.59	0.44	0.45	0.13	0.48	0.27
Spikelets Spike-1 (ct)	0.71	0.45	0.63	0.47	0.72	0.46	0.71	0.47	0.69	0.46
Floret Site Utilization (%)	0.66	0.22	0.45	0.10	0.72	0.17	0.59	0.22	0.60	0.18
Thousand Grain Weight (mg)	0.82	0.54	0.80	0.58	0.80	0.58	0.70	0.56	0.78	0.56
Yield Spike-1 (g)	0.85	0.22	0.80	0.27	0.85	0.06	0.50	0.21	0.75	0.19
Yield Plant-1 (g)	0.83	0.50	0.75	0.49	0.81	0.35	-	-	0.79	0.45

Supplementary Table 1.9. Standardized path coefficients for direct and indirect effects between traits in the structural equation model for St. Paul in 2017, where Trait 1 has the effect (direct or indirect) on Trait 2.

Trait 1	Effect	Trait 2	est[a]	se[b]	z	pvalue	ci.lower[c]	ci.upper[d]
Stem Diameter	Direct	Thousand Grain Weight	0.22	0.03	6.95	0.000	0.16	0.28
Flag Leaf Area	Direct	Thousand Grain Weight	0.06	0.03	1.84	0.066	0.00	0.12
Feekes	Direct	Thousand Grain Weight	-0.03	0.04	-0.88	0.378	-0.11	0.04
Spikelet Density	Direct	Thousand Grain Weight	-0.31	0.03	-9.95	0.000	-0.37	-0.25
Height	Direct	Thousand Grain Weight	0.22	0.04	6.09	0.000	0.15	0.29
Emergence Percent	Direct	Thousand Grain Weight	-0.02	0.04	-0.52	0.606	-0.09	0.05
Reproductive Tillers	Direct	Thousand Grain Weight	0.07	0.03	2.24	0.025	0.01	0.13
Stem Diameter	Direct	Floret Site Utilization	0.03	0.04	0.71	0.475	-0.05	0.10
Flag Leaf Area	Direct	Floret Site Utilization	0.11	0.04	3.00	0.003	0.04	0.18
Feekes	Direct	Floret Site Utilization	0.06	0.04	1.55	0.120	-0.02	0.14
Spikelet Density	Direct	Floret Site Utilization	-0.06	0.03	-1.62	0.104	-0.12	0.01
Height	Direct	Floret Site Utilization	0.11	0.04	2.65	0.008	0.03	0.18
Emergence Percent	Direct	Floret Site Utilization	0.03	0.04	0.85	0.398	-0.04	0.11
Reproductive Tillers	Direct	Floret Site Utilization	0.09	0.04	2.42	0.016	0.02	0.15
Florets per Spikelet	Direct	Floret Site Utilization	-0.26	0.03	-7.56	0.000	-0.33	-0.20
Florets per Spikelet	Direct	Yield per Spike	0.38	0.01	40.84	0.000	0.36	0.40
Floret Site Utilization	Direct	Yield per Spike	0.70	0.01	76.56	0.000	0.68	0.72
Thousand Grain Weight	Direct	Yield per Spike	0.37	0.01	42.87	0.000	0.35	0.38
Spikelets per Spike	Direct	Yield per Spike	0.43	0.01	47.87	0.000	0.41	0.45
Reproductive Tillers	Direct	Yield per Plant	0.59	0.02	30.50	0.000	0.55	0.63
Yield per Spike	Direct	Yield per Plant	0.48	0.02	26.20	0.000	0.44	0.51
Florets per Spikelet	Indirect	Yield per Spike	-0.18	0.02	-7.52	0.000	-0.23	-0.14
Feekes	Indirect	Yield per Spike	0.03	0.03	1.02	0.308	-0.03	0.09
Stem Diameter	Indirect	Yield per Spike	0.10	0.03	3.52	0.000	0.04	0.15
Spikelet Density	Indirect	Yield per Spike	-0.15	0.03	-5.75	0.000	-0.20	-0.10
Height	Indirect	Yield per Spike	0.15	0.03	5.00	0.000	0.09	0.21
Flag Leaf Area	Indirect	Yield per Spike	0.10	0.03	3.49	0.000	0.04	0.15
Reproductive Tillers	Indirect	Yield per Spike	0.09	0.03	3.14	0.002	0.03	0.14
Emergence Percent	Indirect	Yield per Spike	0.02	0.03	0.54	0.590	-0.04	0.08
Florets per Spikelet	Indirect	Yield per Plant	0.09	0.01	7.28	0.000	0.07	0.12

Height	Indirect	Yield per Plant	0.07	0.01	4.91	0.000	0.04	0.10
Stem Diameter	Indirect	Yield per Plant	0.05	0.01	3.39	0.001	0.02	0.07
Feekes	Indirect	Yield per Plant	0.02	0.01	1.02	0.308	-0.01	0.04
Emergence Percent	Indirect	Yield per Plant	0.01	0.01	0.54	0.590	-0.02	0.04
Spikelet Density	Indirect	Yield per Plant	-0.07	0.01	-5.62	0.000	-0.10	-0.05
Flag Leaf Area	Indirect	Yield per Plant	0.05	0.01	3.46	0.001	0.02	0.07
Floret Site Utilization	Indirect	Yield per Plant	0.33	0.01	24.79	0.000	0.31	0.36
Thousand Grain Weight	Indirect	Yield per Plant	0.17	0.01	22.35	0.000	0.16	0.19
Spikelets per Spike	Indirect	Yield per Plant	0.20	0.01	22.98	0.000	0.19	0.22
Reproductive Tillers	Indirect	Yield per Plant	0.04	0.01	3.12	0.002	0.02	0.07

[a] est, standardized path estimate

[b] se, standard error

[c-d] upper and lower 95% confidence interval

Supplementary Table 1.10. Standardized path coefficients for direct and indirect effects between traits in the structural equation model for St. Paul in 2018, where Trait 1 has the effect (direct or indirect) on Trait 2.

Trait 1	Effect	Trait 2	est[a]	se[b]	z	pvalue	ci.lower[c]	ci.upper[d]
Stem Diameter	Direct	Thousand Grain Weight	0.24	0.03	8.24	0.000	0.18	0.29
Flag Leaf Area	Direct	Thousand Grain Weight	0.09	0.03	3.18	0.001	0.04	0.15
Feekes	Direct	Thousand Grain Weight	-0.10	0.04	-2.84	0.005	-0.17	-0.03
Spikelet Density	Direct	Thousand Grain Weight	-0.28	0.03	-9.86	0.000	-0.33	-0.22
Height	Direct	Thousand Grain Weight	0.23	0.03	7.21	0.000	0.17	0.29
Emergence Percent	Direct	Thousand Grain Weight	0.06	0.04	1.57	0.116	-0.01	0.13
Reproductive Tillers	Direct	Thousand Grain Weight	0.04	0.03	1.31	0.190	-0.02	0.10
Stem Diameter	Direct	Floret Site Utilization	0.19	0.03	6.36	0.000	0.13	0.25
Flag Leaf Area	Direct	Floret Site Utilization	0.19	0.03	6.33	0.000	0.13	0.25
Feekes	Direct	Floret Site Utilization	0.17	0.04	4.62	0.000	0.10	0.24
Spikelet Density	Direct	Floret Site Utilization	-0.18	0.03	-6.21	0.000	-0.24	-0.12
Height	Direct	Floret Site Utilization	0.03	0.03	1.05	0.291	-0.03	0.10
Emergence Percent	Direct	Floret Site Utilization	0.15	0.04	4.05	0.000	0.08	0.23
Reproductive Tillers	Direct	Floret Site Utilization	-0.02	0.03	-0.64	0.522	-0.08	0.04
Florets per Spikelet	Direct	Floret Site Utilization	-0.46	0.03	-16.56	0.000	-0.52	-0.41
Florets per Spikelet	Direct	Yield per Spike	0.49	0.01	44.60	0.000	0.47	0.51
Floret Site Utilization	Direct	Yield per Spike	0.70	0.01	64.43	0.000	0.68	0.72
Thousand Grain Weight	Direct	Yield per Spike	0.35	0.01	34.17	0.000	0.33	0.37
Spikelets per Spike	Direct	Yield per Spike	0.39	0.01	36.11	0.000	0.37	0.41
Reproductive Tillers	Direct	Yield per Plant	0.63	0.02	31.39	0.000	0.59	0.67
Yield per Spike	Direct	Yield per Plant	0.51	0.02	27.21	0.000	0.47	0.55
Florets per Spikelet	Indirect	Yield per Spike	-0.32	0.02	-16.04	0.000	-0.36	-0.28
Feekes	Indirect	Yield per Spike	0.08	0.03	2.89	0.004	0.03	0.14
Stem Diameter	Indirect	Yield per Spike	0.22	0.02	9.21	0.000	0.17	0.27
Spikelet Density	Indirect	Yield per Spike	-0.22	0.02	-9.82	0.000	-0.27	-0.18
Height	Indirect	Yield per Spike	0.10	0.03	4.11	0.000	0.05	0.15
Flag Leaf Area	Indirect	Yield per Spike	0.17	0.02	7.05	0.000	0.12	0.21
Reproductive Tillers	Indirect	Yield per Spike	0.00	0.02	0.00	0.997	-0.05	0.05
Emergence Percent	Indirect	Yield per Spike	0.13	0.03	4.32	0.000	0.07	0.19

Florets per Spikelet	Indirect	Yield per Plant	0.09	0.01	7.29	0.000	0.06	0.11
Height	Indirect	Yield per Plant	0.05	0.01	4.07	0.000	0.03	0.08
Stem Diameter	Indirect	Yield per Plant	0.11	0.01	8.46	0.000	0.08	0.14
Feekes	Indirect	Yield per Plant	0.04	0.01	2.87	0.004	0.01	0.07
Emergence Percent	Indirect	Yield per Plant	0.07	0.02	4.27	0.000	0.04	0.10
Spikelet Density	Indirect	Yield per Plant	-0.11	0.01	-9.24	0.000	-0.14	-0.09
Flag Leaf Area	Indirect	Yield per Plant	0.09	0.01	6.83	0.000	0.06	0.11
Floret Site Utilization	Indirect	Yield per Plant	0.36	0.01	25.06	0.000	0.33	0.39
Thousand Grain Weight	Indirect	Yield per Plant	0.18	0.01	21.28	0.000	0.16	0.20
Spikelets per Spike	Indirect	Yield per Plant	0.20	0.01	21.73	0.000	0.18	0.22
Reproductive Tillers	Indirect	Yield per Plant	0.00	0.01	0.00	0.997	-0.02	0.02

[a] est, standardized path estimate

[b] se, standard error

[c-d] upper and lower 95% confidence interval

Supplementary Table 1.11. Standardized path coefficients for direct and indirect effects between traits in the structural equation model for the Land Institute in Salina, KS in 2017, where Trait 1 has the effect (direct or indirect) on Trait 2.

Trait 1	Effect	Trait 2	est[a]	se[b]	z	pvalue	ci.lower[c]	ci.upper[d]
Stem Diameter	Direct	Thousand Grain Weight	0.01	0.03	0.40	0.687	-0.04	0.07
Flag Leaf Area	Direct	Thousand Grain Weight	-0.01	0.03	-0.47	0.641	-0.07	0.04
Feekes	Direct	Thousand Grain Weight	-0.12	0.03	-3.51	0.000	-0.19	-0.05
Spikelet Density	Direct	Thousand Grain Weight	-0.31	0.03	-11.69	0.000	-0.37	-0.26
Height	Direct	Thousand Grain Weight	0.29	0.03	8.74	0.000	0.22	0.35
Emergence Percent	Direct	Thousand Grain Weight	0.10	0.04	2.79	0.005	0.03	0.17
Floret Site Utilization	Direct	Thousand Grain Weight	-0.18	0.03	-6.97	0.000	-0.23	-0.13
Reproductive Tillers	Direct	Thousand Grain Weight	0.04	0.03	1.40	0.161	-0.02	0.11
Stem Diameter	Direct	Floret Site Utilization	0.06	0.03	1.97	0.048	0.00	0.13
Flag Leaf Area	Direct	Floret Site Utilization	-0.06	0.03	-1.84	0.065	-0.12	0.00
Feekes	Direct	Floret Site Utilization	0.12	0.04	3.08	0.002	0.04	0.19
Spikelet Density	Direct	Floret Site Utilization	-0.05	0.03	-1.69	0.090	-0.11	0.01
Height	Direct	Floret Site Utilization	0.15	0.04	4.07	0.000	0.08	0.22
Emergence Percent	Direct	Floret Site Utilization	-0.01	0.04	-0.36	0.719	-0.09	0.06
Reproductive Tillers	Direct	Floret Site Utilization	0.02	0.04	0.45	0.653	-0.05	0.08
Florets per Spikelet	Direct	Floret Site Utilization	-0.26	0.03	-7.89	0.000	-0.32	-0.19
Florets per Spikelet	Direct	Yield per Spike	0.42	0.01	41.99	0.000	0.40	0.44
Height	Direct	Yield per Spike	0.06	0.01	5.47	0.000	0.04	0.08
Floret Site Utilization	Direct	Yield per Spike	0.84	0.01	85.82	0.000	0.82	0.86
Thousand Grain Weight	Direct	Yield per Spike	0.38	0.01	36.92	0.000	0.36	0.40
Spikelets per Spike	Direct	Yield per Spike	0.35	0.01	36.24	0.000	0.33	0.37
Reproductive Tillers	Direct	Yield per Plant	0.58	0.02	31.68	0.000	0.54	0.61
Yield per Spike	Direct	Yield per Plant	0.46	0.02	25.93	0.000	0.43	0.49
Florets per Spikelet	Indirect	Yield per Spike	-0.22	0.03	-7.86	0.000	-0.27	-0.16
Feekes	Indirect	Yield per Spike	0.05	0.03	1.52	0.127	-0.01	0.12
Stem Diameter	Indirect	Yield per Spike	0.06	0.03	1.98	0.047	0.00	0.11
Spikelet Density	Indirect	Yield per Spike	-0.16	0.03	-5.96	0.000	-0.21	-0.11
Height	Indirect	Yield per Spike	0.23	0.03	7.02	0.000	0.17	0.30
Flag Leaf Area	Indirect	Yield per Spike	-0.05	0.03	-1.88	0.060	-0.11	0.00
Floret Site Utilization	Indirect	Yield per Spike	-0.07	0.01	-6.85	0.000	-0.09	-0.05
Empergence Percent	Indirect	Yield per Spike	0.03	0.04	0.72	0.473	-0.04	0.09

Reproductive Tillers	Indirect	Yield per Spike	0.03	0.03	0.95	0.344	-0.03	0.09
Florets per Spikelet	Indirect	Yield per Plant	0.58	0.02	24.73	0.000	0.53	0.62
Height	Indirect	Yield per Plant	0.14	0.02	8.08	0.000	0.10	0.17
Stem Diameter	Indirect	Yield per Plant	0.07	0.01	5.33	0.000	0.05	0.10
Feekes	Indirect	Yield per Plant	0.02	0.02	1.52	0.128	-0.01	0.06
Emergence Percent	Indirect	Yield per Plant	0.01	0.02	0.72	0.474	-0.02	0.04
Spikelet Density	Indirect	Yield per Plant	-0.07	0.01	-5.81	0.000	-0.10	-0.05
Flag Leaf Area	Indirect	Yield per Plant	-0.02	0.01	-1.88	0.061	-0.05	0.00
Floret Site Utilization	Indirect	Yield per Plant	0.35	0.02	23.54	0.000	0.32	0.38
Thousand Grain Weight	Indirect	Yield per Plant	0.17	0.01	21.22	0.000	0.16	0.19
Spikelets per Spike	Indirect	Yield per Plant	0.16	0.01	21.09	0.000	0.15	0.18
Reproductive Tillers	Indirect	Yield per Plant	0.01	0.01	0.94	0.345	-0.01	0.04

[a] est, standardized path estimate

[b] se, standard error

[c-d] upper and lower 95% confidence interval

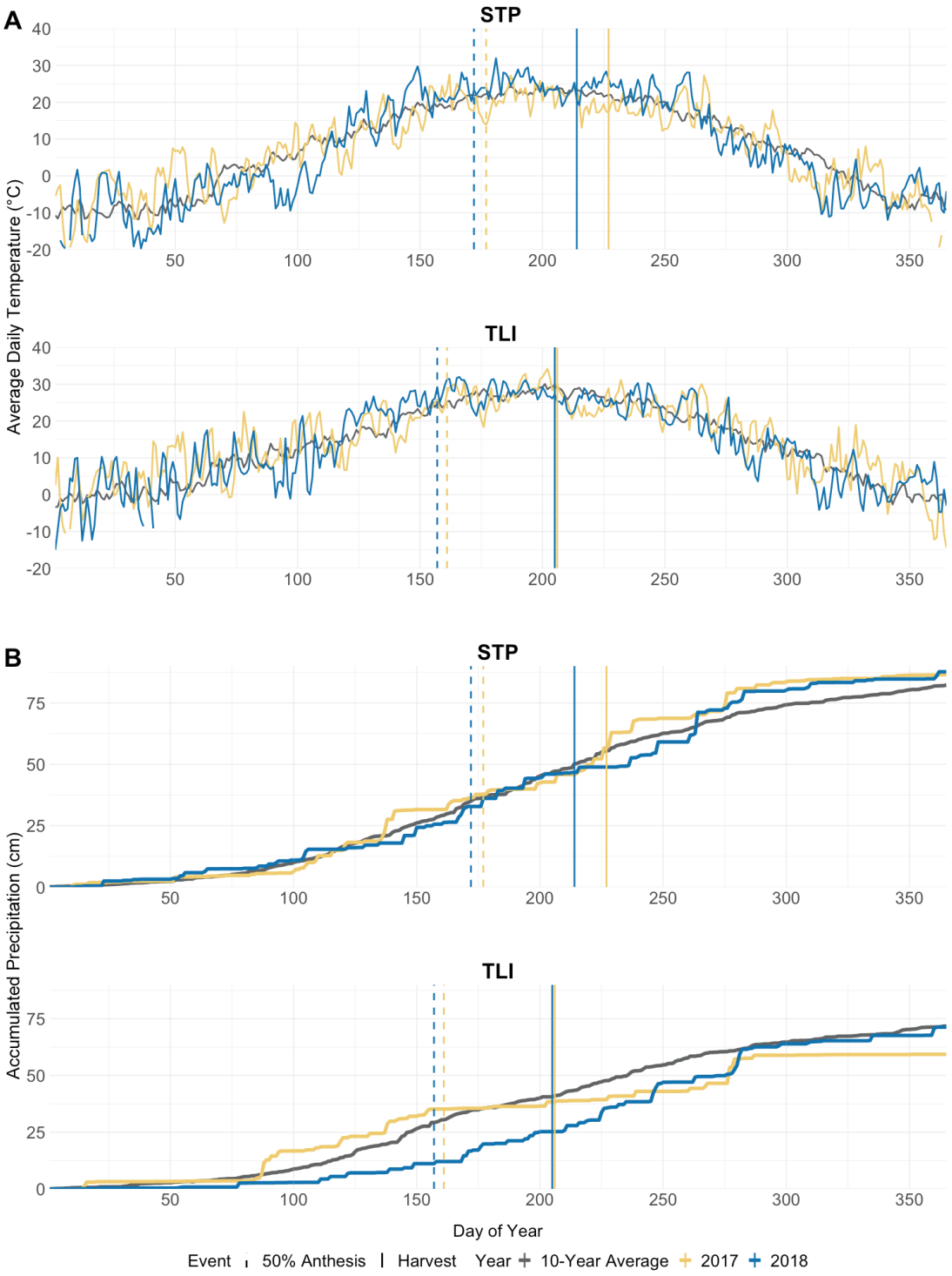
Supplementary Table 1.12. Standardized path coefficients for direct and indirect effects between traits in the structural equation model for the Land Institute in Salina, KS in 2018, where Trait 1 has the effect (direct or indirect) on Trait 2.

Trait 1	Effect	Trait 2	est[a]	se[b]	z	pvalue	ci.lower[c]	ci.upper[d]
Stem Diameter	Direct	Thousand Grain Weight	0.03	0.05	0.48	0.629	-0.08	0.13
Flag Leaf Area	Direct	Thousand Grain Weight	0.05	0.05	1.03	0.303	-0.04	0.13
Feekes	Direct	Thousand Grain Weight	0.08	0.05	1.51	0.132	-0.02	0.18
Spikelet Density	Direct	Thousand Grain Weight	-0.17	0.05	-3.59	0.000	-0.27	-0.08
Height	Direct	Thousand Grain Weight	0.39	0.05	7.54	0.000	0.29	0.49
Emergence Percent	Direct	Thousand Grain Weight	-0.04	0.06	-0.64	0.520	-0.16	0.08
Reproductive Tillers	Direct	Thousand Grain Weight	-0.14	0.05	-2.75	0.006	-0.23	-0.04
Stem Diameter	Direct	Floret Site Utilization	0.07	0.06	1.17	0.241	-0.04	0.18
Flag Leaf Area	Direct	Floret Site Utilization	-0.18	0.05	-3.72	0.000	-0.27	-0.08
Feekes	Direct	Floret Site Utilization	0.35	0.05	6.51	0.000	0.25	0.46
Spikelet Density	Direct	Floret Site Utilization	0.10	0.05	1.97	0.049	0.00	0.20
Height	Direct	Floret Site Utilization	0.22	0.06	4.03	0.000	0.11	0.33
Emergence Percent	Direct	Floret Site Utilization	0.02	0.07	0.27	0.787	-0.11	0.15
Reproductive Tillers	Direct	Floret Site Utilization	0.05	0.05	0.94	0.348	-0.05	0.15
Florets per Spikelet	Direct	Yield per Spike	0.28	0.02	18.49	0.000	0.25	0.31
Floret Site Utilization	Direct	Yield per Spike	0.90	0.01	69.61	0.000	0.88	0.93
Thousand Grain Weight	Direct	Yield per Spike	0.16	0.01	11.69	0.000	0.13	0.19
Spikelets per Spike	Direct	Yield per Spike	0.13	0.02	8.62	0.000	0.10	0.16
Feekes	Indirect	Yield per Spike	0.33	0.05	6.64	0.000	0.23	0.43
Stem Diameter	Indirect	Yield per Spike	0.06	0.05	1.24	0.217	-0.04	0.16
Spikelet Density	Indirect	Yield per Spike	0.06	0.05	1.35	0.176	-0.03	0.16
Height	Indirect	Yield per Spike	0.26	0.05	5.17	0.000	0.16	0.36
Flag Leaf Area	Indirect	Yield per Spike	-0.15	0.04	-3.49	0.000	-0.24	-0.07
Emergence Percent	Indirect	Yield per Spike	0.01	0.06	0.16	0.872	-0.11	0.13
Reproductive Tillers	Indirect	Yield per Spike	0.02	0.05	0.47	0.636	-0.07	0.12

[a] est, standardized path estimate

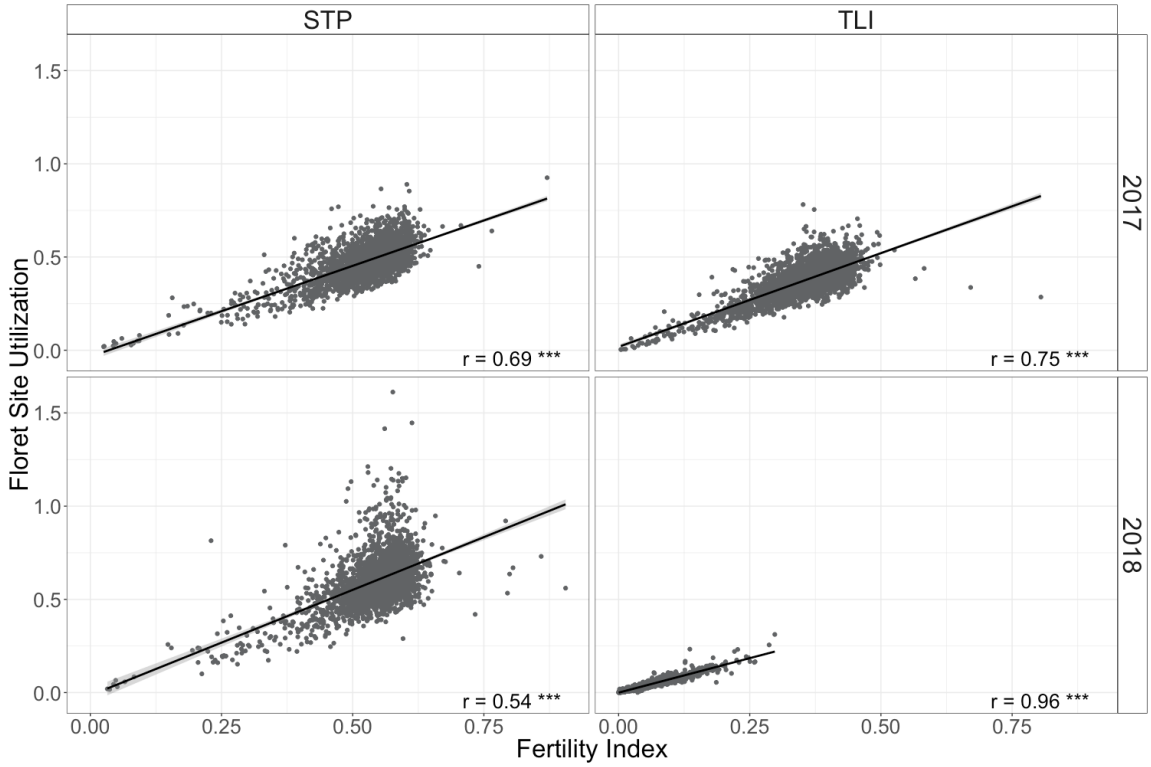
[b] se, standard error

[c-d] upper and lower 95% confidence interval

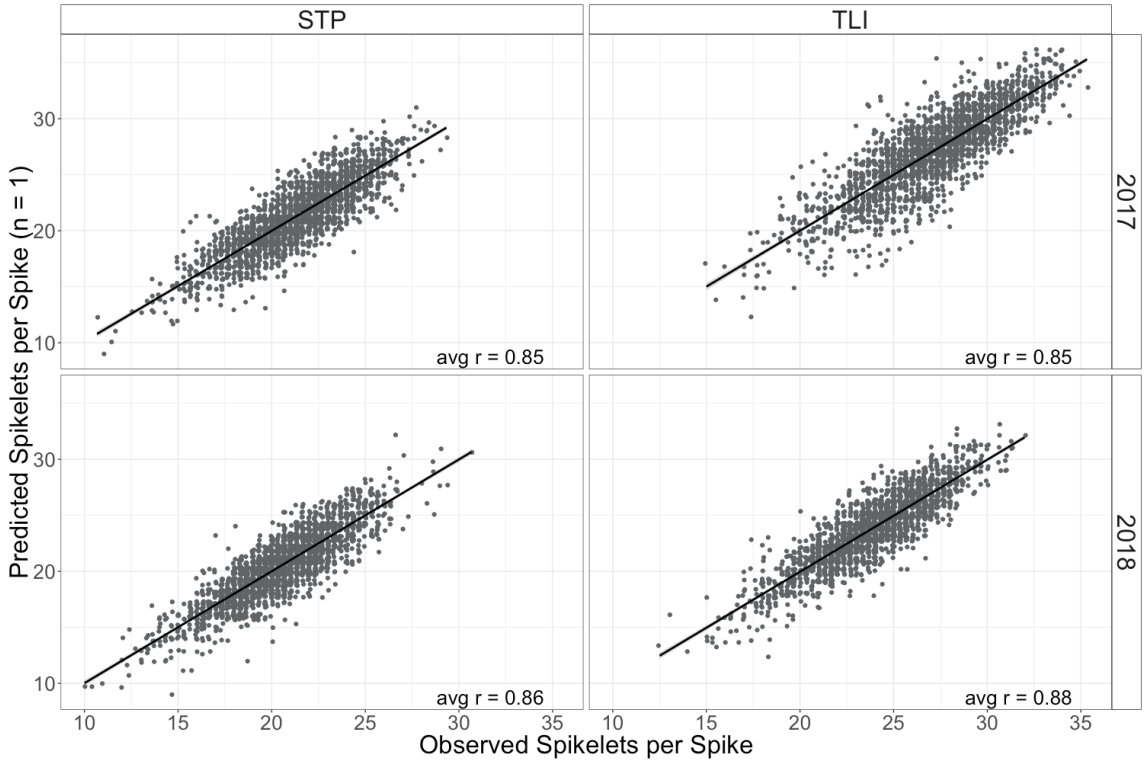

A
STP
TLI
Average Daily Temperature (°C)
40
30
20
10
0
-10
-20
50
100
150
200
250
300
350
B
STP
TLI
Accumulated Precipitation (cm)
75
50
25
0
Day of Year
Event
50% Anthesis
Harvest
Year
10-Year Average
2017
2018

Supplementary Figure 1.1 A) Average daily temperature [(daily min + daily max) / 2)] in °C for St. Paul, MN (STP) and Salina, KS (TLI) during 2017 (yellow), 2018 (blue) compared with the average over the preceding ten years (2006 - 2016) (gray). **B**) Accumulated precipitation (cm). Dashed vertical reference lines indicate 50% anthesis, and solid vertical lines indicate date of harvest. Data were obtained from the National Oceanic and Atmospheric Administration for the following stations; USW00003919 (Lat. 38.80; Long -97.65) and USC00218450 (Lat. 44.99; Long: -93.18) as well as data from the Land Institute Weather Station.

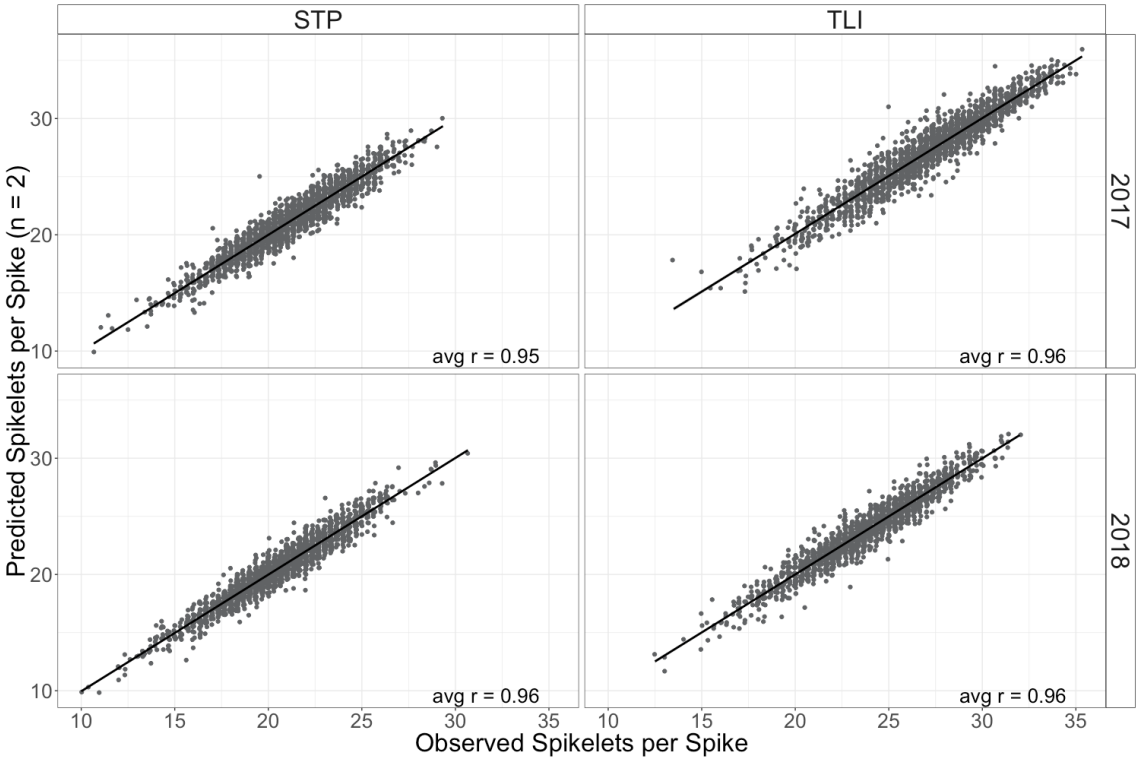
Supplementary Figure 1.2. To estimate reproductive tiller counts per spaced plant, stems were gathered, cut with a large garden shears and the underside of the bundle was photographed using a Canon digital SLR camera. The multi-point function in ImageJ was used to click on and quantify the number of stems.



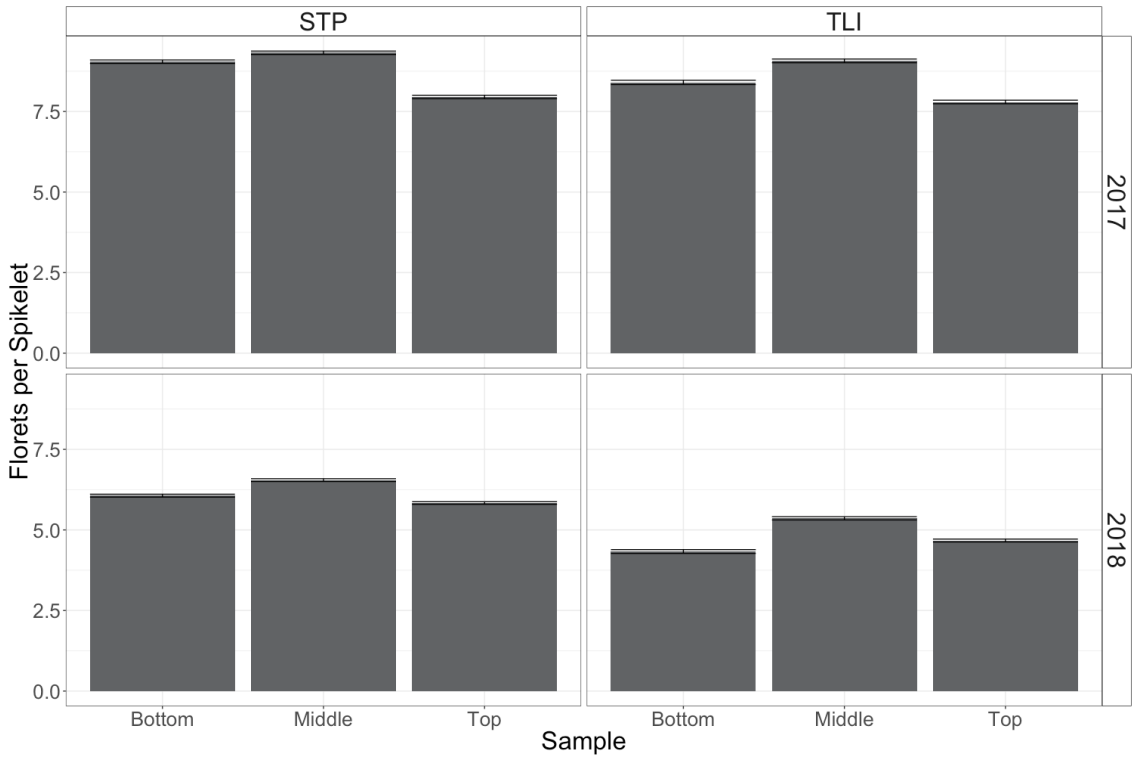
Supplementary Figure 1.3. Fertility index (weight of seeds / head weight) of intermediate wheatgrass spikes correlated to floret site utilization across two locations (STP & TLI) in 2017 and 2018. R is the Pearson correlation coefficient followed by the significance ($P < 0.001$).



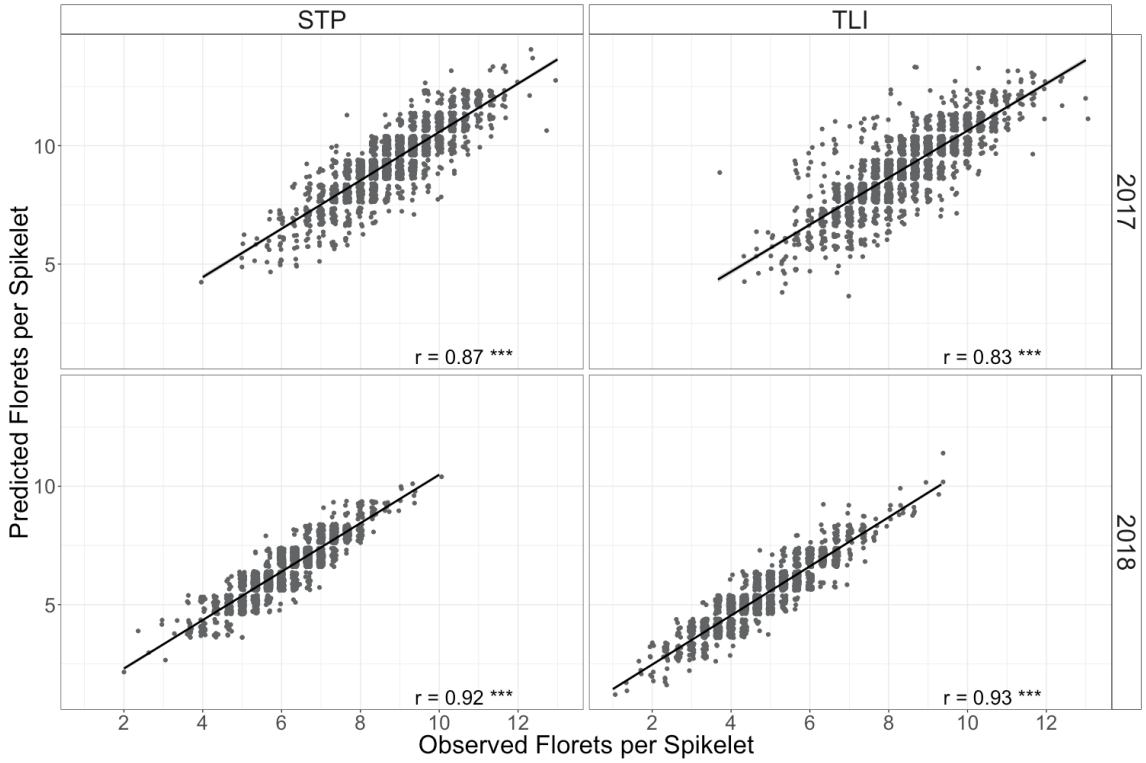
Supplementary Figure 1.4. Predicted spikelets per spike based on a sample size of one, correlated to the observed spikelets per spike, or the average of three spikes across two locations (STP & TLI) in 2017 and 2018. R is the Pearson correlation coefficient and all correlations are highly significant ($P < 0.001$).



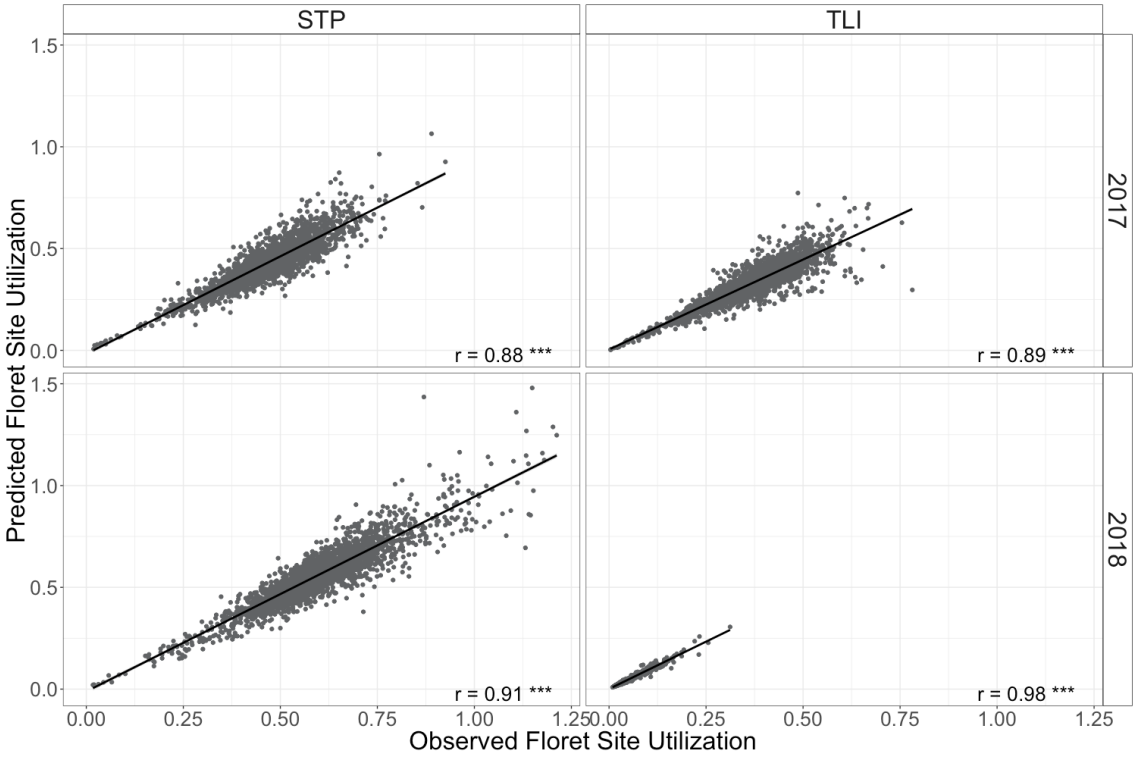
Supplementary Figure 1.5. Predicted spikelets per spike based on a sample size of two, correlated to the observed spikelets per spike, or the average of three spikes across two locations (STP & TLI) in 2017 and 2018. R is the Pearson correlation coefficient and all correlations are highly significant ($P < 0.001$).



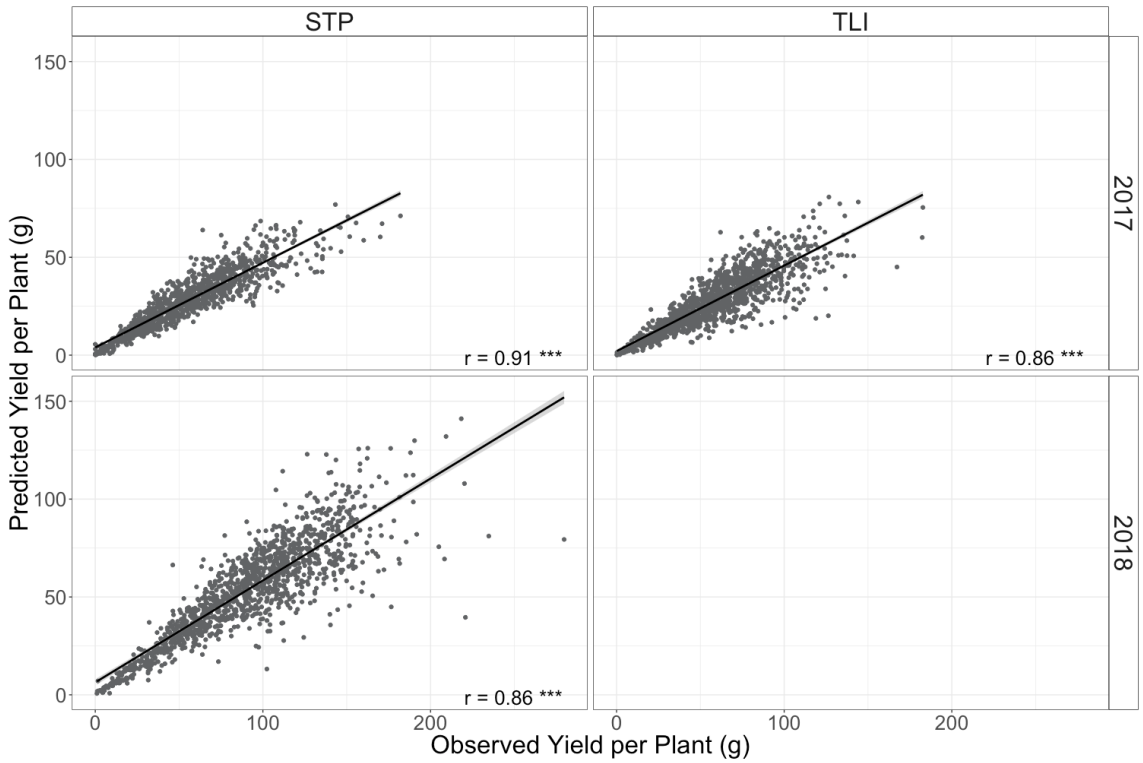
Supplementary Figure 1.6. Florets per spikelet in the bottom, middle and top third of intermediate wheatgrass spikes in two locations (STP & TLI) in 2017 and 2018. Error bars are 95% confidence limits.



Supplementary Figure 1.7. Predicted florets correlated to observed florets per spikelet adjusting for a consistent decrease in floret counts along the spikelet in intermediate wheatgrass spikes in two locations (STP & TLI) in 2017 and 2018. R indicates the correlation coefficient followed by the significance value ($P < 0.001$).



Supplementary Figure 1.8. Floret site utilization calculated using proxies and lower sample numbers correlated to the original calculation in St. Paul, MN (STP) and Salina, KS (TLI) over 2017 and 2018. R indicates the correlation coefficient followed by the significance level ($P < 0.001$).



Supplementary Figure 1.9. Predicted plant yield, calculated by multiplying yield per spike by reproductive tiller count correlated to observed yield per plant in St. Paul, MN (STP) and Salina, KS (TLI) over 2017 and 2018. R indicates the correlation coefficient followed by the significance level ($P < 0.001$). Full plant yields were not recorded at TLI 2018 due to severe drought and lack of seed.

Chapter 2

Nested Association Mapping Reveals the Genetic Architecture of Spike Emergence and Anthesis Timing in Intermediate Wheatgrass (*Thinopyrum intermedium*)

Introduction

Intermediate wheatgrass (*Thinopyrum intermedium;* (Host) Barkworth & D.R. Dewey; IWG hereafter) is a perennial, cool-season grass undergoing direct domestication as a dual-use forage and grain crop for human consumption (DeHaan et al., 2014). IWG is primarily self-incompatible, outcrossing allohexaploid (2n = 6x = 42) (Dewey, 1962). It is native to Europe and Asia, was introduced to North America in 1932 and has since been used primarily as a hay and pasture grass (Dewey 1962; Ogle et al, 2011) and a source disease resistance genes for common wheat (e.g. Brettel et al., 1988; Friebe et al., 1996; Turner et al., 2015). Citing the need for perennial crops to improve agricultural sustainability on highly erodible or marginal lands, the Rodale Institute (Kutztown, PA, USA) in the early 1980s surveyed over 100 perennial grasses for their domestication potential (Wagoner, 1990). They selected IWG due to its relatively large seed size, nutritional similarity to wheat, perennial growth, and its ability to be mechanically harvested. Initial germplasm surveys began in 1987 (Wagoner 1990) and since then, breeding and domestication programs have been established at The Land Institute (TLI; Salina, KS; DeHaan et al., 2018), the University of Minnesota (Zhang et al., 2016), and the University of Manitoba (DeHaan et al., 2014).

IWG breeding programs utilize phenotype-, pedigree- or genotype-based recurrent selection and cultivars are developed as synthetics (DeHaan et al., 2014; Zhang et al., 2016; Bajgain et al., 2020; Crain et al., 2020). IWG is considered to be the most genetically studied of all rangeland grasses (Mortenson et al., 2019) and has considerable genomic resources that have aided breeding and trait dissection efforts. For example, Kantarski et al. (2016) developed a dense consensus genetic map with 10,029 markers from seven full-sib families. Zhang et al. (2016) optimized a protocol for developing genotyping-by-sequencing libraries building on the work of Poland et al., (2012) and Elshire et al., (2011). A reference genome is currently under development by the *Thinopyrum intermedium* Genome Sequencing Consortium (DeHaan et al., 2018; Larson et al., 2019), and genomic selection has become routine in the University of Minnesota and TLI breeding programs (Zhang et al., 2016; Bajgain et al., 2020; Crain et al., 2020).

Breeding and selection have primarily focused on increasing seed size and yield on a per spike basis (DeHaan et al., 2018) and despite the infancy of breeding efforts, gains from selection have been observed. For example, after 5 cycles of selection at TLI, a 143% increase in yield per spike, and 60% in seed mass was predicted in a spaced plant setting (DeHaan et al., 2018). Furthermore, genomic selection has shown great promise in IWG, demonstrating high predictive ability ($r^2 = 0.46$ - 0.67; depending on the trait) and has improved the precision and efficiency of selection (Zhang et al., 2016). Including significant markers from QTL mapping studies as co-factors in genomic selection models improves predictive ability and increases the frequency of favorable alleles (Zhang et al., 2017; Bajgain et al., 2019a; Bajgain et al., 2019b). As a semi domesticated species (i.e. a cultivated crop that has undergone conscious artificial selection; Meyer et al., 2012) breeding targets for this species remain vast. For example, IWG is susceptible to seed shattering (Larson et al., 2019), has low threshability (Zhang et al., 2016), is tall and prone to stem lodging (Frahm et al., 2018), and has low floret site utilization that may contribute to low yields (Altendorf, 2020, Ch. 1). Traits like resistance to bacterial leaf streak and fusarium head blight (Bajgain et al., 2019b) and grain quality for food applications are also of interest (Tyl and Ismail, 2019).

An improved understanding of the genetic architecture of these important traits has the potential to further increase efficiency and precision in IWG breeding efforts through genomics assisted breeding. For example, for Mendelian traits, those controlled by a few large effect genes, it would be advantageous to develop resources for marker assisted selection, allowing breeders to more precisely identify parental material for crossing blocks and increase the frequency of or fix advantageous alleles. By contrast, for more quantitative, polygenic traits controlled by multiple small effect genes, utilizing genomic selection with significant QTL as fixed effects may be a more effective approach (e.g. Spindel et al., 2016). The difficulty lies in the fact that relatively few marker-trait association studies have been conducted in IWG and the genetic control of many important traits remains unknown.

One approach for identifying marker trait associations is bi-parental linkage mapping, in which two divergent parents for a single trait of interest are crossed to

develop a segregating population. The limitations of this method are threefold: 1) mapping resolution is restricted to recombination events that occur between the two parents (Huang & Han, 2014); 2) two parents represent only a subset of the genetic variation in a population (Holland, 2007); and 3) inference of marker-trait associations is generally limited to the trait for which the population was developed. By contrast, genomewide association study (GWAS) or LD-mapping in diverse populations is advantageous because it allows for a greater sampling of genetic diversity and utilizes historic recombination for higher mapping resolution (Huang & Han, 2014). This approach, however, is limited in its ability to detect rare alleles, alleles with low effect, and can lead to false positives arising from underlying population structure (Korte & Farlow, 2013; Huang & Han 2014).

Nested association mapping (NAM) was developed to combine the benefits of both linkage and association mapping by crossing one common parent with a series of diverse donor parents and developing segregating populations in the form of recombinant inbred lines (RILs; Yu et al., 2008). This approach has demonstrated utility and has been used in many crop species including, but not limited to: maize (McMullen et al., 2009; Buckler et al., 2009), rice (Fragoso et al., 2017), wheat (Bajgain et al., 2016; Jordan et al., 2018; Wang et al., 2019), barley (Maurer et al., 2016; Nice et al., 2017; Hemshrot et al., 2019), sorghum (Bouchet et al., 2017), and soybean (Song et al., 2017). Notably, the aforementioned species are all self-compatible where the development of RILs is possible and commonplace. Some of the benefits of this approach are due in part to this ability. For example, generations of selfing in the development of RILs and the associated recombination leads to the breakdown of linkage disequilibrium, increases mapping resolution and essentially eliminates population structure within families (Tian et al., 2011). Inbred parental material can be sequenced at a higher depth and genotypes can be imputed on inbred progeny that segregate for up to two alleles (Tian et al., 2011). Inbred progeny also allow for the reconstruction of parental haplotypes (Hemshrot et al., 2019) and for the assessment of diverse alleles relative to a common genetic background (Cook et al., 2011). Immortal RILs can be maintained in germplasm banks and redistributed for additional study, making the initial investment in population development more

worthwhile. However, there are several practical advantages of NAM that exist irrespective of this feature. First, multiple parents offer more genetic variation and increase the utility of a population by allowing the dissection of more than a single trait. Second, crossing or backcrossing wild or unadapted donor parents with a highly adapted common parent allows for the assessment of diverse germplasm that may not have been possible otherwise due to constraints in phenotyping or growing environment (Poland et al., 2011; Nice et al., 2017). Finally, relative to GWAS in diverse populations, the NAM design offers increased frequency and sampling of rare alleles that may otherwise go undetected or filtered out due to a minimum minor allele frequency (MAF) threshold. A NAM developed from an outcrossing, self-incompatible species may realize the practical advantages of NAM but has its own set of disadvantages. First, immortal lines are not possible, making the population more or less a one-time resource that must be maintained as individual plants. Furthermore, computational programs developed for NAMs and other multi parent populations, including the R packages 'NAM' (Xavier et al., 2015), 'mppR' (Garin et al., 2107), and 'R/qtl2' (Broman et al., 2018), all depend on homozygous parents and progeny to reconstruct haplotypes and phasing and are not suited for heterozygous parents. In many cases, heterozygous sites within the parents and progeny are excluded from the analysis. In an F_1 NAM, instead of two alleles segregating at a locus, bi-parental families from an outbreeding species can segregate from 2-4 alleles, and this may vary from locus to locus (Van Ooijen, 2011). In this case, it is not possible to observe allele combinations in homozygous, identical by descent states as it is in an inbreeding species. Finally, in an F_1 NAM, instead of maximizing genetic divergence between parental material, as would be the case in a traditional NAM, care should be taken to maximize heterozygosity within parents to observe segregation within the progeny. To our knowledge, there have been no formally published NAMs developed in an outcrossing, self-incompatible species, nor for any cool season, perennial grasses (Talduker et al., 2017). Considering the demonstrated utility of the population design and the need within the IWG breeding community to assess the genetic control of many traits of interest, a NAM approach is worth exploring.

Here we introduce an intermediate wheatgrass NAM population and assess its utility by dissecting the genetic control of flowering time. The timing of the transition from vegetative to reproductive phases is a critical step to ensure proper timing for pollination, seed set and dispersal, and variation for this trait has played a critical role in the adaptation of crops to new growing environments (Cockram et al., 2007). Flowering time also plays a role in determining yield in wheat, specifically through mediating the duration of reproductive phases (Guo et al., 2018), and in adapting to stress (Kazan & Lyons 2016). Flowering time in IWG is a target of selection in Kansas where the breeding program selects for early flowering to avoid drought and heat stress in mid-summer (Lee DeHaan, personal communication). As such, the objectives of this work were to: 1) develop and genetically characterize an intermediate wheatgrass NAM population; 2) describe the phenotypic variation in the NAM for flowering time over two years in two distinct locations: St. Paul, MN and Salina, KS; 3) assess the genetic control of flowering time using two approaches: GWAS and linkage mapping within and across multiple populations.

Materials and Methods

Population Development and Establishment

Ten phenotypically diverse genets (donor parents) and one low-shattering genet (common parent) were identified from Cycle 2 of the University of Minnesota breeding program based on their traits of interest (Table 2.1). Parents, and therefore families, were named after their numerical designation within the IWG breeding program preceded with "WGN" for "Wheatgrass NAM." Parental genets were propagated from the field in Fall 2015 into 3-5 clones each, planted in 3.8 L pots, and allowed to re-establish in the greenhouse before vernalization at 4 degrees C for 2 months. After vernalization, plants were placed in a growth chamber (16 hr day, 18-20 degrees C) to induce flowering. In May 2016, multiple reciprocal crosses were made between each donor parent and the common parent by bagging spikes together in custom pollination bags (PBS International, United Kingdom) just before pollen shed. Bags were agitated daily to

maximize pollen flow. Spikes from each cross were harvested in July 2016 and kept separate on the basis of maternal parent identity. Spikes were threshed and cleaned using a belt thresher, sieve (12/64" round; SEEDBURO, Des Plaines, IL) and aspirator (Air Blast Seed Cleaner; ALMACO, Nevada, IA). Approximately 150 seeds from each cross (~75 from each maternal parent) were placed on moist blotter paper (Anchor Paper, St. Paul, MN) in petri dishes and subjected to a cold treatment of 4 degrees C for 3-5 days or until germination occurred. IWG seeds typically do not require cold treatment, but extra care was taken in this case due to the short timeline between harvest and planting. Germinating seeds were transplanted at 1 cm depth into 10 cm 21-count trays set in standard flats with holes. Trays were placed in a misting greenhouse for 4 days and then transferred to an outdoor nursery at the St. Paul Agricultural Experiment Station in St. Paul, MN. Plants were watered daily and fertilized weekly with a standard solution (15 mL per 4.4 L) of 20-20-20 fertilizer, and monthly with slow-release Osmocote (The Scotts Company, Marysville, OH). Plants were clipped to ~5 cm in August to promote tillering, propagated into four ramets per genet and placed into 5 x 5 cm peat pots (Plantation Products, Norton, MA) in Sept. Each genet was labeled with a barcoded tag featuring the following information: family and plant identity, location, and block. Propagules were completely randomized within blocks and were transplanted into spaced plant nurseries in an RCBD with two blocks at the University of Minnesota Agricultural Experiment Station in St. Paul, MN (1-meter centers; STP hereafter) using a mechanical transplanter on 30 Sept, and at The Land Institute in Salina, KS (0.9-meter centers; TLI hereafter) using a jab-type planter on 16 Oct. Several clones of each donor parent (~3 per block) and the common parent (~25 per block) were included for comparison and a two-plant border was established to limit edge effects. Transplants were watered once after transplanting to promote successful establishment. Plots were hand weeded and cultivated with a multivator (Ford Distributing, Marysville, OH) as necessary to control weeds. Additionally, a pre-emergent herbicide, Dual II Magnum (S-metolachlor, Syngenta US), was applied at STP in April 2017 and May 2018 at a rate of 1.75 L ha^{-1}. Herbicides were not used at TLI. At both locations, plots were mowed to 15 cm height

after harvest and fertilized with urea (56.0 kg ha^{-1} at STP; 78.5 kg ha^{-1} at TLI) in fall 2017.

Growing Degree Days

Weather data for St. Paul was obtained from the National Oceanic Atmospheric Administration (https://www.noaa.gov/) from the St. Paul Agricultural Experiment Station (Station ID: USC00218450). Weather data for Salina, KS was obtained from the TLI Weather Station. Growing degree days (GDDs) were calculated in degrees Celsius using the following equation, where T_{max} and T_{min} are the maximum and minimum daily temperatures, and T_{base} is 0°C, or the base temperature for growth used in IWG (Frank, 1996; Jungers et al., 2018):

$$[(T_{max} + T_{min})/2] - T_{base}$$

A maximum threshold of 37°C was set for T_{max}, which is the predicted maximum temperature for growth in wheat (Porter & Gawith, 1999), and GDD accumulation began and ended after five consecutive days where the average daily temperature exceeded T_{base} (Frank & Hoffman, 1989).

Phenotypic Data Collection

Two measures of reproductive growth stage were recorded: spike emergence percent and anthesis stage. IWG spike length is variable both within and among genets, making it challenging to visually estimate the proportion of the spike that has emerged from the boot as is done in a traditional maturity rating scale such as BBCH, Zadoks or Feekes (Lacashire et al., 1991; Large, 1954; Zadoks et al., 1974). Thus, to more precisely capture variation in emergence time, when spikes were approximately 50% emerged on average across the population, the length of the spike emerged from the boot (from the tip of the most apical spikelet to the base of the flag leaf) was measured in cm on one spike per plant. As maturity within a plant can vary, especially in the first year, spikes with larger stem diameter below the base of the spike were chosen for measurement to

minimize experimental error (DeHaan et al., 2018). In cases of high within-plant variability (e.g. ~5 cm), the mean of two spikes was recorded. Harvesting procedures were previously outlined in Altendorf (2020, Ch 1). After harvest, three spikes were aligned end-to-end (from tip of most apical spikelet to base of most basal spikelet, a spikelet defined as having glumes and at least one floret) along a measuring tape and total length was recorded and divided by three to obtain the mean. Length emerged divided by final spike length represents percent spike emergence. In cases where the spike was completely emerged, the length of the emerged peduncle was included, resulting in some estimates exceeding 100%. Early spike emergence is not always coupled with early anthesis in IWG (personal observation). Thus, Feekes flowering time was also recorded for each plant on one occasion per environment when anthers were showing (approx. 1600-1800 h) and when approximately 50% of the plants were in anthesis (Large, 1954). Categories were coded as ordinal variables for data analysis and included: [1] boot stage where no spike is visible; [2] heads emerging or < 25%; [3] heading 25%; [4] heading 50%; [5] heading 75%; [6] heading complete but no anthers visible; [7] beginning flowering, where yellow anthers are beginning to emerge at the center of the spike; [8] flowering 50% ,where anthers are visible through the center and top of spike; [9] flowering 100% where anthers (possibly white or dehiscing) are visible throughout the entirety of the spike including the most basal spikelets; and [10] kernels watery ripe, where anthers are likely dehisced and florets appear plump. The same person recorded anthesis across all locations and years. Dates of data collection events are reported in Figure 2.1.

Phenotypic Data Analysis

All phenotypic data analyses were conducted in R v3.6.1 (R Core Team, 2019). As phenotypic data collection in the second year was technically a repeated measure on the same plants, an initial model with family, location, and year as fixed effects, and genet, the interactions of location by rep, and location by rep by family by genet, as random effects was run to assess interactions of family by environment in the package 'lme4' (Gomez and Gomez 1984; Bates et al., 2015). This analysis revealed highly

significant interactions between family and location, family and year, and location and year (Supplementary Table 2.1). These interactions, differences in plant age between 2017 and 2018, and major differences in climate and weather across the two locations, warranted the separate analysis of all environments (year by location combinations; n = 4). Linear models were fit within each environment for each response variable with fixed terms for genet nested within family and block (n = 2; Supplementary Table 2.2). Estimated marginal means (emmeans) were calculated for genets nested within families using the 'emmeans' package (Bates et al., 2018). Trait correlations were assessed using Pearson's Product Moment correlations. Progeny were separated on the basis of their maternal parent (common or donor) and a t-test was conducted to test for maternal effects. Broad-sense heritability was calculated using a completely random model on a genet mean basis using the following formula:

$$H = \sigma_g^2 / \left(\sigma_g^2 + \sigma_e^2 / r\right)$$

Whereσ_g^2 is the variance of genet (progeny only; parents were excluded), σ_e^2 is the error variance and r is the number of reps or blocks (n = 2). To calculate narrow sense heritability, a separate linear model was fit with family and rep as random effects. Narrow sense heritability is the proportion of additive genetic variance divided by phenotypic variance. In half-sib families, and was calculated using the following equation:

$$h2 = (4 * \sigma_F^2) / [(4 * \sigma_F^2) + (\sigma_e^2)]$$

Genotyping-by-Sequencing

Young leaf tissue was harvested from each genet prior to planting, freeze dried, and genomic DNA was extracted using the BioSprint 96 Plant DNA Kit (QIAGEN, the Netherlands). DNA was quantified using QuantiFluor dsDNA System (Promega Corporation, WI, USA) and normalized to 10 ng/μl. Genotyping by sequencing libraries were developed using PstI/MspI enzymes following the Poland et al. (2012) method

which was optimized for IWG and described by Zhang et al. (2016) and features two unique barcodes per sample. Every 96 samples were pooled, with at least 2 blank wells per plate, creating a total of fifteen 96-plex libraries. The common parent was sampled eight times and the donor parents six each to achieve higher sequencing depth. Libraries were amplified and cleaned using the QIAquick PCR Purification Kit (QIAGEN, the Netherlands), quality control was done using Picogreen (ThermoFisher Scientific, MA, USA), Agilent Bioanalyzer (Agilent, CA, USA) and Kapa qPCR, and subjected to size selection of 160-240bp using PippinHT (3% agarose). Each pool was sequenced in a single lane of a 100 bp single read run on the Illumina HiSeq 2500 HO using v4 chemistry at the University of Minnesota Genomics Center.

Variant Detection & Filtering

Fastq files from the sequencer were demultiplexed using the Barcode Splitter tool from the FastX-Toolkit (unpublished at http://hannonlab.cshl.edu/fastx_toolkit/) where barcodes were matched at the beginning of reads (option: --bol) and no mismatches were allowed (--mismatches 0). Read quality was assessed using FastQC (Babraham Bioinformatics; https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapter sequences, designed according to Poland et al. (2012), were removed using CutAdapt (Martin, 2011). The Quality Trimmer tool from FastX-Toolkit was used to trim reads with a quality score less than 30 (phred+33 scale) and a minimum read length of 30 (-Q 33 -t 30 -l 50 -v). After adapter removal and read trimming, quality was confirmed by re-running FastQC. Bowtie2 (Langmead & Salzberg, 2012) was used to align reads to v2 of the draft IWG reference genome (access provided by The Thinopyrum intermedium Genome Sequencing Consortium), which was indexed prior to analysis using the command: 'bowtie2-build ref.fasta --large-index --threads 20.' Options were set to require reads to align entirely (--end-to-end), and the very-sensitive preset was used, with zero ambiguous reference characters allowed (--n-ceil C, 0, 0). Reads were filtered to include only those that mapped uniquely to the reference, and files were sorted and indexed (option -m 14) using SAMtools (Li et al., 2009). File statistics were calculated before and after filtering using the 'flagstat' option. Within each lane, fastq files with

barcodes corresponding to the same sample were concatenated using a custom BASH script. The Genome Analysis Toolkit (GATK) v4.1.2 (McKenna et al., 2010) was used to call variants, beginning with the HaplotypeCaller tool, which was set to eliminate the duplicate read filter (not applicable for GBS; -DF NotDuplicateReadFilter), run in GVCF mode (-ERC GVCF), and with an expected heterozygosity rate of 0.01. GATK required a dictionary and an 'fai' index of the reference genome; these were created using the GATK 'CreateSequenceDictionary' and the SAMtools 'faidx' commands, respectively. The CombineGVCF tool was used to merge individual sample gvcf files in a hierarchical manner. GVCF files from 18 samples were small in size (~4% of the average), indicating low coverage, and because they caused significant computational delays in the CombineGVCF stage they were removed. The GenotypeGVCF tool was used to perform joint genotyping. The program GNU Parallel (Tange, 2011) facilitated parallel calculations throughout the pipeline. These procedures were conducted using the GBarleyS pipeline described in Neyhart et al., (2019), which is available at https://github.com/neyhartj/GBarleyS. Slight modifications were made to accommodate the two-barcode GBS method used in IWG.

SNP filtering was done using VCFTools v0.1.16 (Danecek et al., 2011) to include only bi-allelic SNPs (--min-alleles 2 --max-alleles 2) with a maximum of 20% missing data (--max-missing 0.8), a minimum allele depth of 5 (--minDP 5), and a minor allele frequency of 0.005 (--maf). A low MAF filter was used to allow rare alleles to segregate in any single family at a rate of 0.5, divided by ten families. Missingness on an individual basis was calculated using --missing-indv, and those that exceeded 70% were removed from the population. The program Cervus v3.0 (Kalinowski et al., 2007) was used to identify unintended outcrosses and selfs using 2,500 markers (the maximum the program can handle) with the lowest percent missing data. Fathers were predicted using the known mothers, then the analysis was rerun using predicted fathers as known and predicting mothers. Selfs or outcrosses (those with unexpected parents resulting from stray pollen or seed contamination) were removed from the population.

Pedigree Relationships Between NAM Parents

Pedigree records from the UMN Breeding Program were used to identify maternal grandmothers and mothers to the NAM Parents. Male parents were identified in Cervus (Kalinowski et al., 2007) using historical GBS sequencing data from UMN Cycle 2 individuals along with data from the NAM parents. TLI Breeding program pedigrees were used to trace the origin of maternal parents to the initial breeding cycles at TLI (Zhang et al., 2016).[1]

Marker Imputation, Population Structure & Linkage Disequilibrium Decay

LinkImpute was used for marker imputation on progeny only (Money et al., 2015). The LD-kNNi method chooses k-nearest neighbors based on linkage disequilibrium between SNPs and is specifically designed to handle data from highly heterozygous species (Money et al., 2015). Imputation accuracy was tested by masking and imputing 10,000 random known genotypes. The program STRUCTURE v 2.3.4 (Pritchard et al., 2000) was used to assess K = 1 through 10, with five replicates each with a length of 25,000 for a burnin period, followed by 75,000 MCM reps. Optimum K values were assessed using Structure Harvester (Dent & vonHoldt, 2012) based on the maximum Delta K value using the Evanno method. The selected K CLUMPP 'indfile' was imported into CLUMPP v1.1.2 (https://rosenberglab.stanford.edu/clumpp.html) to develop an optimal Q matrix over the 5 replicates. Pairwise linkage disequilibrium (LD) using the correlation coefficient r^2 was estimated for pairs of markers within a chromosome and family using the makeGenotypes and LD commands in the 'genetics' package (Warnes et al., 2019). Estimation of LD using genotype data from all families simultaneously proved to yield inaccurate, low numbers. Markers in common with the Round 2 map were used for this calculation to reduce computational load. Pooled R^2 values across all families were plotted over cM distance and the relationships were modeled using a spline approach (Vos et al., 2017) in the 'segmented' package (Muggeo et al., 2019). The decay of LD was estimated when the fitted line intersected with R^2 =

[1] Pedigree analysis was completed by Jared Crain, Kansas State University.

0.2 (Zegeye et al., 2014). Principal components analysis (PCA) was conducted using the 'rrblup' package a.mat function (Endelman, 2011).

Genetic Map Creation

A dense consensus genetic map for IWG was previously developed by Kantarski et al., (2016). Since this map was created prior to the reference genome, GBS tags derived from this study were BLASTed to the reference, and those that matched at a threshold of $1e^{-30}$ were re-named according to their chromosome and base pair position. The number of markers in common between the existing map and the NAM population was determined to be 1,159, with a range of 29 - 77 markers per LG. However, after filtering for informative markers, missing data, and MAF of 0.05 within families, an average of 517 markers per family remained, with a range of 0 - 50 markers per LG. Thus, to obtain more markers, we created a consensus map specifically for the NAM using JoinMap v5 (Van Ooijen et al., 2006). Since parents were genotyped 6-8 times each, the mode call across all samples was selected for the parental genotype at each locus using a custom R script and the 'vcfR' package (Knaus et al, 2018). Families were separated, a minor allele frequency filter of minimum 0.05 was applied, and loci with greater than 20% missing data were removed. Markers were filtered within each family to include only those that segregated as two heterozygous parents (hkxhk), or as one heterozygous and one homozygous (lmxll or nnxnp; Van Ooijen et al., 2006). Parent 1 was always the common parent, WGN59, and Parent 2 the varying donor parent. Genotyping by sequencing data in highly heterozygous species is susceptible to false homozygous calls because it is required that both alleles are annealed with an adapter, amplified, sequenced, aligned, and pass quality filters. JoinMap does not tolerate these low frequency sequencing or call errors (e.g. an nnxnp locus with a progeny genotype pp). Thus, using a custom R script, we calculated the genotype frequency at each locus. If an 'impossible' genotype (pp in this example) persisted at a frequency of less than 0.05, it was assumed to be an erroneous homozygous call and was changed to a heterozygote. If the erroneous calls persisted at a rate of greater than 0.05 at a specific locus, the locus was removed. Using the same custom script, JoinMap .loc files were created on a per

family and per chromosome basis, allowing chromosomes to serve as anchors for grouping markers. An individual project was created in JoinMap for each chromosome, and family .loc files for that chromosome were imported. Within each family, locus genotype frequency was calculated and any markers that displayed a significant level ($\alpha = 0.1$) of segregation distortion were excluded. The groupings tree was calculated and a single group with the highest number of markers was selected at a minimum LOD of 4. Each family group was selected, and consensus maps were calculated using Combine Groups for Map Integration with the Regression Mapping option which conducts three attempts (Rounds) of map creation. More than 250 markers per map proved to be computationally intensive, and thus when more were present, Calculation Options were set to exclude the third round of mapping. If insufficient linkages were detected, the LOD threshold in Calculation Options was lowered by increments of 0.1. To extract phased marker data for each family, all previously excluded markers were unselected, genotype frequency was recalculated and only markers with highly significant ($\alpha = 1.0 \times 10^{-6}$) distortion were excluded. Groups were created using the map node from Round 3 (or Round 2 in cases with 250+ markers) maps, the Maximum Linkages tab was calculated to phase the loci, allowing markers that were initially excluded to phase if they were present on the consensus map.[2] The quality and order of the map was assessed by correlating cM positions of markers shared with the consensus map (Kantarski et al., 2016). It was revealed that several maps were created in JoinMap in reverse order. To maintain consistency across other studies in IWG, maps from these LGs were inverted in JoinMap.

Genomewide Associations

GWAS was conducted using GAPIT software (Lipka et al., 2012) with the default kinship matrix calculation and a minor allele frequency of 0.005. Markers with p-value threshold of 0.00025 (LOD = 3.6) were considered significant. QTL within clusters or peaks were resolved using the mmer function in the R package 'Sommer' (Covarrubias-Pazaran, 2016). Significant markers were fit as fixed effects in a model with a kinship

[2] Genetic map construction was conducted in collaboration with Steve Larson, USDA-ARS.

covariance matrix calculated by the A.mat function as a random effect. Markers clustered on a chromosome within significant QTL peaks were iteratively removed from the model if they were insignificant ($p < 0.001$), and either in LD ($R^2 \geq 0.2$) or within a 21 cM window of a more significant marker. Only SNPs that were significant for more than one unique environment and trait combination, or were in LD or close proximity (using the aforementioned thresholds) to one, are reported. Variance explained by significant QTL was calculated from the model output. Allele frequencies and effects were obtained from the GAPIT output.

Linkage Mapping

Using a custom R script, Round 2 (as recommended in the JoinMap program manual) consensus maps for each chromosome and locus files for each family were appended for use in MapQTL v6 (Van Ooijen, 2009). The traditional cross-pollinated approach in MapQTL yielded excessive singularity errors, likely due to long stretches of insufficiently informative markers in the common parent. Therefore, the two-way pseudo testcross model (TWPT) was used as is suggested in the program manual (Van Ooijen, 2009). This approach is advantageous in that it allows for the assessment of allele effects within parents and has been successfully used to detect QTL in IWG (Larson et al., 2019). The analyses were conducted both within and combined across all families for each location and year (environment, n = 4) separately. For TWPT, map files were created for each parent by excluding any np-type or lm-type markers in one or more families for the common and donor parent, respectively, yielding a total of 42 groups. For .loc files, all hkxhk markers were removed and heterozygous genotypes (lm & np) converted to A and homozygous genotypes (ll & nn) to B. The maximum likelihood mixture analysis procedure was used with maximum 20 iterations, with a maximum number of neighboring markers of 10. For each analysis, a permutation test with 1000 iterations was conducted to determine within-family and combined genome-wide significance thresholds. Each LOD threshold for each family or combined analysis within environment was applied accordingly. Interval mapping was conducted first, and significant LOD peaks, a maximum of one per LG, were selected as cofactors. The

Automatic Cofactor Selection procedure was used to determine the final set of cofactors which were used in a single round of restricted multiple QTL mapping (rMQM). A two-LOD drop off interval was used to establish QTL intervals (Van Ooijen, 2009).

Visualization of Results and Proximity to Orthogenes

Marker position, both physical (Mbp) from the reference genome and genetic (cM) from the NAM consensus map across all linkage groups was normalized and visualized using 'LinkageMapView' (Ouellette et al., 2018). Markers in common between both the physical and the NAM genetic map were connected with a line using the posonleft function. As IWG linkage groups have shown high collinearity with barley (Zhang et al., 2016; Kantarski et al., 2016) known flowering time orthogenes were selected (Supplementary Table 2.3) and aligned to the IWG draft reference genome using BLASTN 2.6.0+ (Altschul et al., 1997) with the online BLAST resource (https://phytozome-next.jgi.doe.gov/blast-search)[3]. Significant ($1e^{-30}$) hits typically corresponded to a known homeologous group (Kantarski et al., 2016). Positions of the orthologous gene hits were plotted. Two-LOD intervals for linkage mapping were plotted from the combined analysis only, and intervals that were significant across traits and years within an environment were collapsed into a single QTL.

Results

Population Establishment

Forty-four spaced plants died between transplanting and the end of the 2017 harvest season at TLI, and 25 additional plants died in the 2018 field season. An additional 203 plants at TLI were deemed not worth harvesting in 2018 due to drought conditions. In St. Paul, 76 plants were lost between transplanting and harvesting in 2017, and an additional 2 in 2018.

Growing Degree Days

[3] BLAST analysis conducted by Steve Larson, USDA-ARS.

Growing degree days (GDD) began accumulating earlier at TLI compared with STP (Figure 2.1). We aimed to collect emergence data when a visual inspection of the field suggested that spikes were on average approximately 50% emerged from the boot, which occurred on June 1 (GDD: 1567) and June 7 (1599) at TLI and June 8 (854) and 7 at STP (794) in 2017 and 2018, respectively. Anthesis notes were taken when plants were approximately 50% beginning anthesis on June 10 (1785) and June 6 (1574) at TLI and June 26 (1219) and June 21 (1111) at STP in 2017 and 2018, respectively. We previously described a severe drought at TLI in 2018, which began towards the end of 2017 and lasted through anthesis in 2018, when there was a maximum 21.7 cm deficit in accumulated precipitation from the ten-year average (Altendorf, 2020, Ch 1). This drought, in combination with high temperatures, resulted in plant stress and highly variable maturity where emergence and anthesis occurred simultaneously in many cases.

Phenotypic Data

Family, location, year, and their interactions were all highly significant for emergence percent and anthesis, providing further support for the separate analysis of unique environments (year by location combinations; Supplementary Table 2.1). In the analysis of unique environments, family and genet nested within family was highly significant ($P < 0.0001$) for both traits in all environments. In every case, genets nested within family explained the majority of variation (avg. 65% and 62% for emergence and anthesis), followed by family (16% and 23%) and rep (avg 1%; Supplementary Table 2.2). Broad and narrow sense heritability estimates were high and averaged 0.73 and 0.46 across environments for emergence and 0.82 and 0.59 for anthesis, respectively (Figures 2.2 & 2.3). Rankings of family means were mostly consistent across environments and traits with WGN46 being the earliest and WGN39 being the latest. The common parent (WGN59) was typically earlier to emerge relative to the other parents, ranking on a scale from early to late 3 of 11 on average across environments for emergence, but ranked in the middle, 6 of 11, for anthesis timing. The maximum difference between parental phenotypes averaged 0.40 for spike emergence (40% of spike emerged), and 1.75 growth stages for anthesis (difference between a few anthers showing and ~75% of anthers

showing). The donor parent with the most divergent phenotype varied depending on the environment and trait. The phenotypic difference between the parents of a family was not a significant predictor of the standard deviation of progeny phenotypes for spike emergence ($r = 0.04$; $P = 0.83$), but it was for anthesis ($r = 0.48$; $P = 0.002$). Furthermore, in every case, the range of progeny phenotypes within families was much greater and exceeded that of the parents, providing evidence for transgressive segregation (Figures 2.2 & 2.3). The average range within a family across environments was 0.98 for emergence (difference of 2% of spike emerged to 100% emerged) and 4.3 growth stages (difference of beginning flowering to kernels watery ripe). Emergence and anthesis were positively associated in all environments (Figure 2.4). In STP 2017 and 2018 where emergence percent data was recorded around 50% emerged, the trend between the two traits was linear ($P < 0.0001$). At TLI in 2017 and 2018, when data was recorded later, around 75-100% emerged, the relationship was best described by a quadratic fit ($P < 0.0001$). In all environments, but especially at TLI, within each category of anthesis stage, spike emergence varied widely, suggesting that early emergence is not always associated with early anthesis. By and large there were no significant maternal effects for either trait, with the exception of one instance, where progeny derived from WGN38 showed a significantly lower anthesis score than progeny derived from the common parent in 3 of 4 environments. Selfs were identified in the population using genomic data after phenotyping was complete. Therefore, we were still able to assess their performance relative to the F_1 progeny. Using the raw phenotypic data (not emmeans) across all environments, values from selfs were compared F_1s using a t-test with unequal variance. Selfs tended to be significantly later for both traits ($P = 2.2 \times 10^{-16}$; Supplementary Figure 2.1).

Marker Data and SNP Filtering

Sequencing generated 3.4 billion reads with 242 M reads per lane on average. With an expected 1.26 million reads per barcode (2.5 million per sample), an average of 94% were recovered with an exact barcode match using FastX BarcodeSplitter. On average across barcodes, CutAdapt removed adapter sequences from 78% of reads and

FastX discarded ~5% that were too short and 0.08% that did not pass the quality threshold. On average across samples, 64.6% of reads aligned to the reference genome, and only those that mapped uniquely (35.6%) were kept for downstream analysis, generating 444,023 SNPs before filtering. Filtering for bi-allelic SNPs, no indels, minimum allele depth, missing data and minor allele frequency reduced the marker count to 8,003 SNPs. Marker count averaged 381 across all chromosomes with an average density of 1 marker every 0.08 Mbp. Two genets were removed because they had greater than 70% missing data. The common parent had a high propensity for sefling with a rate of 8.3% (range of 1.5 - 15.9%, depending on the family), while the donor parents averaged 2.9% (range of 0 - 11.4%, depending on the family). A total of 74 individuals were removed from the final analysis because they were identified as selfs or unintended outcrosses. Final family size on average was 116.8 individuals. Imputation accuracy for LinkImpute was 95.3% with best k = 9 and l = 65. The rate of linkage disequilibrium decay, as defined by $R^2 = 0.2$, varied across linkage groups with a range of 14.5 for LG 20, and 53.5 for LG 18 with a median of 21.08 cM (Figure 2.5). The first two principal components of the genotype matrix explained 22.2 and 15.5% of the variation (Figure 2.6). The distribution of individuals, with the common parent in the center, and progeny distributed approximately mid-way between parents, demonstrated the expected relationship (either half- or full-sibs) between individuals in the population.

Pedigree Relationships

Parents for the NAM were selected based on their divergent phenotypes within Cycle 2 of the University of Minnesota breeding program. Using pedigree records and historical genotype data from the breeding program, we determined multiple shared pedigree relationships between the NAM parents (Supplementary Figure 2.2; Zhang et al., 2016). Importantly, we determined that the common parent, WGN59, was derived from a mating of at least half-siblings (male parents are unknown), leading to a minimum inbreeding coefficient for that individual of $F \geq \frac{1}{8}$. Parent WGN26 is a half-sibling to WGN59 (coefficient of coancestry, $f_{WGN26, WGN59} \geq \frac{1}{8}$). Parents WGN36 and WGN38 are full siblings, and along with WGN15, they share a common grandmother (C3-3471) with

WGN59 ($f_{WGN36, WGN59} = f_{WGN38, WGN59} = f_{WGN15, WGN59} \geq 1/16$). Parent C3-3471 was the first non-shattering, 90% threshability plant derived from The Land Institute's breeding program and was used in numerous crosses in the UMN program and in the creation of the consensus genetic map (Kantarski et al., 2016). We also identified interrelatedness among six donor parents that are half-sibs, sharing either a mother or a father, including: WGN39 and WGN63, WGN07 and WGN46, and WGN45 and WGN55.

Genomewide Association

STRUCTURE results indicated optimal K = 8 (Supplementary Figure 2.3). Results in GAPIT with and without the Q matrix had very similar results ($r = 0.965$) and therefore Q was not used. No additional PCs were added as the model selection feature within GAPIT showed PC = 0 to have the highest BIC for all traits and environments. Using a p-value threshold of -log10(p-value) = 3.6, 35 marker-trait associations were detected for the two flowering time traits across the four environments (Figure 2.7; Supplementary Table 2.4). Five were significant in two trait environment combinations. On average, associations explained a small proportion of the variance, 1.88% for emergence percent, and 3.77% for anthesis. Allele effects varied in their direction, and on average corresponded to a difference of 7.3% spike emerged, and for anthesis, 0.45 fraction of a growth stage. The most QTL were detected on chromosomes 2, 5, 6, 11, and 18. On Chromosome 6, S06_27073072, which was detected in in both STP years for emergence percent and for anthesis in STP 2018, and had a consistently negative allele effect, is 23.9 kb away from a significant BLAST hit for the well-characterized *Ppd-H1* gene in barley that delays flowering time (Turner et al., 2005).

Consensus Map Creation

After filtering for MAF and missing data within families, an average of 3,003 markers per family (average 143 per LG) were considered informative and were input into JoinMap. On average across families, 40% (range: 36 - 43%) of markers displayed significant segregation distortion at $\alpha = 0.1$, and were excluded from the map making step (Supplementary Figure 2.4). A greater proportion of the hkxhk type markers

(average 71%) exhibited distortion, followed by lmxll (28%) and nnxnp (24%). In the case of LG 18, there was an insufficient number of undistorted lmxll (common parent, WGN59) markers to contribute to the map and thus the map for LG 18 only includes nnxnp markers. The final map length was on average 161 cM per LG with a total length of 3,385 cM (Haldane's mapping units) and average marker density of one marker per 0.93 cM (Figure 2.7). Pearson correlation was used to assess the quality of the map order using the positions of markers in common with the consensus genetic map (Kantarski et al., 2016). The average *r* across LGs was 0.94 (Supplementary Figure 2.5). Markers that exhibited distortion in one family, but were included in map making in another, were allowed to phase in the creation of the .loc files for use in MapQTL. The average markers per family in common with the final NAM consensus map was 1,652 per family with an average 78 per LG (Figure 2.5), and after excluding hkxhk markers, which are not used for the TWPT method, there were an average of 994 markers per family and 47 per LG.

Linkage Mapping

In the combined analysis, across all environments, 20 QTL intervals were detected (Figure 2.7; Supplementary Table 2.6). The majority (65%) of the QTL detected in the combined analysis were segregating in the donor parents. In the individual family analyses, the most QTL were detected in families WGN36 (10), WGN15 (10), WGN26 (9) and WGN55 (8) and were most frequently detected on LGs 6 (12 QTL), 17 (8), and 12 (8). A QTL detected in the combined analysis typically overlapped with at least one, but up to 5 significant within-family QTLs. Allele effects and variance explained are only calculated in the within family analyses. Average variance explained for a QTL was 14.2% (range 7.8 - 63%). The significant GWAS marker that mapped close to the *Ppd-H1* gene of barley was positioned on the map at 20.3 cM on LG6, and the QTL markers that clustered on that end had an average peak at 24.9 cM, and these markers were all detected in the donor parents WGN15, WGN26, and WGN55 with a fairly consistently negative allele effect with a few exceptions varying by family. LG17 had the second most significant QTL, with the majority of significant markers clustering between 100 – 136 cM and detected in donor parents as well as in the common parent, specifically in the

analysis of families WGN26, WGN36, WGN38, WGN46 and WGN55. These regions correspond to a physical cluster with several GWAS hits and close to BLAST hits with *Constans 2*, *Constans 5*, and *PHYB* (Figure 2.7). These QTL alleles had a variable effect depending on the family. LG 12 had 8 QTL intervals, many overlapping around a peak of 26 - 65 cM, and detected exclusively in the common parent in several families WGN07, WGN15, WGN26, WGN36, and WGN55, depending on the environment and trait.

Discussion

Intermediate wheatgrass is currently undergoing domestication as a perennial grain crop for human consumption. Selection targets for this crop are numerous and the understanding of the genetic control of important traits remains relatively unknown. Nested association mapping (NAM) as a method for marker-trait dissection has proven useful in other crops, but has not been tested before in a self-incompatible species. Here we present a 10-family F_1 NAM of intermediate wheatgrass developed with phenotypically diverse parents from Cycle 2 of the University of Minnesota breeding program. Considering the importance of variation in flowering time for optimizing yield and performance in new environments, we sought to increase our understanding of the genetic control of flowering time in intermediate wheatgrass and determine the utility of this population for genetic mapping.

Two methods were used to detect QTL for flowering time, including GWAS using GAPIT and two-way pseudo testcross linkage mapping in MapQTL using both a combined and within population analysis and a genetic map created specifically for the NAM. The map was 3,385 cM total, which was slightly smaller than the consensus map in IWG with a total length of 5,061 cM (Kantarski et al., 2016). Significant QTL across the two approaches were relatively consistent in terms of their locations within the genome. QTL detected in the GWAS approach explained a small percent of the variance and typically had low allele effects and percent variation explained, whereas the linkage mapping allowed the assessment of within-family allele effects as well as identified whether the QTL was being donated by the common or donor parent. In almost all cases, regions with several overlapping LOD intervals were consistently either derived from the

common parent or the donor. However, the allele effects of common parent QTL were not detected in every family, suggesting that allele effects may be family-specific. The TWPT approach demonstrated that both the donor parents and the common parents contributed variation for flowering time in the IWG NAM.

Significant regions with QTL aligned closely with orthologs of known flowering time genes in barley, notably the putative ortholog for *Ppd-H1* on Chr06/LG6. Similar to barley, this allele had a mostly consistent negative effect that was clear in the GWAS approach for emergence in STP 2017 and STP 2018, and anthesis score in STP 2018, and in linkage mapping in both STP years for both methods of assessment. The *Ppd-H1* gene appears to have an effect specifically in this environment, as this region was not identified in any instance at TLI. QTL were detected in TLI on LG6, but these were both physically and genetically distant from the *Ppd-H1* region. The QTL on LG17 were detected primarily in the TLI environments in GWAS, but also in STP for linkage mapping, providing further evidence that some QTL are environment specific, which is supported by the vast differences in GDD accumulation between STP and TLI (Figure 2.1). Previous work in IWG has shown that the position of significant QTL can vary depending on whether phenotypic data is tested on an individual environment basis or averaged across environments (Larson et al., 2019; Mortenson et al., 2019). An earlier study of several domestication traits in IWG identified 10 different QTL regions for maturity rating taken at anthesis, several of which were within range of known flowering time genes in Arabidopsis and the Triticeae (Larson et al., 2019), and included QTL on LGs 3, 6, 10, 11, 12, 14, 15, 16, 17, 18. The QTL on LGs 12 and 17 correspond to similar regions as the QTL detected in the present study (Larson et al., 2019). In summary, the population, phenotyping methods and environments tested proved to be successful in identifying both similar QTL regions as previous work in IWG, but also provided substantial evidence for a new QTL corresponding to a likely ortholog of *Ppd-H1* on LG 6.

Two methods of measuring flowering time, emergence and anthesis, were assessed in two distinct environments over two years. The methods exhibited a positive association with one another (Figure 2.4). Recording anthesis can be challenging in IWG

since anthers are on display typically in the evening (~15:00 hrs) and can quickly dehisce with inclement weather such as wind or rain. Once anthers are gone, it becomes very difficult to accurately and efficiently discern between different stages of anthesis. Emergence percent is more time consuming as it involves measuring at least 1 spike in the field and sometimes more in cases with high variability, as well as spike length after harvest. If spike length is already routinely collected, spike emergence may be an easier method for assessing flowering time, but caution should be taken to measure it around 50% emergence so that a linear relationship exists between the two events. Interestingly, more QTL were detected for emergence percent compared with anthesis in both linkage mapping and GWAS.

While the NAM population showed promise for its ability to dissect the genetic control of an important agronomic trait, it also had several challenges. First, the parents were selected initially based on their divergent phenotype, but this alone proved to be an inadequate way to select for genetically divergent parents. The purpose of identifying genetically divergent parents would be to maximize the variation observed in the progeny. However, for the purposes of genetic mapping in an F_1 outcrossing species it is also very important to identify parents that are highly heterozygous so their segregations are effective and can be observed. In the case of this NAM, this was hindered by the fact that the common parent, being a progeny of a cross between at minimum half-siblings likely resulted in increased identity by descent (IBD) and potentially masked recombination. This became apparent in the map making step when certain families and LGs could not be phased, there was severe segregation distortion which reduced the markers that could be used, and limited the number of markers that could be tested within the common parent using the TWPT method. This can also likely be observed in the slow rate of LD decay that was identified across chromosomes. Previous work in IWG has determined LD to decay at $R^2 = 0.20$ at 1 - 5 cM (Zhang et al., 2016; Zhang et al., 2018; Bajgain et al., 2019a), which is much faster than the present estimate of 21 cM. This rate is also much slower than one would expect in an outcrossing species (Flint Garcia et al., 2003). This is due in large part to the population design, where the individuals are all related to one another (as opposed to a traditional, diverse GWAS population where

historical recombination is detectable), the inbred nature of the common parent and the shared pedigree with several of the donor parents (Supplementary Figure 2.1). LG 18 had the slowest rate of LD decay, indicating that in either recombination could not be detected due to homozygosity within the parents, or that recombination rates were very low. Recombination rates have shown to be highly variable across chromosomes (e.g. Bauer et al., 2013). Thus, it is not surprising that LG 18 would not phase for the common parent in JoinMap. Difficulty in creating genetic maps in IWG due to inbreeding has also been reported (Kantarski et al., 2016), and thus it is suggested that future work in IWG ensures high levels of heterozygosity within the chosen phenotypically diverse parents for linkage mapping, which can be achieved through assessing genotype data of the parents. Kantarski et al. (2016) reported segregation distortion in all component maps in the IWG consensus genetic map with presence in variable LGs. One of their maps, C3-3471xS, was derived from self-fertilization of one individual and exhibited rampant distortion across 11 of 21 groups. Interestingly, another population, M26xM35, which was derived from a mating of half-siblings also exhibited high rates of distortion in groups 1, 3, 4, 6, 7, 8, 10, 13, 14, and 20 (Kantarski et al., 2016), but was effective in terms of its mapping potential both using the CP and TWPT approaches (Larson et al., 2019). The populations used in the present population represent a later breeding cycle and thus there have since been additional opportunities for inbreeding and shared pedigree. Kantarski et al. (2016) also reported bias in the segregated markers towards the hkxhk types, which was also found in this case (Supplementary Figure 2.4).

Jungers et al. (2018) developed a model for predicting growth stages in IWG swards in two Minnesota environments, and they estimate that anthesis occurs between 1190-1371 growing degree days (converted from °F to °C). In the present study, 50% anthesis in St. Paul occurred as expected in STP in 2017 at 1219 but slightly earlier in 2018 at 1111 GDD. Personal observation supports this finding, in that IWG plants in spaced plant environments tend to mature earlier than those in swards. Anthesis at TLI occurred on average 15.5 days earlier at TLI compared with STP and at a much higher GDD (1785 and 1574). In IWG, flowering post-vernalization is triggered by long-days (Moore & Moser 1995), which occurs earlier in the year in STP compared with TLI.

However, higher temperatures and corresponding increases in GDDs, as seen at TLI, may have acted synergistically with increasing photoperiod as this has been shown to expedite flowering in grasses (Heide, 1994).

DeHaan et al. (2018) also phenotyped growth stage, and compared with other phenotypic traits, it had a high narrow sense heritability estimate (ranging from 0.29 - 0.67) and a high selection differential (2.6 - 16.5%) that varied depending on the cycle of selection. In the TLI breeding program, early flowering has been selected as a way of reducing heat and drought stress during seed fill. We previously showed that earlier anthesis offered a yield advantage via increased seed set at TLI in 2018 under severe drought conditions (Altendorf, 2020, Ch 1). Earlier spike emergence and anthesis also had a significant, positive effect on yield in STP in the second year (Altendorf, 2020, Ch 1). Using the QTL detected in the present study as fixed effects in genomic selection may be a potential method for altering flowering time in IWG to maximize adaptation to specific environments, or to create more uniformly flowering stands as a way to increase successful pollination. Furthermore, ergot (*Claviceps purpurea*) develops in the unfertilized ovaries of cereal crops, and personal observation suggests this to be the case in IWG as late flowering plants or plants on the edges of fields tend to have higher ergot incidence. In sorghum, increases in pollen viability are associated with increased seed set and lower ergot severity (McLaren, 1997), and fine tuning flowering time may be a way to address a similar problem in IWG.

Tables and Figures

Table 2.1. Parents of the intermediate wheatgrass Nested Association Mapping population, their family sizes separated by maternal parent, and their phenotypic characteristics as recorded in historical breeding program data from St. Paul, MN, which served as the basis of their initial selection.

		Family Size[a] and Maternal Parent			Phenotypic Characterization				
Parent	**Parent Type**	**Common**	**Donor**	**Total**	**Heading Date (1-5)[b]**	**Height (cm)**	**Seed Size (mg)**	**Shattering (0-4)[c]**	**Threshability (1-9)[d]**
WGN07	Donor	59	62	121	3	75	6.83	1	7
WGN15	Donor	60	60	120	5	130	8.88	0	6
WGN26	Donor	62	59	121	3	135	12.52	3	1
WGN36	Donor	51	63	114	3	103	10.1	0.5	4
WGN38	Donor	66	29	95	2	77	9.1	1	0.5
WGN39	Donor	59	63	122	3	131	9.5	3	3
WGN45	Donor	54	64	118	-	120	10.58	0.5	5
WGN46	Donor	58	63	121	-	108	8.28	3	6
WGN55	Donor	61	62	123	-	111	10.48	3	6
WGN59	Common	-	-	-	-	131	10.56	0.5	7
WGN63	Donor	57	56	113	-	126	9.18	3	1

[a] Determined by samples for which there is both phenotypic and genotypic data

[b] Historical data on the heading date was not available for all parents, where one is late and five is early heading

[c] Shattering scale, where zero is low and four is high

[d] Threshability scale, where zero is low and nine is high

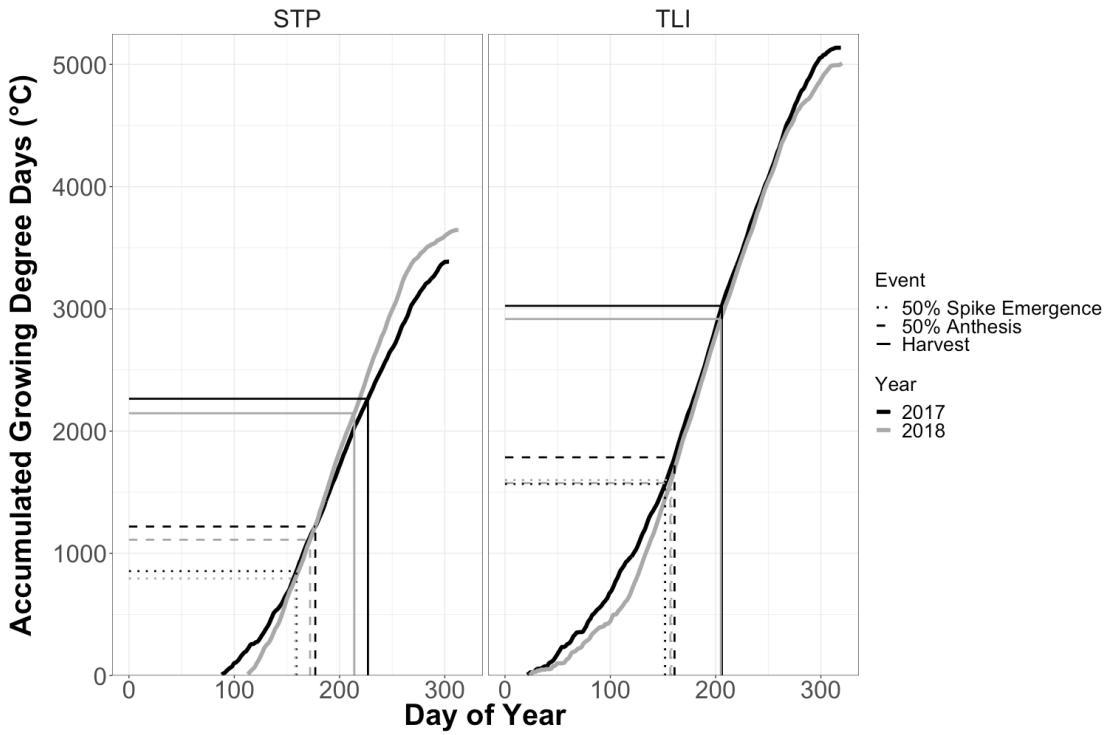


Figure 2.1. Accumulated growing degree days (GDDs) for IWG spaced plants at St. Paul, MN (STP) and the Land Institute in Salina, KS (TLI) for growing seasons 2017 (black) and 2018 (gray). Vertical and horizontal lines indicate timing of major phenological and data collection events including when the plants were at 50% spike emergence, 50% anthesis, and the date at which harvest occurred.

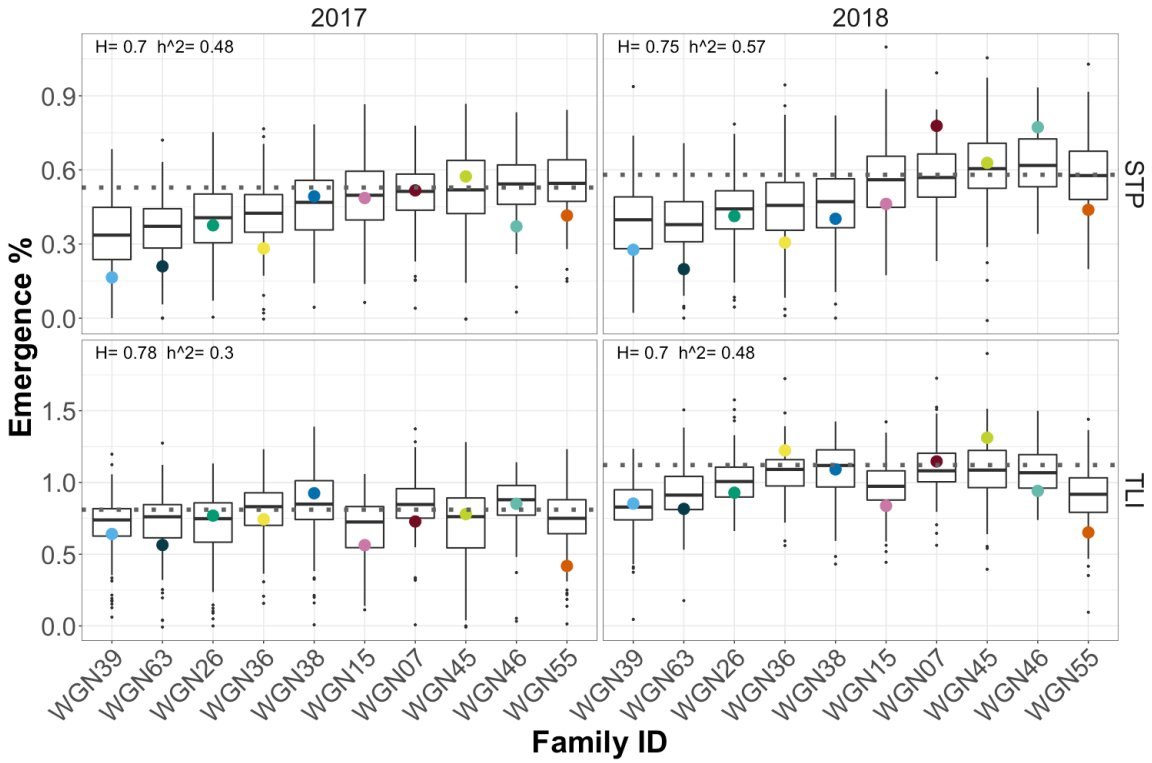


Figure 2.2. Distribution of progeny emmeans within the ten IWG NAM families for emergence percent at St. Paul (STP) and the Land Institute (TLI) in 2017 and 2018. Black horizontal lines within boxplots are progeny means. Horizontal gray dotted line indicates common parent mean and colored dots indicate parent means. Families are ordered based on their ranking for STP 2017.

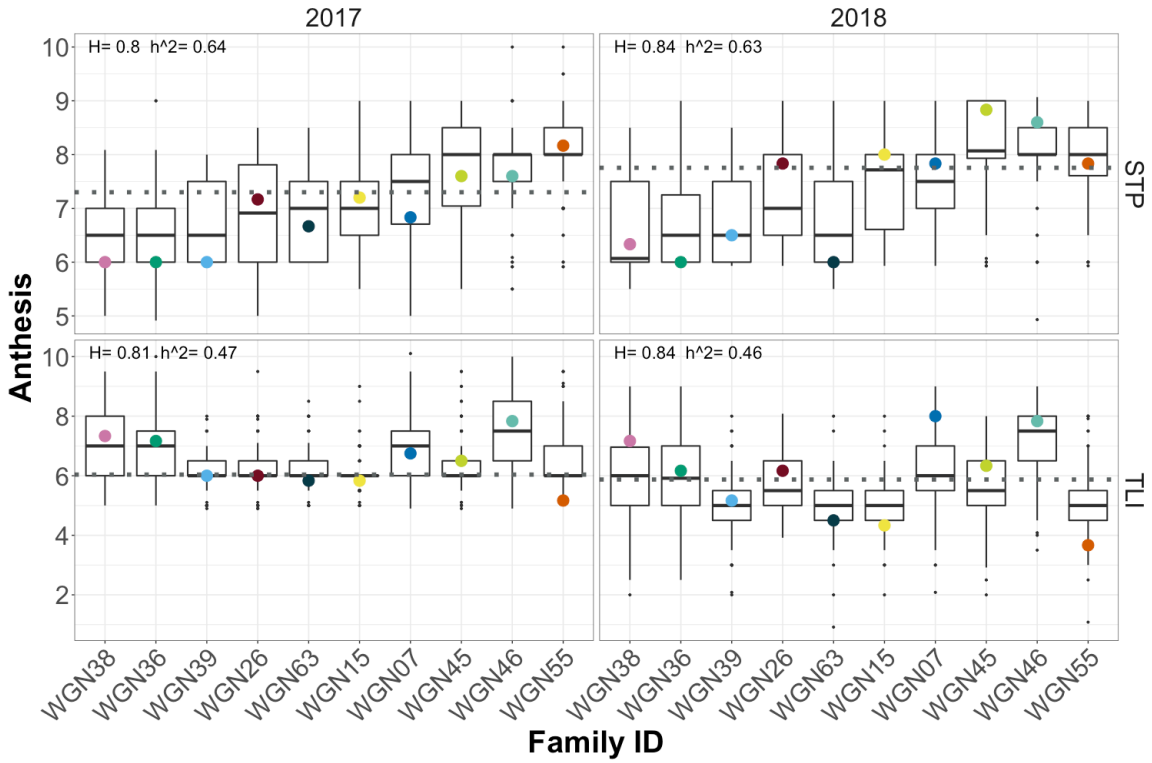


Figure 2.3. Distribution of progeny emmeans within the ten IWG NAM families for anthesis score at St. Paul (STP) and the Land Institute (TLI) in 2017 and 2018. Black horizontal lines within boxplots are progeny means. Horizontal gray dotted line indicates common parent mean and colored dots indicate parent means. Families are ordered based on their ranking for STP 2017.

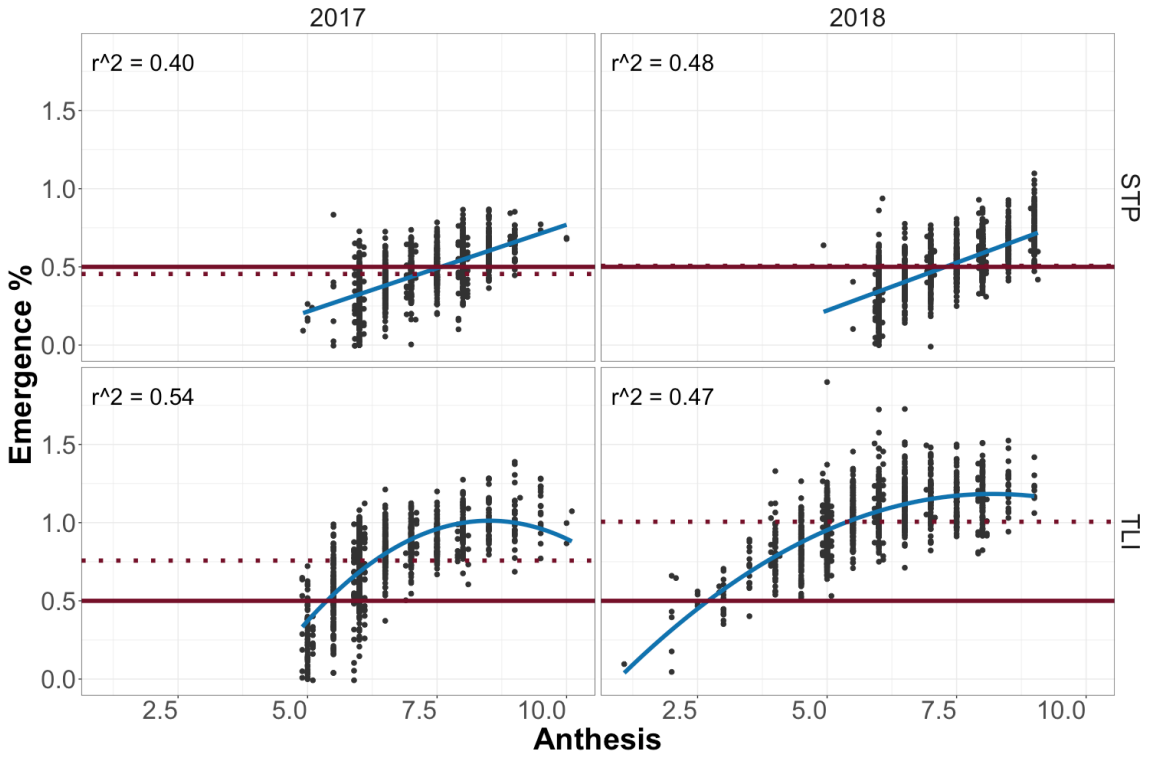


Figure 2.4. Relationships between estimated marginal means for emergence percent and anthesis in the Intermediate wheatgrass Nested Association Mapping population at St. Paul (STP) and the Land Institute (TLI; Salina, KS) in 2017 and 2018. Model fits, linear for STP 2017 and 2018, and quadratic for TLI 2017 and 2018, were highly significant ($P < 0.0001$) and variance explained (r^2) is displayed in the bottom left corner. Horizontal maroon dotted line indicates mean emergence percent for the population to demonstrate the stage at which the data was taken, and the solid maroon line indicates 50% emergence, or the target stage for data collection.

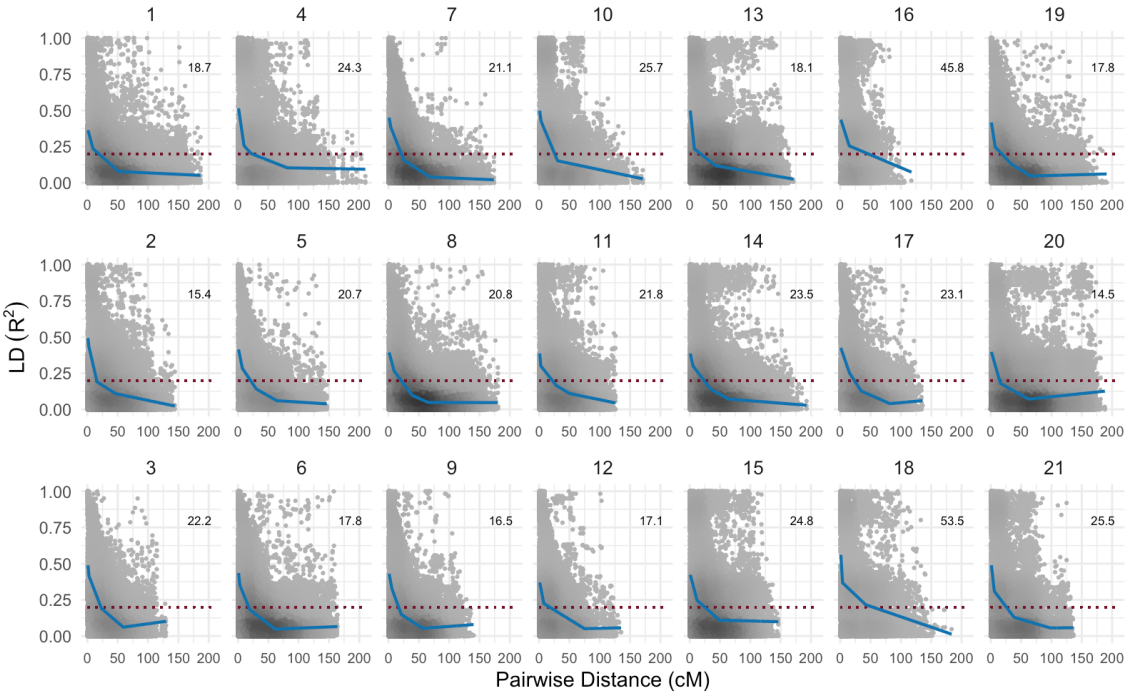


Figure 2.5. Linkage disequilibrium (R^2) across pairwise comparisons of marker distance in cM. Linkage groups are ordered in columns by homeologous groups. Text in top right indicates the cM distance where LD decays below $R^2 = 0.2$ (red line). LD was calculated within each family and linkage group separately and data were pooled for the estimation of decay. Varying shades of gray represents density of the data points (dark gray = higher density).

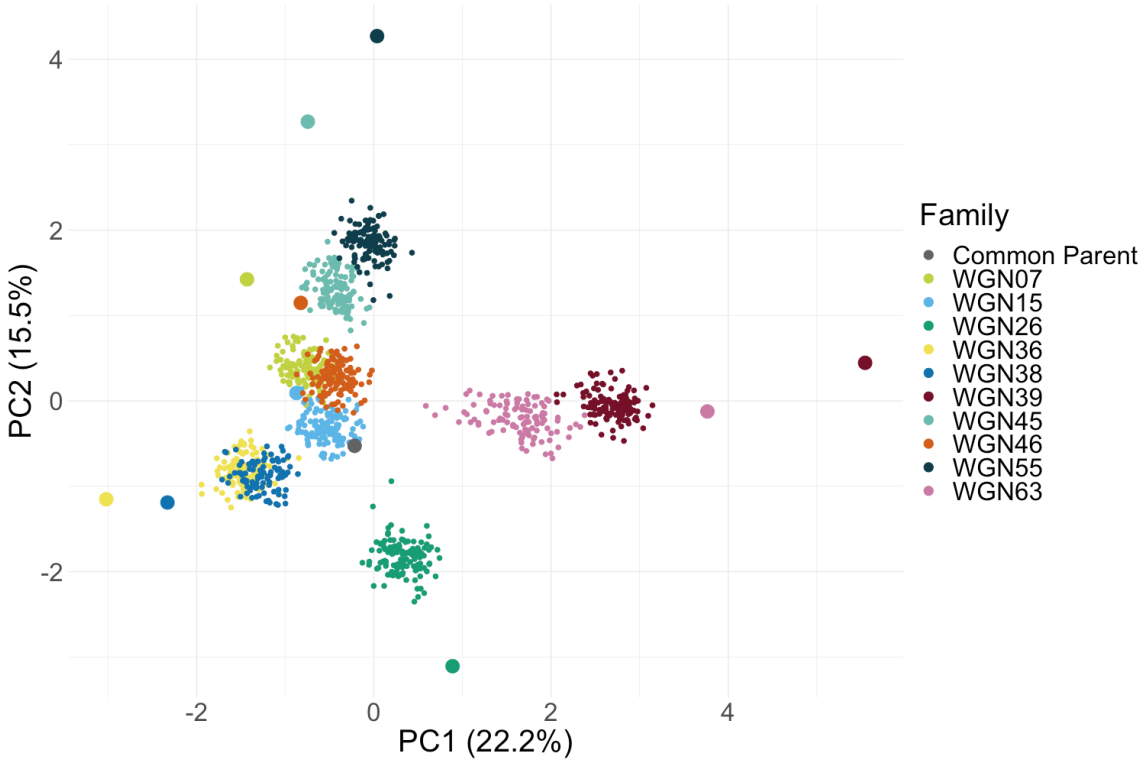
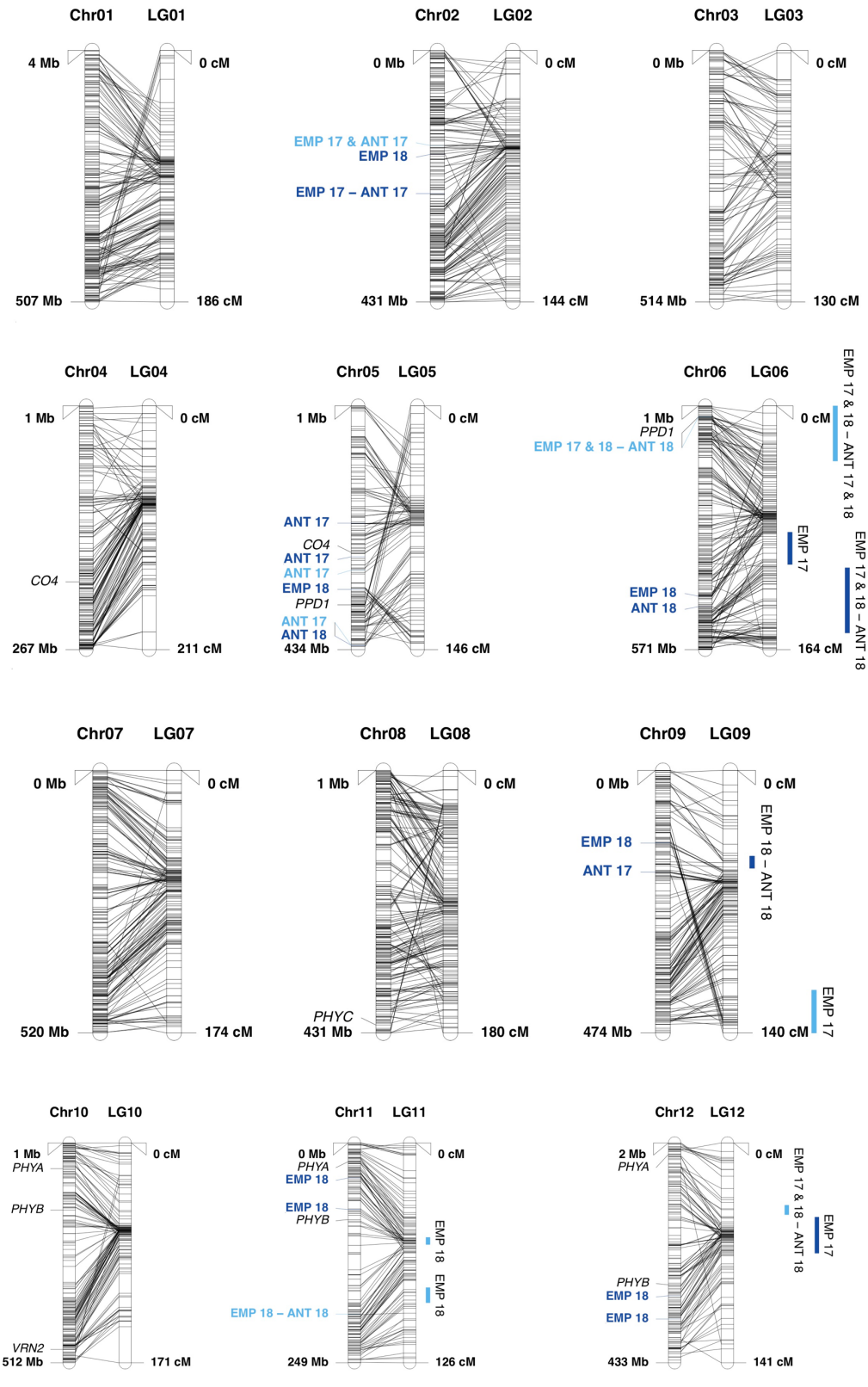


Figure 2.6. Principal components (PC) analysis of the intermediate wheatgrass nested association mapping population families, where large dots indicate parents, small dots indicate progeny and colors indicate family identity. Axis labels include percent variance explained for the two PCs.


Chr01 LG01
4 Mb
0 cM
507 Mb
186 cM
Chr02 LG02
0 Mb
0 cM
EMP 17 & ANT 17
EMP 18
EMP 17 – ANT 17
431 Mb
144 cM
Chr03 LG03
0 Mb
0 cM
514 Mb
130 cM
Chr04 LG04
1 Mb
0 cM
CO4
267 Mb
211 cM
Chr05 LG05
1 Mb
0 cM
ANT 17
CO4
ANT 17
ANT 17
EMP 18
PPD1
ANT 17
ANT 18
434 Mb
146 cM
Chr06 LG06
1 Mb
PPD1
EMP 17 & 18 – ANT 18
0 cM
EMP 17 & 18 – ANT 17 & 18
EMP 17
EMP 17 & 18 – ANT 18
EMP 18
ANT 18
571 Mb
164 cM
Chr07 LG07
0 Mb
0 cM
520 Mb
174 cM
Chr08 LG08
1 Mb
0 cM
PHYC
431 Mb
180 cM
Chr09 LG09
0 Mb
0 cM
EMP 18 – ANT 18
EMP 18
ANT 17
EMP 17
474 Mb
140 cM
Chr10 LG10
1 Mb
PHYA
0 cM
PHYB
VRN2
512 Mb
171 cM
Chr11 LG11
0 Mb
PHYA
EMP 18
0 cM
EMP 18
EMP 18
PHYB
EMP 18
EMP 18 – ANT 18
249 Mb
126 cM
Chr12 LG12
2 Mb
PHYA
0 cM
EMP 17 & 18 – ANT 18
EMP 17
PHYB
EMP 18
EMP 18
433 Mb
141 cM

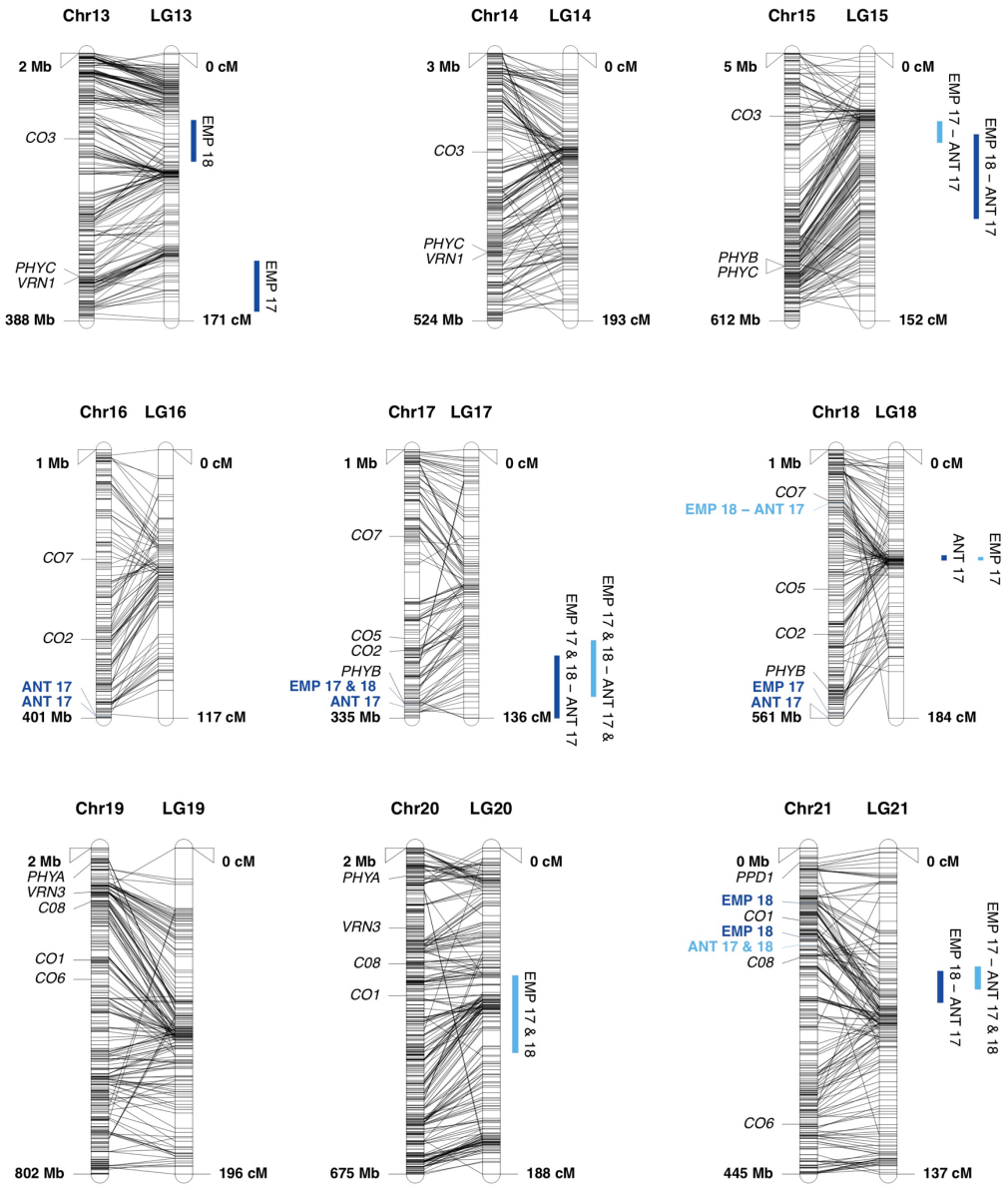
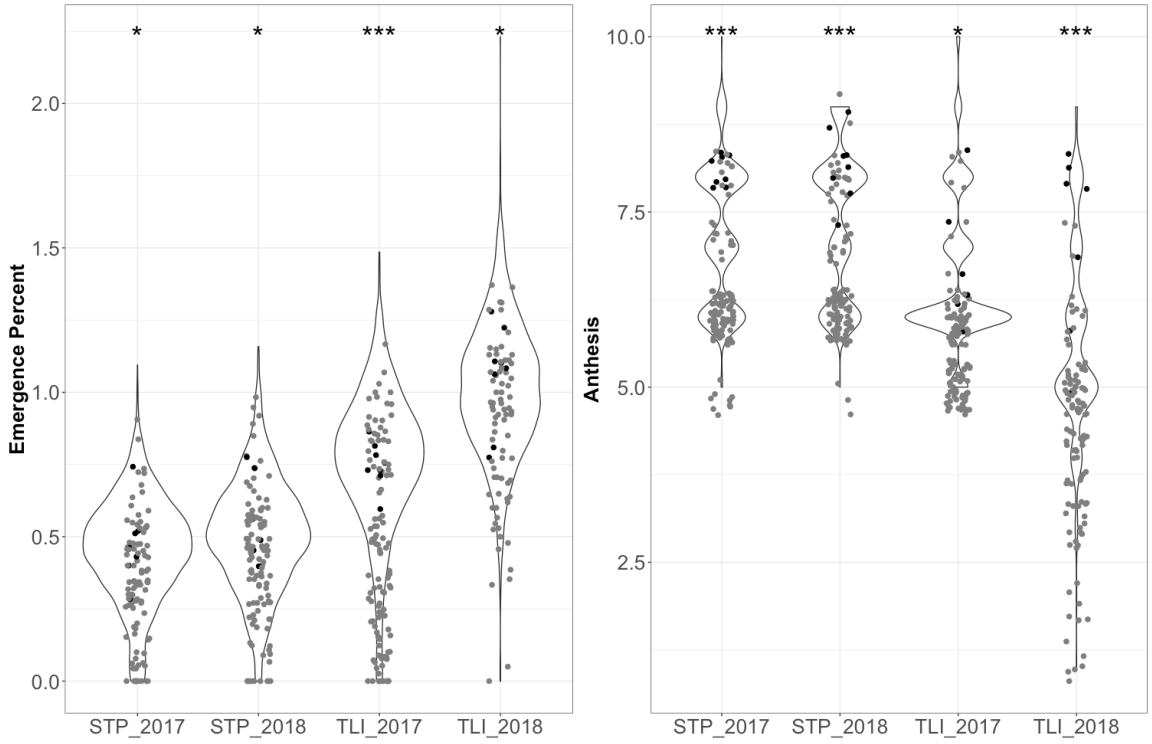
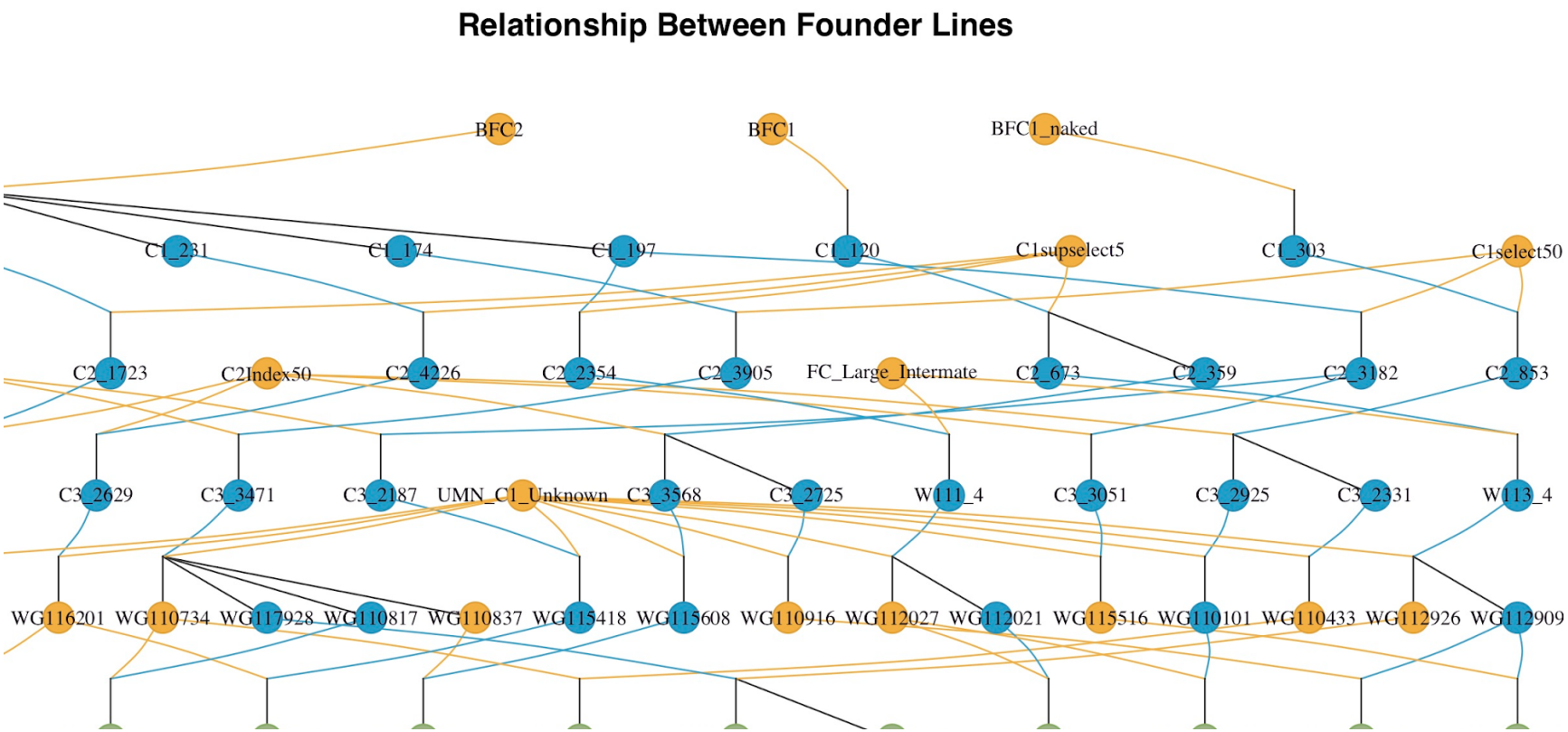


Figure 2.7. Significant markers from GWAS (left), possible candidate orthologous genes (black italic, left) 2-LOD drop off intervals for the combined analysis across populations (bars on right), for STP (light blue) and TLI (dark blue) for emergence percent (EMP) and anthesis (ANT) followed by the years (17 and 18 for 2017 and 2018) in which the marker interval was detected. Markers that were used in both GWAS and linkage

mapping are connected with a black line. Physical distances in megabase pairs (mbp) and genetic distances in centimorgans (cM) are normalized to comparable lengths.

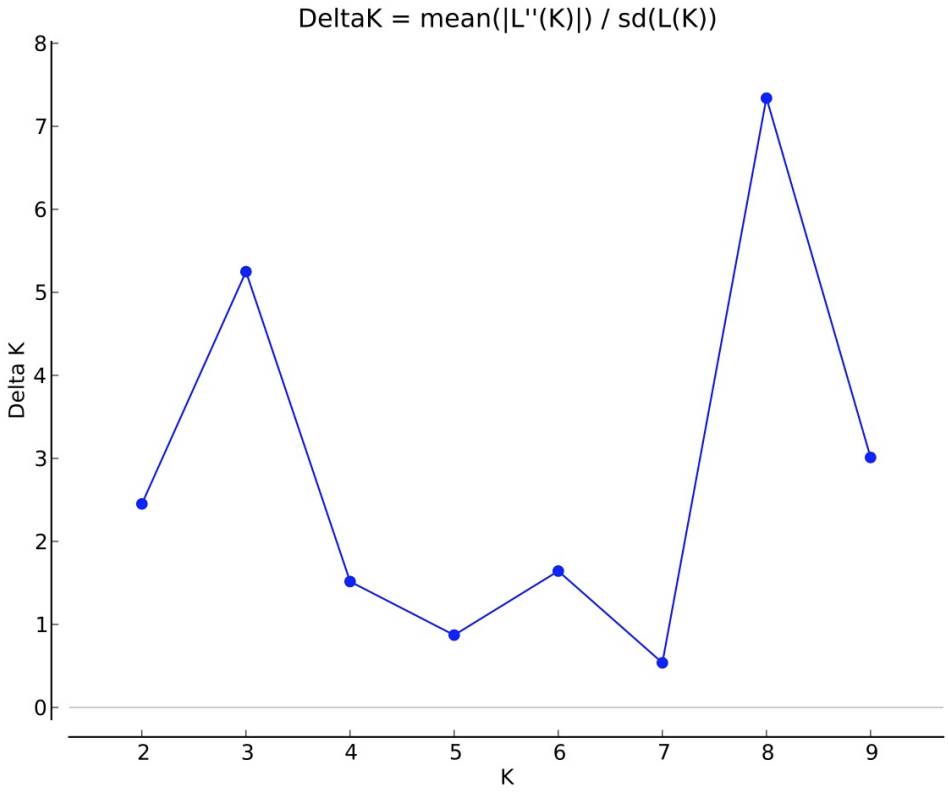


Supplementary Figure 2.1. Performance of selfs (gray dots) and unintended outcrosses (black dots) relative to F1 progeny (distribution represented by the violin plot) for anthesis using the raw data (not emmeans) for both traits across all environments. A "geom_jitter" function was used in ggplot2 to offset data points to increase visibility. Stars at top indicate significance of t-test for differences between selfs and F1 progeny, where 0 = ***; 0.001 = **; 0.01 = *; 0.05 = . ; and not significant = NS.

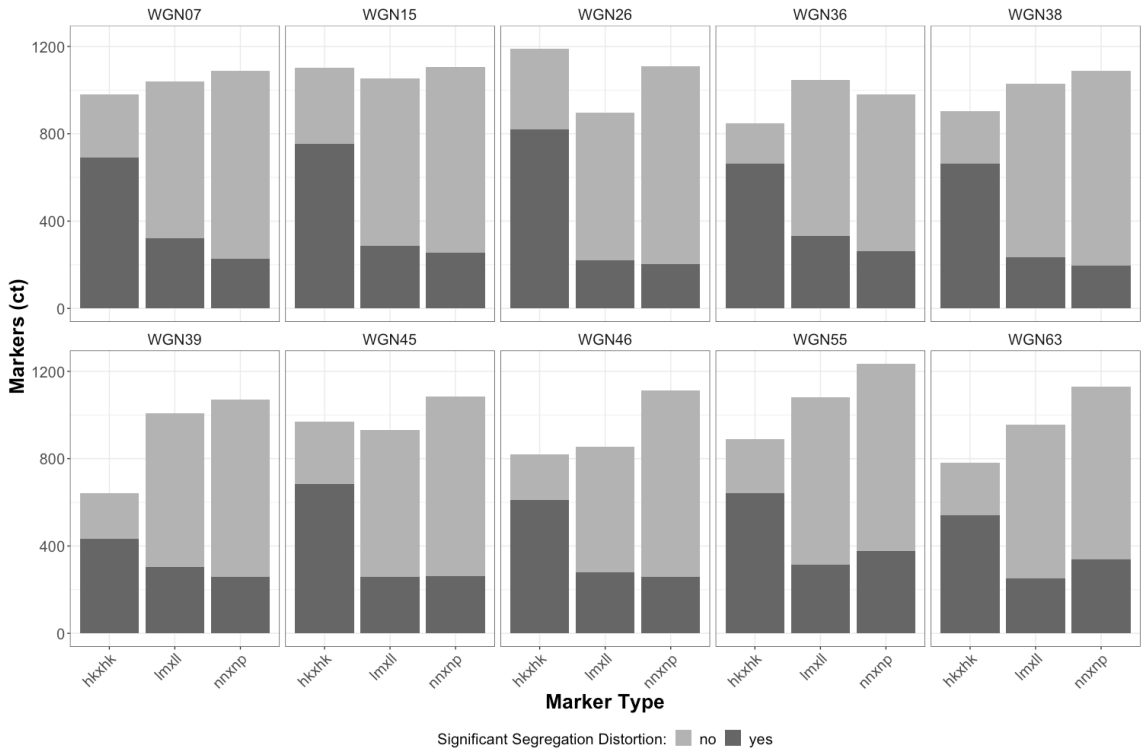


Supplementary Figure 2.2. Pedigree of the eleven Nested Association Mapping founders (green circles) and their male parents (green) and female parents (blue). The common parent is WGN59P20 and all IDs correspond to the Intermediate Wheatgrass Database ID, which are the same except have an additional two digits following (20, 01, or 02). Mothers were traced from pedigree records and

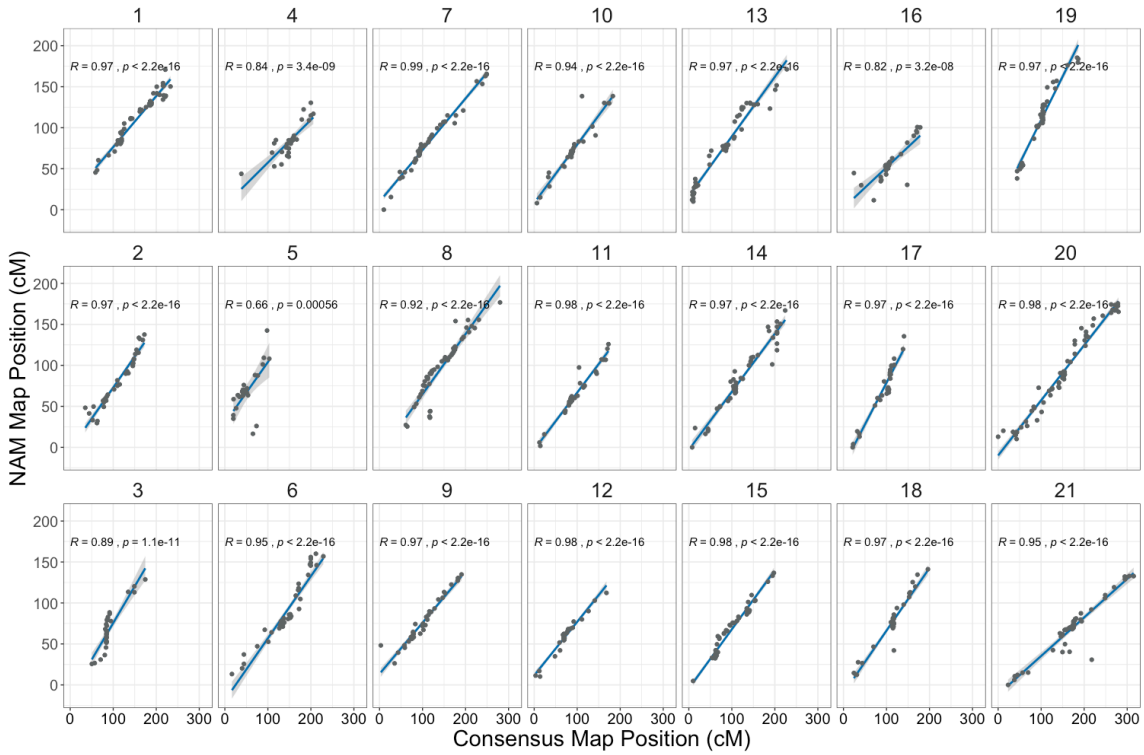
male parents identified through genotypic data using a parentage analysis in Cervus v3.0 (Kalinowski et al., 2007). Parents from UMN Cycle 1 were not genotyped and therefore the male parents were not identifiable and are referred to as UMN_Cycle1_Unknown.



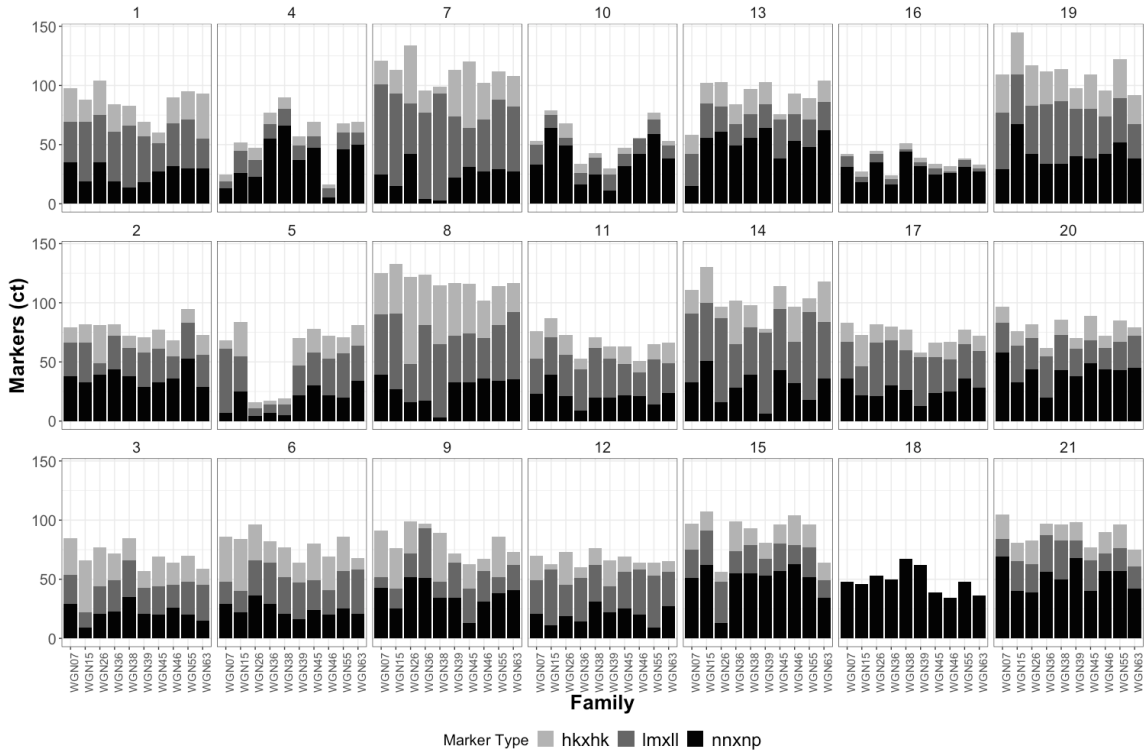
Supplementary Figure 2.3. Delta K for groups k 1-10 from STRUCTURE analysis. Delta K was maximized at k = 8.



Supplementary Figure 2.4. Marker count before filtering and map creation by family and type with gray shades indicating the proportion with significant segregation distortion ($\alpha = 0.1$).



Supplementary Figure 2.5. Correlation of map positions of markers shared (n = 1159) between the nested association mapping population (NAM) and the existing intermediate wheatgrass consensus map (Kantarski et al., 2016) with correlation coefficients and p-value in the upper left corner.



Supplementary Figure 2.6. Marker counts used in linkage mapping by family and linkage group, colored by type. Columns represent homeologous groups.

Supplementary Table 2.1. Analysis of variance results from the mixed effects linear model across all environments for both emergence percent and anthesis score.

Trait	Source	Sum Sq[a]	Mean Sq[b]	NumDF[c]	DenDF[d]	F value	Pr(>F)
Spike Emergence (%)	famID	18.97	2.11	9	4486	100.87	2.74E-172
	loc	5.10	5.10	1	2	244.10	0.00405584
	year	46.66	46.66	1	4478	2233.17	0
	famID:loc	6.15	0.68	9	4455	32.70	3.60E-56
	famID:year	3.50	0.39	9	4478	18.62	8.04E-31
	loc:year	20.20	20.20	1	4478	966.61	2.51E-192
	famID:loc:year	1.71	0.19	9	4478	9.09	9.78E-14
	Source	**Sum Sq**	**Mean Sq**	**NumDF**	**DenDF**	**F value**	**Pr(>F)**
Maturity (1-10)	famID	682.59	75.84	9	4543	148.66	6.29E-247
	loc	89.10	89.10	1	2	174.64	0.00562237
	year	247.97	247.97	1	4609	486.02	1.75E-102
	famID:loc	406.64	45.18	9	4509	88.56	2.66E-152
	famID:year	140.61	15.62	9	4608	30.62	1.67E-52
	loc:year	591.27	591.27	1	4609	1158.90	8.59E-227
	famID:loc:year	42.73	4.75	9	4608	9.31	3.98E-14

[a] Sums of squares
[b] Mean squares
[c] Numerator degrees of freedom
[d] Denominator degrees of freedom

Supplementary Table 2.2. Analysis of variance results from the fixed effects linear model within each unique environment for both emergence percent and anthesis score

Spike Emergence (%)

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	12.4	1.4	97.2	2.95E-133
rep	1	0.0	0.0	3.5	0.061401547
famID:plantID3	1171	42.0	0.0	2.5	3.09E-53
Residuals	1103	15.6	0.0	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	19.1	2.1	155.3	6.97E-190
rep	1	0.2	0.2	12.0	0.000540912
famID:plantID3	1186	46.7	0.0	2.9	1.77E-68
Residuals	1117	15.3	0.0	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	10.9	1.2	64.7	2.05E-95
rep	1	0.2	0.2	13.2	0.000296361
famID:plantID3	1191	87.3	0.1	3.9	2.92E-109
Residuals	1111	20.8	0.0	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	18.6	2.1	92.4	9.69E-124
rep	1	4.4	4.4	196.4	9.51E-41
famID:plantID3	1182	63.3	0.1	2.4	1.03E-43
Residuals	962	21.5	0.0	-	-

Anthesis Score

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	652.8	72.5	216.9	3.41E-238
rep	1	19.3	19.3	57.8	6.07E-14
famID:plantID3	1181	1250.8	1.1	3.2	2.55E-80
Residuals	1119	374.1	0.3	-	-

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	765.1	85.0	273.5	2.10E-278
rep	1	12.6	12.6	40.5	2.89E-10
famID:plantID3	1193	1501.3	1.3	4.0	7.47E-117
Residuals	1149	357.1	0.3	-	-

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	460.9	51.2	138.5	3.84E-175
rep	1	22.7	22.7	61.4	1.07E-14
famID:plantID3	1192	1808.0	1.5	4.1	9.13E-117

Residuals	1121	414.4	0.4	-	-

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
famID	9	1159.8	128.9	256.0	4.84E-265
rep	1	17.7	17.7	35.2	3.95E-09
famID:plantID3	1192	2409.2	2.0	4.0	7.52E-114
Residuals	1125	566.2	0.5	-	-

Supplementary Table 2.3. Flowering time candidate orthogenes and their positions within the IWG v2 draft genome using a BLAST search.

Candidate Gene	Marker Name	Chromosome	Position	Citation
Constans 1	CO1	19	276805349	Griffiths et al., 2003
Constans 1	CO1	21	98669481	Griffiths et al., 2003
Constans 1	CO1	20	306621418	Griffiths et al., 2003
Constans 2	CO2	16	282901202	Griffiths et al., 2003
Constans 2	CO2	17	250837839	Griffiths et al., 2003
Constans 2	CO2	18	385717460	Griffiths et al., 2003
Constans 3	CO3	13	125107091	Griffiths et al., 2003
Constans 3	CO3	14	195239247	Griffiths et al., 2003
Constans 3	CO3	15	147484059	Griffiths et al., 2003
Constans 4	CO4	5	261603082	Griffiths et al., 2003
Constans 4	CO4	4	192966734	Griffiths et al., 2003
Constans 5	CO5	17	236056620	Griffiths et al., 2003
Constans 5	CO5	18	291831437	Griffiths et al., 2003
Constans 6	CO6	21	376610724	Griffiths et al., 2003
Constans 6	CO6	19	324075807	Griffiths et al., 2003
Constans 7	CO7	16	163883113	Griffiths et al., 2003
Constans 7	CO7	17	109119421	Griffiths et al., 2003
Constans 7	CO7	18	107516879	Griffiths et al., 2003
Constans 8	C08	20	240725516	Griffiths et al., 2003
Constans 8	C08	21	149256186	Griffiths et al., 2003
Constans 8	C08	19	133513194	Griffiths et al., 2003
Photoperiod-H1	PPD1	5	354869737	Turner et al., 2005
Photoperiod-H1	PPD1	6	27049222	Turner et al., 2005
Photoperiod-H1	PPD1	21	19461507	Turner et al., 2005
Phytochrome A	PHYA	10	61216437	Szucs et al., 2006
Phytochrome A	PHYA	11	21312497	Szucs et al., 2006
Phytochrome A	PHYA	12	24842223	Szucs et al., 2006
Phytochrome A	PHYA	20	59481522	Szucs et al., 2006
Phytochrome A	PHYA	19	39332590	Szucs et al., 2006
Phytochrome B	PHYB	11	86834965	Szucs et al., 2006
Phytochrome B	PHYB	18	501568393	Szucs et al., 2006
Phytochrome B	PHYB	12	280435707	Szucs et al., 2006
Phytochrome B	PHYB	10	157061869	Szucs et al., 2006
Phytochrome B	PHYB	10	157064095	Szucs et al., 2006
Phytochrome B	PHYB	17	317317668	Szucs et al., 2006
Phytochrome B	PHYB	15	487687325	Szucs et al., 2006
Phytochrome C	PHYC	13	323480455	Nishida et al., 2013
Phytochrome C	PHYC	15	487687247	Nishida et al., 2013
Phytochrome C	PHYC	14	390253842	Nishida et al., 2013
Phytochrome C	PHYC	8	418114556	Nishida et al., 2013
Vernalization 1	VRN1	13	324177081	Fu et al., 2005
Vernalization 1	VRN1	14	391042141	Fu et al., 2005
Vernalization 2	VRN2	10	481014250	Yan et al., 2004
Vernalization 3	VRN3	19	109866574	Yan et al., 2006
Vernalization 3	VRN3	20	166756031	Yan et al., 2006

Supplementary Table 2.4. Significant QTL detected in GWAS, including their alleles, position, minor allele frequency and the families in which the allele segregates above 0.05.

Trait	SNP[a]	Alleles[b]	Chr[c]	Pos[d]	MAF[e]	Segregating Families
Emergence Percent	S02_163115345	T/C	2	163115345	0.11	15, 39, 63
	S02_178305818	C/T	2	178305818	0.16	26, 38, 39, 45, 46, 55, 63
	S02_245307556	C/T	2	245307556	0.02	38
	S05_327412330	A/G	5	327412330	0.05	36, 38, 39
	S06_27073072	T/C	6	27073072	0.44	07, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S06_445976596	G/T	6	445976596	0.34	07, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S09_131144770	G/A	9	131144770	0.25	07, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S11_39753177	G/A	11	39753177	0.05	45, 55
	S11_75489017	A/T	11	75489017	0.16	26, 38, 39, 45, 46, 55, 63
	S11_193858632	G/A	11	193858632	0.04	39, 55
	S12_301921172	C/T	12	301921172	0.03	45
	S12_347517636	G/A	12	347517636	0.13	15, 26, 39, 45, 46, 55, 63
	S17_317374690	C/G	17	317374690	0.36	07, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S18_112460873	G/A	18	112460873	0.04	36, 38
	S18_550223545	G/C	18	550223545	0.16	07, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S21_75063069	C/T	21	75063069	0.2	07, 36, 38, 39, 45, 46, 55, 63
	S21_121306193	A/T	21	121306193	0.02	07, 15, 36, 39, 45, 55
Anthesis	S02_163115345	T/C	2	163115345	0.11	15, 39, 63
	S02_245307556	C/T	2	245307556	0.02	38
	S05_209580093	C/G	5	209580093	0.02	26
	S05_269372528	C/T	5	269372528	0.02	26
	S05_293089678	C/T	5	293089678	0.02	26
	S05_426524999	C/G	5	426524999	0.25	07, 15, 26, 36, 38, 39, 45, 46
	S06_27073072	T/C	6	27073072	0.44	07, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S06_470381731	G/T	6	470381731	0.45	07, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S09_183303350	C/T	9	183303350	0.02	46
	S11_193858632	G/A	11	193858632	0.04	39, 55
	S16_397588073	C/T	16	397588073	0.4	07, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S16_397627839	G/A	16	397627839	0.39	07, 15, 26, 36, 38, 39, 45, 46, 55, 63

S17_321627073	C/T	17	321627073	0.05	36, 38
S18_112460873	G/A	18	112460873	0.05	36, 38
S18_560536334	G/T	18	560536334	0.09	36, 38
S21_131911561	G/C	21	131911561	0.3	07, 15, 36, 38, 39, 45, 46, 55

[a] SNP, single nucleotide polymorphism
[b] Alleles at the SNP including reference and alternate, respectively
c Chr, chromosome
[d] Pos, position of SNP in base pairs
e MAF, minor allele frequency

Supplementary Table 2.5. Results from genomewide association mapping results for all environments organized by trait. Instances in which the QTL was not detected in the present environment are indicated by missing values ("-").

Loc & Year		STP 2017			STP 2018			TLI 2017			TLI 2018		
Trait	**SNP[a]**	**-log10(p)[b]**	**Allele Effect[c]**	**PVE[d]**	**-log10(p)**	**Allele Effect**	**PVE**	**-log10(p)**	**Allele Effect**	**PVE**	**-log10(p)**	**Allele Effect**	**PVE**
Emergence Percent	S02_163115345	4.1	0.08	3.96	-	-	-	-	-	-	-	-	-
	S02_178305818	-	-	-	3.9	-0.05	0.58	-	-	-	-	-	-
	S02_245307556	-	-	-	-	-	-	5.6	0.19	5.78	-	-	-
	S05_327412330	-	-	-	3.8	-0.07	1.53	-	-	-	-	-	-
	S06_27073072	4	-0.03	2.31	8.8	-0.05	3.25	-	-	-	-	-	-
	S06_445976596	-	-	-	6	-0.06	0.9	-	-	-	-	-	-
	S09_131144770	-	-	-	-	-	-	-	-	-	4.4	-0.06	2.27
	S11_39753177	-	-	-	4.1	0.08	0.27	-	-	-	-	-	-
	S11_75489017	-	-	-	-	-	-	-	-	-	3.7	0.05	1.84
	S11_193858632	-	-	-	3.8	0.07	0.63	-	-	-	-	-	-
	S12_301921172	-	-	-	7	0.14	1.07	-	-	-	-	-	-
	S12_347517636	-	-	-	3.8	0.05	0.33	-	-	-	-	-	-
	S17_317374690	-	-	-	-	-	-	4.4	0.05	2.03	3.7	0.04	1.98
	S18_112460873	-	-	-	4.9	-0.09	1.7	-	-	-	-	-	-
	S18_550223545	-	-	-	-	-	-	4.9	0.07	1.74	-	-	-
	S21_75063069	-	-	-	3.7	0.05	2.65	-	-	-	-	-	-
	S21_121306193	-	-	-	-	-	-	-	-	-	3.8	0.1	0.81
Anthesis	S02_163115345	3.6	0.4	3.21	-	-	-	-	-	-	-	-	-
	S02_245307556	-	-	-	-	-	-	4.9	0.8	2.29	-	-	-
	S05_209580093	3.7	0.56	0.52	-	-	-	-	-	-	-	-	-
	S05_269372528	6	0.76	22.67	-	-	-	-	-	-	-	-	-
	S05_293089678	4.6	0.66	9.69	-	-	-	-	-	-	-	-	-
	S05_426524999	4.2	0.24	1.23	4.1	0.25	1.93	-	-	-	-	-	-
	S06_27073072	-	-	-	9.1	-0.31	5.37	-	-	-	-	-	-
	S06_470381731	-	-	-	-	-	-	-	-	-	-	-	-
	S09_183303350	-	-	-	-	-	-	4.5	0.7	1.19	-	-	-
	S11_193858632	-	-	-	4.6	0.44	0.81	-	-	-	-	-	-
	S16_397588073	-	-	-	-	-	-	4.2	0.21	0.78	-	-	-

S16_397627839	-	-	-	-	-	-	3.7	-0.21	2.17	-	-	-
S17_321627073	-	-	-	-	-	-	9.1	-0.76	2.9	-	-	-
S18_112460873	4.2	-0.42	1.78	-	-	-	-	-	-	-	-	-
S18_560536334	-	-	-	-	-	-	6.8	-0.33	0.92	-	-	-
S21_131911561	4.3	-0.33	3.85	3.7	-0.32	2.8	-	-	-	-	-	-

[a] SNP, single nucleotide polymorphism
[b] Level of significance
[c] Allele effect in trait units
[d] PVE, percent variance explained by the QTL

Supplementary Table 2.6. Results from linkage mapping analyses combined across families and within families for emergence percent and anthesis, including two-LOD drop off intervals, peak loci, maximum LOD, variance explained and allele effects.

Trait	Loc[a]	Yr[b]	LG[c]	Map[d]	Left[e]	Peak[f]	Right[g]	Peak Locus[h]	Max LOD[i]	Analysis[j]	r2	μ[k]	α[l]	γ[m]
Emergence Percent	STP	2017	6	DP	0.0	24.9	31.9	-	11.78	combined	-	-	-	-
Emergence Percent	STP	2017	9	CP	117.2	125.1	140.2	-	8.38	combined	-	-	-	-
Emergence Percent	STP	2017	9	CP	110.5	118.7	140.2	CP_Chr09_130411564	3.64	WGN15	8.00	0.48	-0.05	-
Emergence Percent	STP	2017	10	DP	106.0	116.0	99.2	DP_Chr10_394388469	4.32	WGN15	9.00	0.49	-	-0.05
Emergence Percent	STP	2017	12	CP	53.0	65.1	70.6	-	9.75	combined	-	-	-	-
Emergence Percent	STP	2017	15	DP	38.7	49.1	50.8	DP_Chr15_356914772	10.07	combined	-	-	-	-
Emergence Percent	STP	2017	15	DP	51.3	54.1	65.8	DP_Chr15_371353327	5.27	WGN15	12.20	0.49	-	-0.06
Emergence Percent	STP	2017	17	CP	52.5	59.3	82.9	CP_Chr17_18919651	8.03	combined	-	-	-	-
Emergence Percent	STP	2017	18	DP	73.4	75.4	75.6	DP_Chr18_170459443	8.19	combined	-	-	-	-
Emergence Percent	STP	2017	18	DP	87.9	89.8	95.4	DP_Chr18_396163694	3.82	WGN36	10.20	0.42	-	-0.04
Emergence Percent	STP	2017	20	CP	105.1	73.4	87.2	-	8.07	combined	-	-	-	-
Emergence Percent	STP	2017	21	DP	49.8	58.3	59.4	DP_Chr21_87225445	9.23	combined	-	-	-	-
Anthesis	STP	2017	5	DP	17.6	17.6	33.2	-	18.46	combined	-	-	-	-
Anthesis	STP	2017	5	DP	15.0	15.0	33.2	DP_Chr05_353731146	9.75	WGN26	62.80	7.02	-	0.66
Anthesis	STP	2017	6	DP	16.7	20.0	37.2	-	15.79	combined	-	-	-	-
Anthesis	STP	2017	6	DP	0.0	21.0	39.3	-	5.30	WGN55	16.80	8.09	-	-0.27
Anthesis	STP	2017	12	CP	54.5	57.6	69.0	CP_Chr12_261998642	3.55	WGN07	18.10	7.32	0.30	-
Anthesis	STP	2017	15	DP	51.0	51.3	51.9	DP_Chr15_357630002	17.00	combined	-	-	-	-
Anthesis	STP	2017	15	DP	36.4	50.2	56.6	DP_Chr15_201840460	3.46	WGN46	11.60	7.83	-	0.22
Anthesis	STP	2017	17	DP	116.5	124.8	96.4	-	17.43	combined	-	-	-	-
Anthesis	STP	2017	21	DP	54.9	58.6	60.9	DP_Chr21_88073865	19.09	combined	-	-	-	-
Anthesis	STP	2017	21	DP	36.2	58.6	66.1	DP_Chr21_88073865	4.03	WGN07	13.60	7.32	-	-0.32
Emergence Percent	STP	2018	4	CP	134.1	147.3	158.5	CP_Chr04_261391029	3.45	WGN38	8.20	0.47	0.04	-
Emergence Percent	STP	2018	6	DP	0.0	25.1	28.0	DP_Chr06_445976596	19.52	combined	-	-	-	-
Emergence Percent	STP	2018	6	DP	0.0	24.9	32.9	-	5.53	WGN15	14.10	0.54	-	-0.06
Emergence Percent	STP	2018	6	DP	0.0	19.0	32.9	DP_Chr06_29323161	3.57	WGN26	9.30	0.44	-	0.04
Emergence Percent	STP	2018	6	DP	0.0	25.1	28.0	DP_Chr06_445976596	7.58	WGN55	21.40	0.57	-	-0.07
Emergence Percent	STP	2018	8	CP	28.8	31.3	35.4	CP_Chr08_97687643	11.80	combined	-	-	-	-
Emergence Percent	STP	2018	8	CP	12.6	41.6	44.8	CP_Chr08_31449650	4.31	WGN36	9.10	0.47	-0.06	-
Emergence Percent	STP	2018	11	CP	82.8	84.3	91.5	CP_Chr11_223223131	10.02	combined	-	-	-	-

Emergence Percent	STP	2018	11	DP	54.0	54.8	58.2	DP_Chr11_90001198	15.47	combined	-	-	-	-
Emergence Percent	STP	2018	11	DP	61.3	70.5	86.2	DP_Chr11_220432635	3.74	WGN26	9.10	0.44	-	0.04
Emergence Percent	STP	2018	11	DP	49.9	54.8	58.2	DP_Chr11_90001198	8.59	WGN45	23.90	0.64	-	-0.07
Emergence Percent	STP	2018	12	CP	59.1	60.3	61.7	CP_Chr12_262720774	22.00	combined	-	-	-	-
Emergence Percent	STP	2018	12	CP	54.5	59.8	78.8	CP_Chr12_332128678	3.78	WGN07	10.60	0.58	0.05	-
Emergence Percent	STP	2018	12	CP	35.2	46.1	53.6	CP_Chr12_147552391	3.83	WGN15	8.60	0.54	-0.05	-
Emergence Percent	STP	2018	12	CP	25.9	26.9	31.9	CP_Chr12_51058733	4.99	WGN26	12.20	0.44	0.04	-
Emergence Percent	STP	2018	12	CP	50.0	50.9	52.6	CP_Chr12_305515847	5.06	WGN36	11.70	0.46	0.06	-
Emergence Percent	STP	2018	15	CP	111.7	69.9	98.5	-	3.38	WGN46	11.40	0.63	0.04	-
Emergence Percent	STP	2018	17	CP	52.5	59.3	80.0	CP_Chr17_18919651	11.78	combined	-	-	-	-
Emergence Percent	STP	2018	17	DP	111.0	135.7	98.3	-	9.27	combined	-	-	-	-
Emergence Percent	STP	2018	17	DP	110.0	135.7	95.9	-	3.59	WGN46	12.50	0.63	-	-0.05
Emergence Percent	STP	2018	20	CP	106.0	114.0	117.9	CP_Chr20_537718179	8.69	combined	-	-	-	-
Emergence Percent	STP	2018	21	DP	52.7	63.4	65.0	DP_Chr21_208186773	3.70	WGN63	9.60	0.38	-	0.05
Anthesis	STP	2018	2	DP	113.1	122.6	75.7	DP_Chr02_425918121	3.34	WGN45	11.30	8.32	-	0.23
Anthesis	STP	2018	6	DP	18.6	23.0	27.1	-	30.78	combined	-	-	-	-
Anthesis	STP	2018	6	DP	0.0	23.9	34.2	DP_Chr06_42939512	4.88	WGN15	10.90	7.46	-	-0.34
Anthesis	STP	2018	6	DP	0.0	23.9	30.9	DP_Chr06_42939512	5.30	WGN55	13.90	7.96	-	-0.28
Anthesis	STP	2018	12	CP	39.8	42.4	45.9	-	23.99	combined	-	-	-	-
Anthesis	STP	2018	12	CP	39.8	46.1	50.8	CP_Chr12_147552391	6.01	WGN15	12.50	7.45	-0.34	-
Anthesis	STP	2018	12	CP	36.2	45.9	59.1	CP_Chr12_137591505	3.69	WGN55	7.80	7.97	0.24	-
Anthesis	STP	2018	17	CP	50.2	52.2	55.1	CP_Chr17_54933892	19.92	combined	-	-	-	-
Anthesis	STP	2018	21	DP	56.8	58.3	59.2	DP_Chr21_87225445	21.75	combined	-	-	-	-
Emergence Percent	TLI	2017	6	CP	133.5	145.6	152.9	CP_Chr06_507506304	8.18	combined	-	-	-	-
Emergence Percent	TLI	2017	6	CP	127.4	145.6	145.6	CP_Chr06_507506304	4.12	WGN36	12.80	0.82	0.06	-
Emergence Percent	TLI	2017	6	DP	106.8	85.0	99.3	DP_Chr06_396830702	9.64	combined	-	-	-	-
Emergence Percent	TLI	2017	6	DP	125.4	74.7	99.3	DP_Chr06_396830702	3.46	WGN36	14.40	0.82	-	0.07
Emergence Percent	TLI	2017	9	DP	107.5	60.8	73.1	DP_Chr09_355495163	3.62	WGN36	10.10	0.82	-	0.06
Emergence Percent	TLI	2017	12	CP	47.5	50.8	52.6	CP_Chr12_305515836	11.44	combined	-	-	-	-
Emergence Percent	TLI	2017	12	CP	52.6	53.6	58.5	CP_Chr12_253921298	4.77	WGN36	13.10	0.81	0.07	-
Emergence Percent	TLI	2017	13	DP	132.9	143.4	165.3	DP_Chr13_362372796	8.77	combined	-	-	-	-
Emergence Percent	TLI	2017	13	DP	132.9	157.3	168.3	-	5.42	WGN38	19.70	0.87	-	0.10
Emergence Percent	TLI	2017	15	DP	56.6	76.9	93.9	DP_Chr15_428831151	7.96	combined	-	-	-	-
Emergence Percent	TLI	2017	15	DP	55.2	71.9	87.6	DP_Chr15_5517656	5.62	WGN15	20.10	0.68	-	-0.09
Emergence Percent	TLI	2017	17	CP	122.2	130.7	135.7	-	16.86	combined	-	-	-	-

Emergence Percent	TLI	2017	17	CP	108.6	120.9	135.7	CP_Chr17_321627070	4.69	WGN55	15.50	0.74	0.08	-
Emergence Percent	TLI	2017	17	DP	113.0	135.3	135.7	-	12.54	combined	-	-	-	-
Emergence Percent	TLI	2017	17	DP	108.0	135.3	135.7	-	4.69	WGN38	18.90	0.87	-	0.09
Emergence Percent	TLI	2017	17	DP	100.1	119.1	135.7	-	3.55	WGN46	12.70	0.86	-	-0.06
Anthesis	TLI	2017	2	CP	112.5	90.6	98.4	-	8.34	combined	-	-	-	-
Anthesis	TLI	2017	9	CP	112.5	126.3	140.2	CP_Chr09_450543915	2.99	WGN15	10.10	6.07	-0.18	-
Anthesis	TLI	2017	15	DP	107.1	114.0	123.5	DP_Chr15_521971228	3.85	WGN26	12.20	6.30	-	0.28
Anthesis	TLI	2017	17	CP	123.6	133.7	135.7	-	22.51	combined	-	-	-	-
Anthesis	TLI	2017	17	CP	123.4	135.7	135.7	CP_Chr17_324713190	4.03	WGN26	13.80	6.30	0.30	-
Anthesis	TLI	2017	17	CP	111.2	126.7	135.7	-	4.83	WGN55	18.80	6.48	0.43	-
Anthesis	TLI	2017	17	DP	111.0	118.5	135.7	DP_Chr17_318579995	12.27	combined	-	-	-	-
Anthesis	TLI	2017	17	DP	109.0	128.7	135.7	-	4.32	WGN36	16.40	6.94	-	0.44
Anthesis	TLI	2017	18	DP	72.3	73.6	75.7	DP_Chr18_322420106	9.21	combined	-	-	-	-
Anthesis	TLI	2017	21	DP	57.4	62.8	65.0	DP_Chr21_187723404	10.44	combined	-	-	-	-
Anthesis	TLI	2017	21	DP	53.7	64.0	77.2	DP_Chr21_203177933	3.80	WGN26	10.50	6.28	-	0.26
Emergence Percent	TLI	2018	2	CP	10.7	24.5	40.9	CP_Chr02_91989386	3.59	WGN07	11.80	1.10	-0.06	-
Emergence Percent	TLI	2018	6	DP	109.0	115.8	129.4	DP_Chr06_487707179	8.16	combined	-	-	-	-
Emergence Percent	TLI	2018	6	DP	115.8	131.0	85.0	DP_Chr06_487707179	3.51	WGN36	10.30	1.09	-	0.05
Emergence Percent	TLI	2018	6	DP	56.3	75.6	95.5	DP_Chr06_152302233	3.67	WGN46	9.90	1.08	-	0.05
Emergence Percent	TLI	2018	9	CP	133.3	68.4	97.4	CP_Chr09_416113771	3.79	WGN63	14.50	0.92	-0.06	-
Emergence Percent	TLI	2018	9	DP	45.7	47.6	52.3	DP_Chr09_81355774	9.57	combined	-	-	-	-
Emergence Percent	TLI	2018	13	CP	42.9	59.2	69.4	CP_Chr13_104225517	9.20	combined	-	-	-	-
Emergence Percent	TLI	2018	13	CP	35.6	59.2	69.4	CP_Chr13_104225517	3.44	WGN26	8.30	1.00	0.05	-
Emergence Percent	TLI	2018	13	CP	147.0	37.6	96.1	-	4.22	WGN55	30.50	0.91	0.11	-
Emergence Percent	TLI	2018	14	DP	118.9	148.1	97.5	-	3.57	WGN38	13.50	1.10	-	-0.07
Emergence Percent	TLI	2018	17	DP	104.1	125.8	135.7	-	7.81	combined	-	-	-	-
Emergence Percent	TLI	2018	17	DP	102.1	115.5	135.7	-	4.94	WGN46	16.50	1.07	-	-0.07
Emergence Percent	TLI	2018	21	DP	51.7	54.1	65.0	DP_Chr21_90120652	7.78	combined	-	-	-	-
Emergence Percent	TLI	2018	21	DP	43.8	54.1	63.3	DP_Chr21_90120652	4.02	WGN36	13.70	1.08	-	0.06
Anthesis	TLI	2018	6	DP	110.7	115.8	151.8	DP_Chr06_487707179	8.56	combined	-	-	-	-
Anthesis	TLI	2018	6	DP	121.1	143.8	143.8	DP_Chr06_510359842	4.18	WGN26	17.80	5.81	-	0.46
Anthesis	TLI	2018	6	DP	0.0	28.0	73.2	DP_Chr06_13655058	3.86	WGN55	10.30	5.12	-	-0.34
Anthesis	TLI	2018	9	DP	42.9	44.8	53.0	DP_Chr09_25220856	8.57	combined	-	-	-	-
Anthesis	TLI	2018	9	DP	33.8	45.7	57.5	DP_Chr09_165133955	3.80	WGN07	13.00	6.25	-	0.42
Anthesis	TLI	2018	14	DP	114.3	127.3	142.1	DP_Chr14_373735432	7.81	combined	-	-	-	-

Anthesis	TLI	2018	14	DP	116.6	130.3	142.1	-	3.40	WGN15	14.50	5.05	-	-0.29
Anthesis	TLI	2018	15	DP	46.1	49.3	50.8	DP_Chr15_193256422	9.59	combined	-	-	-	-
Anthesis	TLI	2018	17	DP	118.5	123.8	95.9	DP_Chr17_318579995	9.11	combined	-	-	-	-
Anthesis	TLI	2018	19	DP	78.1	82.5	87.2	DP_Chr19_540414045	8.97	combined	-	-	-	-
Anthesis	TLI	2018	19	DP	71.9	81.5	88.2	DP_Chr19_522394713	4.78	WGN63	17.70	5.01	-	-0.27
Anthesis	TLI	2018	21	DP	54.9	63.3	78.0	DP_Chr21_208887568	3.26	WGN63	10.20	4.96	-	0.21

[a] Loc, location
[b] Yr, year
[c] LG, linkage group
[d] Map used in linkage mapping, DP for donor parent, or CP for common parent
[e-g] Genetic positions in centimorgans (cM) of the 2-LOD drop-off interval
[h] Locus at the LOD peak where "-" indicates the peak was in an interval
[i] Maximum LOD significance value
[j] Analysis whether the QTL was detected in the combined (including all families) analysis or within the analysis of a specific family. Allele effects are not reported in the combined analysis
[k] Overall mean
[l] Difference between alleles in the common parent
[m] Difference between alleles in the donor parent

Chapter 3

Nested Association Mapping Reveals the Genetic Architecture of Domestication Traits in Intermediate Wheatgrass (*Thinopyrum intermedium*)

Introduction

Perennial cropping systems have the potential to provide extensive environmental service benefits compared with annual cropping systems (Glover et al., 2010 & Pimentel et al., 2012). The challenge is that approximately 69% of the world's agricultural land is dedicated to the production of annual crops, including cereal grains, legumes and oilseeds (Cox et al., 2006). There are few viable perennial alternatives for these crops, as herbaceous perennials were not domesticated during the first agricultural revolution (Van Tassel et al., 2010). Today, plant breeders are relying on intense, conscious artificial phenotype and genotype-based selection to develop crops to fill this void (DeHaan et al., 2004; DeHaan et al., 2014; Van Tassel et al., 2010). There are two distinct methods for developing new perennial crops. The first is interspecific hybridization, in which a cultivated annual species is crossed to a related perennial and attempts are made to introgress perenniality. The second is direct domestication, where selection for domestication traits takes place within an undomesticated perennial species (Cox et al., 2006; Kantar et al., 2016). In the case of direct domestication, much emphasis is placed on fixing domestication syndrome traits, which are part of a series of traits that distinguish most cultivated species from their wild progenitors (Harlan et al., 1973). These broadly include characteristics associated with improved establishment, such as reduced dormancy and larger seeds, and increased harvest efficiency such as reduced shattering and improved threshability (i.e. ease of removal of grain from the lemma and palea; Purugganan and Fuller, 2009).

One species currently undergoing direct domestication that has generated considerable interest is the perennial grass species intermediate wheatgrass or *Thinopyrum intermedium* (2n = 6x = 42; Barkworth & D.R. Dewey; IWG hereafter). Grain from improved lines of IWG is being marketed under the name Kernza, a trademark owned by the Land Institute (DeHaan et al., 2017). IWG is a cool season, rhizomatous, outcrossing and largely self-incompatible grass native to Europe and Asia (Dewey, 1962). It was introduced to North America in 1932 and has since been used extensively as a hay and pasture grass (Ogle et al., 2011). In 1988, the Rodale Institute (Kutztown, PA) selected IWG from a panel of nearly 100 perennial grasses for its

domestication potential as it had relatively large seeds with nutritional similarity to wheat, vigorous perennial growth, and could be mechanically harvested (Wagoner, 1990). Breeding programs, using both phenotype-and genotype-based recurrent selection, have been established at the Land Institute (Salina, KS; TLI hereafter), within the University of Minnesota's Forever Green Initiative, and the University of Manitoba (Cox et al., 2010; DeHaan et al., 2014; Zhang et al., 2016).

In contrast to ancient plant domestication events, modern day efforts are aided by an extensive body of literature generated over the last century that has sought to understand precisely where, why, and how annual crops cereals like maize, rice, barley and wheat, and others, were domesticated. Plant domestication is studied to gain understanding of human history, artificial and natural selection, and the co-evolution of humans and plants. Most relevant in this case, however, is the applicability of this knowledge to new breeding efforts (Doebley et al., 2006). These resources are especially advantageous in the case of IWG, as it is a member of the Poaceae family and Triticeae tribe, where gene order and content among species are highly conserved (Moore et al., 1995), a phenomenon recently demonstrated specifically between IWG, barley (*Hordeum vulgare*), and *Aegilops tauschii* (Kantarski et al., 2016; Zhang et al., 2016). In addition to what can be considered a blueprint for plant domestication, breeders also have access to affordable DNA sequencing, methods for trait discovery, high throughput phenotyping methodology, and precise, efficient and proven methods for selection.

Selection efforts thus far in the development of IWG have focused on improving grain yield on a per spike basis and seed size (DeHaan et al., 2018), but targets remain vast, and include a variety of domestication and improvement traits. The domestication traits needing immediate improvement include shattering resistance, threshability, reduced height, increased seed size, uniform flowering time, and reduced tillering (DeHaan et al., 2018; Zhang et al., 2017; Bajgain et al., 2019a; Larson et al., 2019; Altendorf, 2020, Ch 2). Improvement traits such as baking quality and disease resistance are also of interest (Bajgain et al., 2019b; Tyl & Ismail, 2019).

Shattering is the most significant of the domestication syndrome traits as it marked the transition of a plant to dependence upon humans for seed dispersal and

therefore reproductive success (Purugganan and Fuller, 2009). Non-shattering was one of the first traits to be selected in modern domestication of wild rice (Hayes et al., 1989; Kennard et al., 2002) and is suggested as an important criteria for selecting new candidate species for domestication (DeHaan et al., 2016). In perennial grasses, shattering can result in yield losses, narrow the harvest timing window, and lead to stand contamination in perennial or non-rotational crops (Berdahl & Frank, 1998; Mueller-Warrant & Rosato, 2002). IWG has a spike inflorescence and seed shattering can occur in two distinct ways (Larson et al., 2019). The first, which will be referred to as floret shattering, is through disarticulation between florets within the rachilla axis, which is the secondary axis of the spike. In this case, the dispersal unit is one or more caryopses enclosed in a lemma and palea (Figure 3.1A). The second is wedge-type brittle rachis, sometimes also referred to as spikelet shattering, where the disarticulation event occurs within the rachis above the spikelet, and the dispersal unit is a spikelet with a rachis segment attached, or a series of spikelets still attached to another portion of the rachis (Figure 3.1B). Floret shattering occurs in about half of the species in the Poaceae (Sakuma et al., 2011), whereas brittle rachis is unique to species in the Triticeae tribe (Pourkheirandish et al., 2015). It is speculated that floret shattering may be a more ancient form more commonly associated with grasses with panicle inflorescence structure (Pourkheirandish et al., 2015). In barley, the brittle rachis trait is controlled by two genes, *Btr1* and *Btr2* on chromosome 3H, where a mutation at either locus confers a non-brittle rachis (Pourkheirandish et al., 2015). In wheat, the recessive, non-brittle rachis conferring alleles *Br-A1* and *Br-B1*, located on the short arm of chromosomes 3A and 3B, were fixed in domesticated emmer (*Triticum dicoccum*) (Watanabe et al., 2002).

Selection for free-threshing was likely driven by an increase in efficiency. For example, in emmer wheat, the non-brittle rachis and free-threshing traits together conferred an 85% reduction in time required for threshing (Tzarfati et al., 2013). The mechanism of free threshing ability varies depending on the species. For example, in non-free threshing wheat, the caryopses are enclosed in tough, tenacious glumes that are difficult to separate from the spikelet to release the kernel (Kerber & Rowland, 1974). Whereas in barley, the lemma and palea are adhered to the hull and it can also be difficult

to separate the awns and rachis segments from the spikelet (Schmalenbach et al., 2011). In barley, this trait is controlled by the locus *nud* on chromosome 7HL (Taketa et al., 2008) and in diploid wheat via soft glume (*sog* on $2A^{m}S$) and polyploid wheat via tenacious glume (*Tg* on 2DS) (Sood et al., 2009) and the domestication gene *Q* on 5A (Simons et al., 2006).

Grain yield in IWG grown in sward environments declines with age (Jungers et al., 2017) and changes in reproductive tillering dynamics have been suggested as a potential cause (Frahm et al., 2018). For example, under highly competitive conditions, IWG reduces resource allocation to sexual reproductive structures in favor of vegetative growth (Hunter et al., 2020). When given adequate space, such as in a spaced plant nursery, IWG tillers aggressively and yield component traits decrease from the first to the second year (Altendorf, 2020, Ch 1). Genotypes that tiller less and allocate more biomass to floret fertility and seed fill may be a potential ideotype for grain-producing IWG (Altendorf, 2020, Ch 1).

Floret site utilization, or the percent of florets that produce economically viable seed (Elgersma, 1985) was shown to be a major limiting factor in IWG yield in spaced plants (Knowles, 1977; Altendorf, 2020, Ch 1) as well as in other crops like wheat and ryegrass (Philip et al., 2018; Elgersma and Śnieżko, 1988). In IWG, reported rates of FSU are low, and vary from 12 - 48%, depending on the environment (Larson et al., 2019; Altendorf, 2020, Ch 1). In a QTL mapping study of domestication traits in wheatgrass, seed size, fertility and yield-related traits aligned closely with a self-incompatibility (Z) gene (Larson et al., 2019), which suggests self-incompatibility may play a role. A recent review also suggested the *GNI1* genes in wheat and barley as potential targets for increasing fertility in IWG (DeHaan et al., 2020).

Height is considered an improvement trait and is a target of IWG breeding and selection. IWG plants are tall (~ 114 cm) and prone to lodging (Altendorf, 2020, Ch 1; Frahm et al., 2018), which can reduce yield and grain quality. Fixation of the *Reduced Height-1* genes or the 'green revolution genes' of wheat resulted in yield increases and their putative orthologs are suggested targets for selection or mutagenesis in IWG

(DeHaan et al., 2020). A better understanding of this and other domestication traits could greatly aid in our efforts to improve IWG.

We previously described a 1,168 individual F_1 Nested Association Mapping population (NAM) of IWG that was developed from 11 phenotypically divergent parents from Cycle 2 of the University of Minnesota breeding program (Yu et al., 2008; Altendorf, 2020, Ch 2). NAM has demonstrated great potential for uncovering marker-trait associations in other species but no published reports existed for F_1 progeny in a largely self -incompatible species. In the dissection of flowering time, using phenotypic data collected over two years at two locations, including St. Paul, MN and Salina, KS, we demonstrated that this population and the described methods were effective at identifying both previously detected QTL and a new one that aligns closely with well-characterized flowering time ortholog from barley, *Ppd-H1*. Here we map seven domestication traits that are major targets of selection in IWG breeding: brittle rachis, floret shattering, threshability, reproductive tiller number, floret site utilization, plant height, and thousand grain weight. The objectives of this work were threefold: 1) evaluate the phenotypic variation for these six domestication traits in the IWG NAM across four growing environments; 2) assess the relationship among these traits as well as others phenotyped in the population; 3) dissect their genetic control using genome wide association mapping (GWAS), and in linkage mapping using both within family and combined across families analyses using the two-way pseudo testcross approach in MapQTL6 (TWPT; Van Ooijen, 2009).

Materials and Methods

Population Development and Establishment

Parental selection, methods for crossing, propagating and establishing the IWG NAM were previously described in detail (Altendorf, 2020, Ch 2). Briefly, ten phenotypically diverse donor genets were identified from Cycle 2 of the UMN breeding program and crossed to WGN59, which is a low-shattering common genet, in a reciprocal manner. Approximately 130 seeds from each cross were germinated, allowed to establish and propagated into four clones and transplanted in St. Paul, MN (STP hereafter) and

Salina, KS at the Land Institute (TLI hereafter) in an RCBD design with two blocks each in fall 2016. A two-plant border was established to limit edge effects, and parents were propagated and replicated 3 (donor) or 25 (common) times per block. The population was phenotyped during the 2017 and 2018 field seasons. It is important to note that at TLI in 2018 there was a severe drought, where there was a maximum of 21 cm deficit of precipitation from the 10-year average at the time of anthesis (Altendorf, 2020, Ch 1). In combination with high temperatures, this drought resulted in plant death and poor performance of the spaced plants which played a major role in the interpretation of the results. The final NAM included 1,168 F1 progeny with an average of 117 individuals per family (Altendorf, 2020, Ch 2).

Phenotypic Data Collection

This study focuses on seven domestication traits, but to better understand their relationship with agronomic and phenological traits we ran a correlation analysis including some traits that were previously described in the NAM and here we include only a brief description of those methods (Altendorf, 2020, Chs 1 & 2). Two measures of reproductive growth stage were recorded: spike emergence percent and anthesis stage, and methods were previously described in detail (Altendorf, 2020, Ch 2). Briefly, spike emergence percent was estimated by measuring the length of the spike that emerged from the boot when spikes were approximately 50% emerged, divided by the final length of the spike at harvest. Anthesis was scored on the Feekes scale on one occasion per environment per year when anthers were showing (1600-1800 h) and when approximately 50% of the plants were in anthesis (Large, 1954). Categories were coded as ordinal variables from 1-10, ranging from boot stage (1) to kernels watery ripe (10). Plant height was recorded in cm at maturity when the plants reached maximum height.

Harvest methods, timing and growing degree days were previously described in detail (Altendorf, 2020, Chs 1 & 2). Briefly, ten of the larger, most uniform and fully intact spikes from each spaced plant were cut ~5 cm below the peduncle at harvest and placed in paper baguette bags. Uniform spikes were harvested not for selection but to reduce experimental error (DeHaan et al., 2018). Sample bags were carefully stored

upright, peduncle down in cardboard boxes to limit shattering or breakage in the bag and were dried in a 35 deg C dryer for 1 week. The remaining spikes from each spaced plant were gathered and clipped ~10 cm below the peduncle using large garden shears. The underside of the cut bundle was photographed in the field using a Canon digital SLR camera and stems were later counted on the computer by manually clicking with the "multi-point" function in ImageJ as previously described (Schneider et al., 2012). The total number of spikes in a ten-spike sample was verified and recorded to limit calculation errors. Three representative spikes were chosen from the ten-spike sample and used to determine spike length, spikelets spike^{-1} and florets spikelet^{-1}. Spikes were aligned end-to-end (from tip of most apical spikelet to base of most basal spikelet) along a measuring tape and total length was recorded and divided by three to obtain the mean. The number of spikelets was counted on three spikes and averaged. To estimate the number of florets spikelet^{-1}, the florets were counted in three spikelets on one representative spike, one from the bottom, middle and top one-third and the average was calculated according to guidelines outlined in Altendorf (2020, Ch 1). Shattering was measured by selecting three representative spikes from the ten spike sample and dropping them horizontally from a 20 cm height onto a 9 x 24" foam tray on an individual basis three times (method adapted from DeHaan et al., 2018; Supplementary Figure 3.1). The number and type of disarticulation events were counted and recorded. An event was classified as a single break, thus when a spike broke into two pieces it was considered a single event. Disarticulation events were classified as floret shattering within the rachilla axis using a scale of 0: 0 events; 1: 1-3 events; 2: 4-6; and increasing thereafter in increments of three, and brittle rachis disarticulation events were counted (Figure 3.1 A & B). The mean score for each type across the three spikes was calculated and rounded to the nearest integer. All shattered seeds and spike fragments were returned to the bag and were later threshed on a Wintersteiger LD350 (Ried, Austria) with the following settings: air flow: 3.5; RPM: 1000, 660, 380; concave: 0140-133 2x6 mm. Samples were sieved (12/64"; Seedburo, Des Plaines, IL) and cleaned using a fractionating aspirator (Carter Day International, Minneapolis, MN) set at 50 RPM (top setting) and 20 RPM (bottom setting). Threshability was visually estimated as the percent naked seeds of a clean

sample by comparing to known standards in petri dishes. Samples were rated in increments of 5%, except with resolution of 1% above 90% and below 10%. Since moisture content of a spike can influence shattering and threshability (personal observation), care was taken to ensure that samples were transferred from storage and placed back into a 95 deg F dryer for at least 12 hours prior to conducting shattering tests and threshing, and samples were removed from the dryer on an hourly basis to achieve consistent moisture content during daily processing. To determine seed count, area, and thousand grain weight, cleaned seed samples from the ten-spike-sample were imaged and weighed using a Marvin Seed Analyzer (https://www.marvitech.de/en/optical-seed-analyze/). Floret site utilization (FSU) in an economical sense (Elgersma, 1985) was estimated using the following equation:

$$FSU = \text{Total seed count} / (\text{Florets spikelet}^{-1} * \text{Spikelets spike}^{-1} * \text{Spikes in the ten spike sample})$$

Variant Detection and Filtering

We previously described the sequencing and variant detection methods in detail in Altendorf (2020, Ch 2). Briefly, genomic DNA was extracted from each genet using BioSprint 96 Plant DNA Kit (QIAGEN, the Netherlands) and 96-plex genotyping by sequencing (GBS; Elshire et al., 2011; Poland et al., 2012) libraries were developed using a two-enzyme (*Pst*I and *Msp*I), two-barcode approach described in Zhang et al. (2016) and sequenced on Illumina HiSeq 2500. The pipeline described in Neyhart et al., (2019) was used. Reads were demultiplexed with FastX Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/), adapters removed in CutAdapt (Martin, 2011), and quality assessed in FastQC (Babraham Bioinformatics; https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Reads were aligned to the v2 IWG reference genome (access provided by The Thinopyrum intermedium Genome Sequencing Consortium) using Bowtie2 (Langmead & Salzberg, 2012). Options were set to require reads to align entirely (--end-to-end), and the very-sensitive preset was used, with zero ambiguous reference characters allowed (--n-ceil C, 0, 0). Reads that mapped

uniquely were run through the Genome Analysis Toolkit HaplotypeCaller, CombineGVCF, and GenotypeGVCF tools (McKenna et al., 2010). Markers were filtered in VCFTools (Danecek et al., 2011) to include only bi-allelic SNPs with maximum 20% missing data, minimum allele depth of 5, and minor allele frequency > 0.005 (0.05 / 10 families). Individuals were removed if they had greater than 70% missing data or were determined to be selfs or unintended outcrosses as identified using Cervus v3.0 (Kalinowski et al., 2007). Marker imputation was conducted using LinkImpute (Money et al., 2015).

Phenotypic Data Analysis

All phenotypic data analyses were conducted in R v3.6.1 (R Core Team, 2019). As phenotypic data collection in the second year was essentially a repeated measure on the same plants, an initial model with family, location, and year as fixed effects, and genet, the interactions of location by rep, and location by rep by family by genet, as random effects was run to assess interactions of family by environment in the package 'lme4' (Gomez and Gomez 1984; Bates et al., 2015). Linear models were fit within each environment for each response variable with fixed terms for genet nested within family and block (n = 2). A square root transformation was applied to reproductive tiller counts and an arcsine square root to threshability to meet assumptions of ANOVA. For rachis breaks and floret score, where data was zero-inflated and of count nature, a generalized linear model was used. Estimated marginal means were calculated for genets nested within families on response and transformed scales as applicable using the 'emmeans' package (Length et al., 2018). Transformed scale emmeans were used for marker-trait association analyses in cases where data were transformed. Correlations between traits were assessed using the rcorr function in the package 'Hmisc'. The primary objective of the correlation analysis was to assess relationships between brittle rachis, floret shattering and threshability with yield component and maturity traits. Correlations between yield component and biomass traits as well as maturity are explored in detail in our previous work (Altendorf, 2020, Chs 1 & 2). T-tests were used to assess the phenotypic differences for each trait between progeny derived from common and donor parents

within each cross. To assess whether phenotypically divergent parents yielded F_1 families with high variability, the phenotypic difference between the parents was correlated with standard deviation among the progeny in each family using the rcorr function. Broad-sense heritability was calculated using a completely random model on a genet mean basis using the following formula:

$$H = \sigma_g^2 / \left(\sigma_g^2 + \sigma_e^2 / r\right)$$

Where σ_g^2 is the variance of genet (progeny only; parents were excluded), σ_e^2 is the error variance and *r* is the number of reps or blocks (n = 2). To calculate narrow sense heritability, a separate linear model was fit with family and rep as random effects. Narrow sense heritability is additive genetic variance divided by phenotypic variance and was calculated for half-sib families sharing a common parent using the following equation (Falconer & Mackay, 1996):

$$h^2 = (4 * \sigma_F^2) / [(4 * \sigma_F^2) + (\sigma_e^2)]$$

Where σ_F^2 is the variance of family and σ_e^2 is the error variance.

Genomewide Association Mapping and Linkage Mapping

Genomewide association mapping was conducted using GAPIT software (Lipka et al., 2012) with the default kinship matrix calculation. We previously described additional approaches to account for population structure using a Q-matrix and adding principal components, which produced essentially similar results (Altendorf, 2020, Ch 2) and thus were excluded from the model. Markers with a p-value below 0.00025 (LOD = 3.6) were considered significant. QTL within clusters or peaks were resolved using the mmer function in the R package 'Sommer' (Covarrubias-Pazaran, 2016) as described in Altendorf (2020, Ch 2). Allele effects and frequencies were obtained from the GAPIT output. To be considered significant and reported, a QTL was required to be detected in at least two of the four environments. The NAM consensus genetic map developed in Altendorf (2020, Ch 2) was used for linkage mapping in MapQTL v 6 (Van Ooijen,

2009) using the two-way pseudo testcross model (TWPT). In the TWPT method, QTL are mapped within each individual parent using separate maps, indicated as CP for the common parent, and DP for the donor parent. The maximum likelihood mixture analysis procedure was used with a maximum of 200 iterations and a maximum neighboring markers of 10. For each analysis, a permutation test with 1000 iterations was conducted to determine within-family and combined genomewide significance thresholds. Each LOD threshold for each family or combined analysis within environment was applied accordingly. Interval mapping was conducted first, and significant peak LOD markers for the combined analysis were evaluated and a maximum of one significant marker per LG was selected as a cofactor. The automatic cofactor selection procedure was used to select cofactors, which were then used in one round of restricted multiple QTL mapping (rMQM). A two-LOD drop off interval was used to establish QTL intervals (Van Ooijen, 2009). Only QTL for which an interval was detected for a trait and linkage group combination from at least two environments in the combined analysis are reported. LOD intervals and allele effects were summarized across all analyses using a custom R script.

Visualization of Results and Proximity to Orthogenes

Marker position, both physical (mbp) from the reference genome and genetic (cM) from the NAM consensus map across all chromosomes and linkage groups were normalized and visualized using 'LinkageMapView' (Ouellette et al., 2018). Markers in common between both the physical and the NAM genetic map were connected with a line using the posonleft function. GWAS and linkage mapping results were plotted. QTL intervals from the combined analysis that overlapped were collapsed into a single interval for the results figure and their precise cM values can be found in the results table. Because IWG linkage groups have shown high collinearity with barley and wheat (Zhang et al., 2016; Kantarski et al., 2016; DeHaan et al., 2020) known domestication orthogenes were selected (Supplementary Table 3.1) and aligned to the IWG draft reference genome using BLASTN 2.6.0+ (Altschul et al., 1997) with the online BLAST resource

(https://phytozome-next.jgi.doe.gov/blast-search)[4]. Significant ($1e^{-30}$) hits typically corresponded to a known homoeologous group (Kantarski et al., 2016). Positions of the orthologous gene hits were plotted.

Results

Population, Genotype Data and Genetic Maps

We previously described results from the development of the NAM population in detail (Altendorf, 2020, Ch 2). Briefly, a total of 1,168 F_1 progeny with an average of 117 progeny per family were used. 8,003 SNP markers were curated and imputed, and used in GWAS. In linkage mapping, the NAM map was a total of 3,385 cM (Haldane's) with one marker per 0.93 cM. Using the TWPT method, which excludes markers that are heterozygous in both parents, a total of 994 markers were used in linkage mapping with an average of 47 per linkage group (Altendorf, 2020, Ch 2).

Analysis of Variance

Family, location, year, and their interactions were all highly significant for all traits (year by location combinations; Supplementary Table 3.2). The year nested within location term explained the majority (avg. 70%) of the variance in every case except thousand grain weight, where year explained the most (68%), providing further support for the separate analysis of unique environments. The variance explained by family by year, or family by location was low (avg. 0.6%, 1.3%, respectively). In the analysis of unique environments, family and genet nested within family was highly significant ($P < 0.0001$; Supplementary Table 3.3). In every case, genets nested within family explained the majority of variation (average across all traits 65%), followed by family (15%) and replicate (avg 1.7%).

Brittle Rachis

Brittle rachis had the lowest heritability (average H = 0.37; $h^2 = 0.14$; Figure 3.2). Excluding TLI 2018, where brittle rachis occurred very rarely, the common parent,

[4] BLAST analysis conducted by Steve Larson, USDA-ARS.

WGN59, was the most susceptible, exhibiting on average 1.3 rachis disarticulation events per spike tested. For donor parents, WGN63 and WGN36 and WGN39 were the most susceptible, while WGN07 was the most resistant and the progeny means of these families followed this similar trend and displayed evidence for transgressive segregation (Figure 3.2; Supplementary Table 3.4). Phenotypic difference between the parents, however, was only a significant predictor of progeny variation in 1 of 4 environments (Supplementary Table 3.5). Brittle rachis did not share any notable (absolute value $r >$ 0.3) correlations with the traits surveyed (Supplementary Tables 3.6 – 3.9). In the GWAS analysis, four QTL were detected for brittle rachis on chromosomes 7, 8, 14, and 18. While they explained only a small percent (avg 3%) of the total variation they were consistently detected across at least two environments, but up to four in the case of chromosome 7. Linkage mapping also identified the QTL region on LG 8 segregating in both the common and donor parents (Figure 3.3; Table 3.7) which explained on average 18% of the variation.

Floret Shattering

Floret shattering was moderately heritable (average H = 0.55; $h^2 = 0.36$). WGN59 ranked among the most resistant, while WGN39 and WGN26 exhibited an average of 4.5 and 4.06 on the scale, which is equivalent to 13.5 and 12.2 floret disarticulation events per spike tested (Figure 3.4; Supplementary Table 3.4). The phenotypic difference between parental phenotypes was a significant predictor of progeny variation in all four environments ($P < 0.1$; Supplementary Table 3.5). It shared a moderate, significant correlation with seed area and thousand grain weight (avg r across environments, excluding TLI 2018 was 0.31 and 0.33, respectively; Supplementary Tables 3.6 – 3.9). Floret shattering had the highest number of QTL of all traits reported here, including 10 QTL across 9 chromosomes in GWAS and 4 significant regions in linkage mapping, suggesting highly polygenic control (Tables 3.3 – 3.7). Chromosomes 2 and 11 were in common across the two analyses, where the allele effects were primarily positive and negative, respectively. The QTL on chromosome 11, which explained 16% of the variation, segregated in the common parent and provided a consistent negative allele

effect in every family and environment analysis, corresponding to the distribution of progeny and parent phenotypic means for this trait (Figure 3.4). Linkage mapping identified two other significant regions on LGs 4 and 5, which were not detected in GWAS. The QTL on LG 4 explained on average 18% of the variation across families, segregated in the donor parents and had a positive allele effect (Table 3.7). QTL on LG 5 segregated in both the common and donor parents where the allele effect varied in its direction depending on the family (Table 3.7).

Reproductive Tiller Count

We previously reported high heritability for reproductive tiller count (average H = 0.84; $h^2 = 0.32$), and phenotype values that nearly doubled in the second year (Altendorf, 2020, Ch 1). Parent WGN07 had the fewest reproductive tillers (avg = 54) compared with WGN39 which had the most (avg = 142) (Figure 3.5; Supplementary Table 3.11). The effect of the maternal parent was variable across families, and three families, WGN07, WGN26 and WGN38, exhibited, on average across environments a significant effect ($P < 0.10$), where progeny derived from the common parent as the mother had more reproductive tillers (Supplementary Table 3.10). GWAS detected significant QTL on chromosomes 5, 6, 9, 15, 16, and 20, whereas linkage mapping identified only QTL on LGs 5 and 9. The QTL on chromosome 9 was the most significant (avg across environments -log10(p) = 8) and explained the most variation (avg = 4%) and was detected in linkage mapping analysis of both the common and donor parents (Table 3.7) with an average variance explained of 18%.

Threshability

Threshability, or percent dehulled seed, was highly heritable (average H = 0.79; $h^2 = 0.36$; Figure 3.6), and was generally high across the entire NAM population. Excluding TLI 2018 where variation in threshability was likely confounded by low seed number, WGN59 had the highest threshability (average across environments = 94%) and WGN63 had the lowest threshability (62%), and progeny values followed a similar trend. Variable parents were a significant predictor of progeny standard deviation in 3 of 4

environments ($P < 0.1$; Supplementary Table 3.5) and there was no effect of maternal parent (avg $P = 0.44$; Supplementary Table 3.10). Threshability and seed area were negatively correlated (avg $r = -0.28$). Threshability QTL were detected on chromosomes 5, 6, and 18 in GWAS and LGs 5 and 11 in linkage mapping. The GWAS marker on LG 5 was the most significant of all QTL detected in this study (-log10(p) = 12.7) and explained a high percentage of the variation (avg PVE = 12%) and was consistently detected in 3 of 4 environments. In linkage mapping this region was detected in the analysis of both the common parent and donor parent WGN15 and explained on average 27% of the variation.

Floret Site Utilization

Floret site utilization (FSU) had moderate broad and low narrow sense heritabilities (average H = 0.60; $h^2 = 0.18$; Figure 3.7), which was previously reported (Altendorf, 2020, Ch 1). Parents WGN55 and WGN59 had the highest FSU across environments, excluding TLI 2018, of 0.54 and 0.53, respectively (Figure 3.7; Supplementary Table 3.11). Parents WGN63 and WGN38 had the lowest FSU of 0.30 and 0.32 and progeny means followed a similar trend. Phenotypic differences between parents were not a significant predictor of progeny standard deviation (Supplementary Table 3.5) and there were no significant maternal effects (Supplementary Table 3.10). Few QTL for FSU were consistent across environments. In the case of GWAS specifically, results were highly variable across environments and many QTL were below the significance threshold. GWAS detected QTL across more than one environment on chromosomes 5 and 15 and linkage mapping on LGs 14 and 5 (Tables 3.3-3.7; Figure 3.3).

Height

We previously reported that height was highly heritable (average: H = 0.81; $h^2 = 0.59$) and variable across years and environments (Altendorf, 2020, Ch 1). WGN38 was the shortest parent (avg: 74 cm) and WGN63 the tallest (avg: 128 cm) (Figure 3.8; Supplementary Table 3.11). Progeny exhibited transgressive segregation, and

phenotypically divergent parents were not a significant predictor of progeny variation (Supplementary Table 3.5). In families WGN07, WGN38, and WGN46, there was a significant difference between the maternal parents, and progeny derived from the common parent as the mother were significantly taller on average across environments ($P < 0.1$; Supplementary Table 3.10). QTL for height were also highly variable across environments, with few meeting the required threshold for detection in more than one environment. GWAS identified QTL for height on chromosomes 4 and 5, and linkage mapping on LG 5 where allele effects were consistently negative and were detected in families WGN26 and WGN38.

Thousand Grain Weight

Heritability for thousand grain weight was high (average: $H = 0.78$; $h^2 = 0.56$) and the performance of parents and progeny was consistent across environments (Figure 3.9). Variation between parental phenotypes was not a significant predictor of progeny variation (Supplementary Table 3.5). Maternal effects on thousand grain weight were primarily nonsignificant with a few exceptions (Supplementary Table 3.10). GWAS detected significant QTL on chromosomes 9, 15 and 17, whereas linkage mapping only detected the QTL on chromosome 9 and only in STP environments.

Selfs

We previously described the identification of 74 progeny in the NAM that were unintended selfs or outcrosses and were therefore removed from the analysis. Since these plants were identified after the fact, they were still phenotyped. Selfs are typically known to perform poorly in the field (personal observation) and being able to identify them for culling purposes would be useful. In our previous work, we showed that selfs on average tend to mature later than F_1s (Altendorf, 2020, Ch 2). Here we show that their rate of brittle rachis shattering, threshability percent, and thousand grain weight (with the exception of STP 2017) were indistinguishable from F_1s (Supplementary Figure 3.2). However, they had fewer reproductive tillers, lower floret site utilization, shorter height, and lower floret shattering in most cases. Despite the significant differences, in many

cases they exhibited values very similar to F_1s and phenotype data alone may not be sufficient to discriminate selfs using the traits reported here.

Discussion

In several crop species domestication syndrome traits are under the control of single or few large effect QTL (e.g. Simons et al., 2006; Taketa et al., 2008; Asano et al., 2011). Considering the highly conserved gene order across species in the Triticeae tribe, and the decades of research into understanding the genetics of domestication, it is possible to utilize this knowledge to expedite the domestication of new crops, in particular in the case of intermediate wheatgrass (DeHaan et al., 2020). Utilizing a large Nested Association Mapping population (NAM) of IWG to investigate the genetic control of domestication traits, we demonstrated that QTL for brittle rachis, floret shattering, and threshability are either physically close or on the same chromosome or homeologous groups as significant BLAST hits to major domestication orthologous genes in barley, wheat and rice. For example, in the case of QTL for brittle rachis, Chromosomes 7 & 8 are in IWG homeologous group 3 and BLAST hits with the barley *Btr1* and *Btr2* sequences from 3H were 1.3 mbp and 216 bp away from the most significant GWAS marker on Chromosome 7, and 14 mbp away from the peak locus in linkage mapping on LG 8 (Figure 3.3; Tables 3.3 – 3.7; Candidate orthogene positions in Supplementary Table 3.1). The GWAS hit for brittle rachis on chromosome 14, which was detected in two environments, was 31 mbp away from a BLAST hit for the domestication gene, *Q*, of wheat which is known to affect rachis fragility (Simons et al., 2006). This region, however, was not detected using linkage mapping. For floret shattering, the peak locus for the QTL segregating in the common parent on Chromosome 2, which explained an average of 11.6% of the variation, aligned with a candidate ortholog BLAST hit for *SH5*, which enhances seed shattering in rice, within 8.6 mbp (Yoon et al., 2014; Table 3.7; Supplementary Table 3.1). The floret shattering QTL on LG 11 aligns closely with a previously reported QTL in IWG for shattering (Larson et al., 2018). Interestingly, a significant BLAST hit for the candidate ortholog, *sh4* of rice, was

on the opposite end of the chromosome (Li et al., 2006), which suggests that it is unlikely causative, and provides evidence that this locus may be novel, and unique to IWG. On chromosome 4, A BLAST hit for the potential candidate ortholog of rice, *SHAT1*, is within 26 mbp of the peak LOD marker for floret shattering, detected in two environments (Zhou et al., 2012). Threshability QTL were detected on chromosome 5 in both GWAS and LM. Chromosome 5 is part of the homeologous group 2 of the Triticeae, where wheat threshability genes, *sog* and *Tg*, are located (Sood et al., 2009) though these have not been cloned (Haas et al., 2019) and could therefore not be easily aligned to the IWG genome. BLAST hits for the reduced height (*rht*) gene in wheat mapped to group homoeologous 4 chromosomes, where no QTL for height were detected. In most cases, domestication traits in IWG are under more complex genetic control compared to its diploid relatives, with more QTL detected across multiple chromosomes, in some cases on more than one chromosome in a homoeologous group, as was specifically demonstrated in the case of brittle rachis.

While previous researchers have defined the mechanisms of seed shattering in IWG (Larson et al., 2018; DeHaan et al., 2020), this study is the first to distinguish between brittle rachis and floret shattering in phenotyping and data analysis. To do this, we adapted a spike drop test method described by DeHaan et al. (2018) and expanded the scale to characterize and record the disarticulation events separately. The correlation analysis demonstrated that there was no correlation between the two types of shattering (Supplementary Tables 3.6 – 3.9), suggesting independent genetic control. Furthermore, neither trait showed a strong correlation with maturity or traits that might exhibit a mechanical influence on shattering (e.g. spike length, seed size, yield, and headweight). This finding, in combination with high heritability estimates, suggests that the variation was due primarily to genetics and less-so to environmental variation. QTL mapping results showed that the two are indeed under distinct genetic control in IWG. Brittle rachis is under the control of fewer QTL, primarily on chromosomes 7 and 8. Floret shattering appears to be more complex with QTL detected across 9 chromosomes in GWAS and 4 in linkage mapping. The QTL detected for floret shattering corresponds to others detected in Larson et al., (2018) on chromosomes 6, 10, 11, 12, at 14. The Larson

et al. (2018) study, however, did not detect seed shattering QTL on homoeologous group 3, where the *Btr* orthologs are located. It is possible this trait was not segregating in their M26 x M35 IWG bi-parental population, or the method of phenotyping failed to capture the variation for this trait. Interestingly, using the original phenotyping method of scoring both traits together, The Land Institute breeding program was able to greatly improve the non-brittle rachis trait (Lee DeHaan, personal communication). For example, spikes of some wild types completely disarticulated by harvest maturity, whereas in current material, it no longer occurs to that degree. However, this trait is still segregating in the UMN program, or at least was in the C2 material from which this population was derived. The relatively simple genetic control would suggest that it may be easy to purge with the correct method of phenotyping and selection. Floret shattering, with its more complex genetic control, will be more difficult to purge and with more precise phenotyping can be a target for the routine genomic selection programs in IWG.

The NAM parents were chosen based on their phenotypic values for domestication traits including flowering time, shattering, threshability, height and seed size (Altendorf, 2020, Ch 2). The performance of the parental phenotypes corresponded with their reasons for inclusion in the population. For example, WGN07 and WGN26 were selected for their low and high shattering, respectively, and WGN63 for its low threshability. WGN59, the common parent, was selected for its low shattering. At the time, both types of shattering were evaluated collectively and therefore we missed the fact that WGN59 was highly susceptible to brittle rachis, providing further support for the importance of evaluating shattering types separately.

Previous work in IWG has suggested that the barley domestication gene for seed nakedness, *nud*, is a putative candidate ortholog for the free-threshing trait (Larson et al., 2018; DeHaan et al., 2020). Interestingly, no threshability QTL were detected on homeologous group (HG) 7 chromosomes, where the BLAST hits were (Figure 3.3). We did, however, detect large effect QTL on group 2 chromosomes, specifically chromosome 5, notably where the wheat *sog* and *Tg* loci are located. The mechanism for free threshing varies depending on the species, and in IWG threshability is typically described as similar to barley, where seeds are hulled, and not like wheat where the caryopses are difficult to

remove from the glumes. In this study, threshability was rated as percent clean seed after threshing in a Wintersteiger LD350, which is drum-style thresher with a concave. Overall, threshability was very high across the population (avg = 88% clean seed). Whether the high rating was due to the population itself being highly free threshing, or due to the method used, is not clear. The strong signal from HG 2 and the absence of QTL in HG 7 would suggest that threshability in the case of this population and these environments is more likely controlled in a fashion similar to that in wheat. It is possible that easier release from the glumes in IWG, if allowed sufficient threshing time, may result in a higher percentage of clean, dehulled seed. Furthermore, only very rarely did we observe seeds with hulls physically adhered to the caryopsis as is seen in hulled barley. Hulls in this population could be removed with relative ease. Since the UMN breeding program was derived from a small subset of improved material from the Land Institute program (Zhang et al., 2016) and this population represents an even smaller portion of that variation, it is possible that the adherent hull characteristic was not segregating in the NAM. Further sequence similarity analysis across wheat and IWG using the SSR markers linked with *Tg* and *sog* in wheat would need to be conducted to provide further evidence for this hypothesis. As evidenced by the tendency for IWG to exhibit floret shattering, its rachilla axis is fragile and long, sometimes containing up to 12 florets (Altendorf, 2020, Ch 1). Furthermore, there was a negative correlation between threshability and floret shattering (average *r* across environments = -0.21), suggesting that individuals with greater shattering resistance had greater percent naked seed, a finding also reported by Larson et al., (2018). Perhaps a more rigid rachilla axis with lower floret shattering keeps the lemma and palea in place, and allows for easier removal of the caryopsis in the thresher. In rice, fine tuning of the expression of the shattering genes, *Sh3* and *Sh4* was important for maintaining low shattering but also high threshability of the grain (Li et al., 2006; Lv et al., 2018), and a similar approach may be required for shattering or threshability genes in IWG.

Frequently the GWAS QTL detected in multiple environments are not supported by linkage mapping results. In general, the GWAS analysis paints a much more complex picture of the genetic control of these traits than does the linkage mapping analysis. This

could be due to a few reasons. First, linkage mapping may have been limited in its effectiveness due to the severe segregation distortion we observed in the linkage map creation process (Altendorf, 2020, Ch 2). Due to the severe segregation distortion, and singularity errors associated with long spans of uninformative markers, we were required to utilize the TWPT method, which excludes any markers that are heterozygous in both parents, and uses only markers that are heterozygous in one and homozygous in the other, further reducing the marker count. It is possible that due to low marker numbers some QTL were missed. For example, it is puzzling that GWAS for brittle rachis detected QTL on chromosome 7 in all four environments, even in TLI 2018 where brittle rachis barely occurred at all under the severe drought conditions, but this region was not detected in the linkage mapping analysis. We previously reported marker types and counts for the linkage map creation step, and LG 7 had especially low marker counts (Altendorf, 2020, Ch 2). Second, the v2 IWG genome is considered a draft and there are various errors, several of which are visible in the ordering of markers on the genetic and physical maps (Figure 3.3). It is also possible that there are errors in the placement of markers on different chromosomes within homoeologous groups, and perhaps even across homoeologous groups resulting in what may appear to be spurious significant associations in GWAS.

The magnitude of family by environment interaction was low in the analysis of variance across environments, suggesting that rank in the performance of families was relatively consistent and that GxE was minimal. The term explaining the greatest proportion of the variance was year nested within location, which suggests that the overall performance of the population varied substantially across years within locations (Supplementary Table 3.2). The consistency of QTL detection across environments varied depending on the trait, where traits like threshability, brittle rachis, floret shattering, and fertile tiller number were more consistent, but height, floret site utilization, and thousand grain weight were less consistent. In cases where they were less consistent, it is possible that the QTL were of smaller effect and that these traits are more susceptible to environmental variation. The requirement that a QTL must occur in more than one environment to be considered significant may have inherently excluded QTL

unique to the TLI environment since the drought in 2018 limited the ability to detect QTL. This may specifically be the case for thousand grain weight, a trait likely greatly influenced by environmental factors, where only one QTL on Chromosome 9 was detected in LM and only in STP.

The large number of linkage mapping and GWAS QTL for several domestication traits on chromosome 5, detected across multiple environments, is puzzling and somewhat unexpected considering the lack of candidate orthologs in that region. Interestingly, these QTL have negative effects for FSU and height but positive effects for reproductive tiller number, and variable effects for floret shattering. We previously described that the IWG NAM is more inbred and the parents more closely related than we expected after selecting them based solely on their phenotypic divergence (Altendorf, 2020, Ch 2). Possibly, this region, when highly homozygous or identical by descent, confers a reduction in some fitness traits. For example, we showed that selfed plants tend to have low FSU and are shorter in height. To further investigate, we looked at the GWAS SNP, S05_426524999, which was significant for three traits in GWAS across two or three environments. In some cases this SNP is supported by a QTL peak or cluster of neighboring SNPs, but not all. Furthermore, this SNP segregated in 8 of 10 families and was only significantly distorted in 2 of 8, but it did not easily group with LG 5 in the creation of the linkage map and was therefore not used in linkage mapping. This may be an example of a misplaced marker in the genome. However, this alone does not explain the density of QTL in this region, as the significant LOD intervals span a larger region in linkage mapping. Alternatively, this region may harbor many domestication QTL. In a GWAS and linkage mapping study on seed size in IWG, LG 5 harbored QTL for seed weight, seed area, seed length and seed size (Zhang et al., 2017). In a GWAS study of yield component traits in IWG, a QTL on chromosome 5 was significant for seed length and thousand grain weight, however these peak markers were physically distant, nearly 60 mbp apart (Bajgain et al., 2019a). The Larson et al. (2018) study identified QTL on LG 5 for threshability, or what they called seed nakedness, and plant height in which the intervals cover approximately the same region as the QTL found here. Therefore, there is

evidence the chromosome 5 may play an important role in the selection for domestication and seed size traits in IWG.

In the development of bi-parental populations for QTL mapping, it is common to select individuals that vary phenotypically for a trait of interest. In the case of the NAM, we showed that there were only select cases where divergent parents resulted in divergent progeny and that divergence can vary depending on the environment. For example, divergent parents for floret shattering and threshability resulted in divergent progeny, whereas brittle rachis varied by environment, and reproductive tillers, floret site utilization, height, and thousand grain weight were consistently insignificant. In a highly heterozygous species like IWG it is of equal importance to choose highly heterozygous parents in order to achieve segregating progeny. The IWG NAM was more inbred than we anticipated, which may have played a role in the segregation and variation observed (Altendorf, 2020, Ch 2). In several cases, significant and large effect QTL were detected in families where the parents exhibited very similar phenotypic values. One example is threshability in family WGN15, where the parents performed similarly, but progeny were highly variable (Figure 3.6), and QTL were consistently detected on LG 5 and explained a large proportion of the variance in that family.

Previous work in IWG has demonstrated the potential for using significant QTL markers as fixed effects in genomic selection to increase the frequency of beneficial alleles (Zhang et al., 2017; Bajgain et al., 2019a). In the case of complex traits like floret shattering, height, and thousand grain weight, this approach may be advantageous. In cases of traits like brittle rachis and threshability, where the genetic control appears to be less complex, working towards the validation of these QTL and developing resources for marker assisted selection may allow breeders to select parental material with the advantageous alleles for crossing blocks. An alternative and potentially complementary approach may be exploring the use of targeted mutagenesis in known domestication gene sequences to alter gene function (DeHaan et al., 2020), and efforts to do so may be better informed by the results of this study which provides substantial evidence in support of potential candidate orthogenes.

Tables and Figures

Table 3.1. Results from genomewide association mapping (GWAS) organized by trait. Only significant SNPs that were detected in more than one environment are reported.

Trait	Unit	SNP[a]	Alleles[b]	Chr[c]	Pos[d]	MAF[e]
Brittle Rachis	Count	S07_66140981	G/A	7	66140981	0.08
		S08_77402910	G/A	8	77402910	0.46
		S14_386756967	G/A	14	386756967	0.07
		S18_214242119	G/A	18	214242119	0.05
Floret Shattering	Scale	S02_181520123	A/T	2	181520123	0.03
		S02_296782379	G/C	2	296782379	0.47
		S03_133897959	G/A	3	133897959	0.03
		S06_524746973	T/C	6	524746973	0.37
		S09_106857528	G/A	9	106857528	0.03
		S10_436139244	A/G	10	436139244	0.46
		S11_244241276	A/C	11	244241276	0.47
		S15_454588273	C/A	15	454588273	0.05
		S18_40836512	G/A	18	40836512	0.46
		S20_593264859	C/T	20	593264859	0.09
Reproductive Tillers	Count	S05_426524999	C/G	5	426524999	0.25
		S06_52435107	T/C	6	52435107	0.03
		S09_450570468	G/C	9	450570468	0.35
		S15_344515633	C/T	15	344515633	0.02
		S16_326253169	G/C	16	326253169	0.47
		S20_174499953	C/A	20	174499953	0.46
Threshability	%	S05_156358046	G/T	5	156358046	0.35
		S06_147094010	C/G	6	147094010	0.27
		S06_424161336	T/A	6	424161336	0.05
		S18_165923396	C/A	18	165923396	0.08
		S18_17030514	G/A	18	17030514	0.1
Floret Site Utilization	%	S05_426524999	C/G	5	426524999	0.25
		S15_212606143	G/A	15	212606143	0.02
Height	cm	S04_173993463	C/T	4	173993463	0.03
		S05_426524999	C/G	5	426524999	0.25
Thousand Grain Weight	g	S02_262333799	A/G	2	262333799	0.03
		S07_462262924	C/A	7	462262924	0.03
		S09_447878309	A/G	9	447878309	0.37

[a] SNP, single nucleotide polymorphism
[b] Alleles at the SNP including reference and alternate, respectively
[c] Chr, chromosome
[d] Pos, position of SNP in base pairs
[e] MAF, minor allele frequency

Table 3.2. Intermediate wheatgrass nested association mapping (NAM) families that segregate with a minor allele frequency above 0.005 for the QTL detected as represented by their numerical identification.

Trait	SNP[a]	Segregating Families
Brittle Rachis	S07_66140981	26, 39, 45, 46, 55
	S08_77402910	7, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S14_386756967	39, 45, 63
	S18_214242119	26, 63
Floret Shattering	S02_181520123	39
	S02_296782379	7, 15, 26, 36, 38, 39, 46, 55, 63
	S03_133897959	39
	S06_524746973	7, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S09_106857528	39
	S10_436139244	7, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S11_244241276	7, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S15_454588273	39
	S18_40836512	7, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S20_593264859	15, 36, 38, 39, 55, 63
Reproductive Tillers	S05_426524999	7, 15, 26, 36, 38, 39, 45, 46
	S06_52435107	7
	S09_450570468	7, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S15_344515633	45
	S16_326253169	7, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S20_174499953	7, 15, 26, 36, 38, 39, 45, 46, 55, 63
Threshability	S05_156358046	7, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S06_147094010	7, 15, 26, 36, 38, 39, 45, 46, 55, 63
	S06_424161336	15, 39, 45
	S18_165923396	15, 45, 63
	S18_17030514	39, 45, 63
Floret Site Utilization	S05_426524999	7, 15, 26, 36, 38, 39, 45, 46
	S15_212606143	39
Height	S04_173993463	46
	S05_426524999	7, 15, 26, 36, 38, 39, 45, 46
Thousand Grain Weight	S02_262333799	39
	S07_462262924	36, 39, 45
	S09_447878309	7, 15, 26, 36, 38, 39, 45, 46, 55, 63

[a] SNP, single nucleotide polymorphism

Table 3.3. Results from genomewide association mapping results for St. Paul, 2017 organized by trait. Only significant SNPs that were detected in more than one environment are reported. Instances in which the QTL was not detected in the present environment are indicated by missing values ("-").

Trait	SNP[a]	-log10(p)[b]	Effect[c]	PVE[d]
Brittle Rachis	S07_66140981	9	0.302	4.29
	S08_77402910	4.3	0.115	3.72
	S14_386756967	-	-	-
	S18_214242119	5.1	-0.29	2.9
Floret Shattering	S02_181520123	5.2	1.28	46.65
	S02_296782379	6.8	0.515	1.07
	S03_133897959	4.9	1.197	0.24
	S06_524746973	7.9	0.541	1.92
	S09_106857528	4.3	1.119	3.45
	S10_436139244	4.8	0.324	0.97
	S11_244241276	6.1	-0.353	5.42
	S15_454588273	7.1	-2.25	17.24
	S18_40836512	5.2	0.582	0.54
	S20_593264859	3.7	0.383	1.42
Reproductive Tillers	S05_426524999	5.5	0.426	2.75
	S06_52435107	-	-	-
	S09_450570468	10.3	0.681	5.06
	S15_344515633	-	-	-
	S16_326253169	6.3	0.29	1.4
	S20_174499953	5.2	0.191	1.15
Threshability	S05_156358046	12.7	0.008	10.36
	S06_147094010	5.5	0.006	1.91
	S06_424161336	6.3	0.011	1.56
	S18_165923396	4.8	0.007	2.67
	S18_17030514	4	0.003	2.81
Floret Site Utilization	S05_426524999	5.6	0.03	2.93
	S15_212606143	4	0.066	0.26
Height	S04_173993463	9.3	14.704	4.84
	S05_426524999	4.3	3.471	1.75
Thousand Grain Weight	S02_262333799	3.8	-0.861	0.98
	S07_462262924	4.1	-0.707	0.92
	S09_447878309	3.6	-0.317	3.93

[a] SNP, single nucleotide polymorphism
[b] Level of significance
[c] Allele effect in trait units
[d] PVE, percent variance explained by the QTL

Table 3.4. Results from genomewide association mapping results for St. Paul, 2018 organized by trait. Only significant SNPs that were detected in more than one environment are reported. Instances in which the QTL was not detected in the present environment are indicated by missing values ("-").

Trait	SNP[a]	-log10(p)[b]	Effect[c]	PVE[d]
Brittle Rachis	S07_66140981	6.5	0.357	2.93
	S08_77402910	4.5	0.176	2.89
	S14_386756967	4.4	0.299	1.65
	S18_214242119	-	-	-
Floret Shattering	S02_181520123	4	0.905	5.49
	S02_296782379	-	-	-
	S03_133897959	3.9	0.869	0.35
	S06_524746973	-	-	-
	S09_106857528	4.6	0.958	8.84
	S10_436139244	-	-	-
	S11_244241276	5.5	-0.273	2.12
	S15_454588273	5.7	-1.643	5.63
	S18_40836512	-	-	-
	S20_593264859	4.5	0.364	1.58
Reproductive Tillers	S05_426524999	4.4	0.418	2.38
	S06_52435107	3.7	2.289	2.43
	S09_450570468	8.8	0.715	4.49
	S15_344515633	4	2.293	1.71
	S16_326253169	5	0.29	1.11
	S20_174499953	3.9	0.177	1
Threshability	S05_156358046	11.3	0.011	11.28
	S06_147094010	-	-	-
	S06_424161336	4.9	0.012	1.43
	S18_165923396	-	-	-
	S18_17030514	6	0.008	3.41
Floret Site Utilization	S05_426524999	-	-	-
	S15_212606143	-	-	-
Height	S04_173993463	-	-	-
	S05_426524999	2.3	2.211	0.84
Thousand Grain Weight	S02_262333799	4.1	-0.771	0.62
	S07_462262924	-	-	-
	S09_447878309	3.9	-0.285	4.2

[a] SNP, single nucleotide polymorphism
[b] Level of significance
[c] Allele effect in trait units
[d] PVE, percent variance explained by the QTL

Table 3.5. Results from genomewide association mapping results for The Land Institute in Salina, KS in 2017 organized by trait. Only significant SNPs that were detected in more than one environment are reported. Instances in which the QTL was not detected in the present environment are indicated by missing values ("-").

Trait	SNP[a]	-log10(p)[b]	Effect[c]	PVE[d]
Brittle Rachis	S07_66140981	10.3	0.403	3.2
	S08_77402910	4.2	0.146	2.49
	S14_386756967	-	-	-
	S18_214242119	6.9	-0.435	4.45
Floret Shattering	S02_181520123	-	-	-
	S02_296782379	3.7	0.272	0.42
	S03_13387959	-	-	-
	S06_524746973	5.3	0.319	0.65
	S09_106857528	-	-	-
	S10_436139244	4.7	0.234	0.8
	S11_244241276	6.7	-0.27	2.69
	S15_454588273	-	-	-
	S18_40836512	3.6	0.353	0.21
	S20_593264859	-	-	-
Reproductive Tillers	S05_426524999	-	-	-
	S06_52435107	4.9	1.998	2.02
	S09_450570468	5.4	0.266	2.66
	S15_344515633	4.2	1.416	1.89
	S16_326253169	-	-	-
	S20_174499953	4.3	0.124	1.19
Threshability	S05_156358046	13.3	0.01	13.79
	S06_147094010	5.1	0.006	1.39
	S06_424161336	-	-	-
	S18_165923396	4.2	0.007	2.46
	S18_17030514	-	-	-
Floret Site Utilization	S05_426524999	3.6	0.024	1.2
	S15_212606143	4.2	0.075	2.96
Height	S04_173993463	5.3	13.092	1.19
	S05_426524999	-	-	-
Thousand Grain Weight	S02_262333799	-	-	-
	S07_462262924	4.4	-0.689	1.97
	S09_447878309	-	-	-

[a] SNP, single nucleotide polymorphism
[b] Level of significance
[c] Allele effect in trait units
[d] PVE, percent variance explained by the QTL

Table 3.6. Results from genomewide association mapping results for The Land Institute in Salina, KS in 2018 organized by trait. Only significant SNPs that were detected in more than one environment are reported. Instances in which the QTL was not detected in the present environment are indicated by missing values ("-").

Trait	SNP[a]	-log10(p)[b]	Effect[c]	PVE[d]
Brittle Rachis	S07_66140981	6	0.099	2.79
	S08_77402910	-	-	-
	S14_386756967	4.5	0.09	1.67
	S18_214242119	-	-	-
Floret Shattering	S02_181520123	-	-	-
	S02_296782379	-	-	-
	S03_133897959	-	-	-
	S06_524746973	-	-	-
	S09_106857528	-	-	-
	S10_436139244	-	-	-
	S11_244241276	-	-	-
	S15_454588273	-	-	-
	S18_40836512	5.6	0.081	1.03
	S20_593264859	-	-	-
Reproductive Tillers	S05_426524999	-	-	-
	S06_52435107	-	-	-
	S09_450570468	-	-	-
	S15_344515633	-	-	-
	S16_326253169	-	-	-
	S20_174499953	-	-	-
Threshability	S05_156358046	-	-	-
	S06_147094010	-	-	-
	S06_424161336	-	-	-
	S18_165923396	-	-	-
	S18_17030514	-	-	-
Floret Site Utilization	S05_426524999	4.4	0.009	1.82
	S15_212606143	-	-	-
Height	S04_173993463	-	-	-
	S05_426524999	-	-	-
Thousand Grain Weight	S02_262333799	-	-	-
	S07_462262924	-	-	-
	S09_447878309	-	-	-

[a] SNP, single nucleotide polymorphism
[b] Level of significance
[c] Allele effect in trait units
[d] PVE, percent variance explained by the QTL

Table 3.7. Results from linkage mapping analyses combined across families and within families for reproductive tillers. Only QTL intervals that were detected for the same trait on the same linkage group and in overlapping QTL intervals in more than one environment are reported.

Trait	Units	LG[a]	Map[b]	Left[c]	Peak[d]	Right[e]	Peak Locus[f]	Max LOD[g]	Analysis[h]	r^2	μ[i]	α[j]	γ[k]	Loc	Year
Brittle Rachis	Count	8	CP	61.0	68.7	74.8	CP_Chr08_23347234	9.73	combined	-	-	-	-	STP	2018
		8	CP	64.7	75.6	87.4	CP_Chr08_22429423	11.85	combined	-	-	-	-	TLI	2017
		8	CP	64.7	77.0	106.3	-	3.60	WGN36	14.1	0.58	0.23	-	TLI	2017
		8	DP	55.5	82.0	89.6	-	8.14	combined	-	-	-	-	STP	2017
		8	DP	64.1	82.0	89.6	-	11.08	combined	-	-	-	-	STP	2018
		8	DP	58.5	82.0	89.6	-	6.26	WGN45	23.3	0.62	-	0.33	STP	2018
Floret Shattering	Scale	2	CP	56.3	76.8	89.1	CP_Chr02_343055305	7.98	combined	-	-	-	-	STP	2017
		2	CP	86.1	103.6	130.6	-	3.77	WGN39	12.5	2.54	0.52	-	STP	2018
		2	CP	55.3	77.5	86.1	CP_Chr02_338749407	12.37	combined	-	-	-	-	TLI	2017
		2	CP	41.3	76.8	102.6	CP_Chr02_343055305	3.57	WGN38	14.6	1.14	0.38	-	TLI	2017
		2	DP	76.9	85.0	97.5	-	9.16	combined	-	-	-	-	STP	2017
		2	DP	51.0	52.8	107.7	DP_Chr02_9740407	3.23	WGN26	8.7	1.26	-	0.33	STP	2017
		2	DP	75.9	91.4	112.4	-	8.65	combined	-	-	-	-	STP	2018
		2	DP	70.6	79.8	79.8	DP_Chr02_338749432	3.57	WGN38	10.6	0.71	-	-0.24	STP	2018
		2	DP	66.3	77.9	109.7	-	9.21	combined	-	-	-	-	TLI	2017
		4	DP	195.5	196.5	208.5	-	14.68	combined	-	-	-	-	STP	2017
		4	DP	195.5	201.5	201.5	-	4.77	WGN15	14.5	0.67	-	0.36	STP	2017
		4	DP	188.2	196.5	196.5	-	5.47	WGN55	11.9	1.52	-	0.49	STP	2017
		4	DP	182.2	195.5	195.5	DP_Chr04_260663237	12.21	combined	-	-	-	-	STP	2018
		4	DP	177.2	187.2	187.2	-	5.42	WGN55	28.2	1.25	-	0.42	STP	2018
		5	CP	126.2	135.0	138.3	-	18.40	combined	-	-	-	-	STP	2017
		5	CP	95.6	128.9	146.4	-	3.33	WGN26	15.4	1.21	-0.41	-	STP	2017
		5	CP	84.6	102.3	113.3	-	6.47	WGN55	16.5	1.46	-0.53	-	STP	2017
		5	CP	128.9	136.0	146.4	-	16.03	combined	-	-	-	-	STP	2018
		5	CP	6.1	63.0	83.6	CP_Chr05_86148562	3.18	WGN26	9.2	1.12	0.33	-	STP	2018
		5	CP	116.9	136.0	146.4	-	4.47	WGN38	16.2	0.67	0.29	-	STP	2018
		5	CP	113.3	118.2	146.4	-	5.26	WGN55	13.8	1.24	-0.33	-	STP	2018

		5	CP	129.9	136.0	146.4	-	16.89	combined	-	-		-	-		TLI	2017
		5	CP	126.9	135.0	146.4	-	4.30	WGN45	15.3		1.15	-0.31	-		TLI	2017
		5	DP	115.8	127.8	132.8	-	8.09	combined	-	-		-	-		STP	2017
		5	DP	101.8	128.8	132.8	DP_Chr05_392031395	3.73	WGN15	29.6		0.52	-		-1.13	STP	2017
		5	DP	101.8	109.8	131.3	-	4.50	WGN63	12		1.95	-		-0.57	STP	2017
		5	DP	143.4	146.4	146.4	DP_Chr05_339720810	9.36	combined	-	-		-	-		STP	2018
		5	DP	143.4	146.4	146.4	DP_Chr05_339720810	7.37	WGN07	41		0.72	-		1.78	STP	2018
		11	CP	106.8	109.2	111.3	-	31.98	combined	-	-		-	-		STP	2017
		11	CP	101.4	109.4	121.3	CP_Chr11_238097236	4.64	WGN26	15.1		1.21	-0.41	-		STP	2017
		11	CP	101.4	106.8	112.3	CP_Chr11_236555758	4.75	WGN39	14.8		3.05	-0.66	-		STP	2017
		11	CP	93.6	97.6	111.3	-	5.74	WGN45	23.7		1.12	-0.49	-		STP	2017
		11	CP	102.4	115.3	122.3	CP_Chr11_231841692	3.62	WGN46	12.9		1.17	-0.46	-		STP	2017
		11	CP	87.4	103.4	115.3	-	5.70	WGN55	15.2		1.49	-0.49	-		STP	2017
		11	CP	107.8	109.2	111.3	-	35.04	combined	-	-		-	-		STP	2018
		11	CP	101.4	108.2	112.3	CP_Chr11_244241276	4.53	WGN07	15.4		0.87	-0.39	-		STP	2018
		11	CP	80.6	101.4	115.3	CP_Chr11_228740879	5.38	WGN15	14.9		0.66	-0.31	-		STP	2018
		11	CP	96.6	109.2	109.2	-	3.37	WGN26	11.1		1.15	-0.36	-		STP	2018
		11	CP	106.8	109.4	109.4	CP_Chr11_238097236	5.53	WGN38	19.8		0.68	0.32	-		STP	2018
		11	CP	90.8	104.4	122.3	-	4.32	WGN39	15.4		2.44	-0.55	-		STP	2018
		11	CP	106.8	122.3	122.3	-	6.43	WGN45	27.2		1.09	-0.48	-		STP	2018
		11	CP	80.6	91.8	106.8	CP_Chr11_225969790	5.57	WGN55	12.4		1.21	-0.31	-		STP	2018
		11	CP	104.4	109.2	117.3	-	21.73	combined	-	-		-	-		TLI	2017
		11	CP	85.6	109.2	109.2	-	4.21	WGN36	15.4		1.11	-0.36	-		TLI	2017
		11	CP	94.6	113.3	113.3	CP_Chr11_234190857	4.47	WGN45	15.8		1.13	-0.32	-		TLI	2017
Reproductive Tillers	Count	5	DP	135.6	144.4	146.4	-	11.94	combined	-	-		-	-		STP	2017
		5	DP	54.0	136.6	146.4	-	3.60	WGN38	12		75	-		0.38	STP	2017
		5	DP	48.0	112.8	146.4	-	3.35	WGN26	43.3		114.7	-		2.4	STP	2018
		5	DP	103.8	140.3	146.4	-	10.02	combined	-	-		-	-		TLI	2017
		5	DP	59.9	111.8	146.4	-	4.75	WGN38	51.9		81	-		1.77	TLI	2018
		9	CP	127.2	130.9	136.8	CP_Chr09_447878309	19.07	combined	-	-		-	-		STP	2017
		9	CP	127.2	130.4	137.8	CP_Chr09_450004747	7.11	WGN15	24.3		71.4	0.42	-		STP	2017
		9	CP	113.8	125.2	140.2	-	4.59	WGN36	15.4		62.73	0.36	-		STP	2017
		9	CP	123.2	130.9	137.8	-	17.20	combined	-	-		-	-		STP	2018
		9	CP	123.2	133.4	140.2	CP_Chr09_111605290	6.41	WGN15	21.6		174.24	0.61	-		STP	2018

		9	CP	115.8	129.7	140.2	CP_Chr09_450543915	4.69	WGN36	14.9	140.66	0.48	-		STP	2018
		9	CP	94.3	123.2	140.2	-	3.34	WGN63	17.6	195.44	0.36	-		STP	2018
		9	DP	112.4	136.0	140.2	DP_Chr09_450543908	8.77	combined	-	-	-	-		STP	2017
		9	DP	117.4	136.0	140.2	DP_Chr09_450543908	5.84	WGN26	18.5	46.65	-		0.58	STP	2017
		9	DP	100.2	132.9	140.2	DP_Chr09_450543925	9.95	combined	-	-	-	-		STP	2018
		9	DP	85.3	111.1	140.2	DP_Chr09_419434458	4.00	WGN07	15.6	142.32	-		0.66	STP	2018
Threshability	%	5	CP	64.2	65.8	66.6	CP_Chr05_184445451	45.77	combined	-	-	-	-		STP	2017
		5	CP	2.1	65.9	71.9	CP_Chr05_219902957	5.92	WGN15	20	0.89	0.01	-		STP	2017
		5	CP	54.7	65.5	78.2	CP_Chr05_111147787	10.79	WGN26	28	0.92	0.01	-		STP	2017
		5	CP	23.5	65.5	84.6	CP_Chr05_111147787	4.57	WGN36	16.6	0.95	0	-		STP	2017
		5	CP	42.0	60.7	77.2	-	10.02	WGN38	36.3	0.92	0.01	-		STP	2017
		5	CP	63.0	65.5	66.6	CP_Chr05_111147787	10.35	WGN63	34	0.84	0.01	-		STP	2017
		5	CP	63.0	65.5	66.6	CP_Chr05_111147787	47.04	combined	-	-	-	-		STP	2018
		5	CP	38.0	65.5	74.0	CP_Chr05_111147787	5.49	WGN15	19.2	0.85	0.01	-		STP	2018
		5	CP	56.8	63.0	74.0	CP_Chr05_86148562	12.33	WGN26	36.6	0.88	0.02	-		STP	2018
		5	CP	57.8	57.8	89.6	-	3.87	WGN36	15.3	0.9	0	-		STP	2018
		5	CP	56.8	65.5	76.0	CP_Chr05_111147787	13.52	WGN38	48.4	0.84	0.04	-		STP	2018
		5	CP	51.7	65.5	76.0	CP_Chr05_111147787	4.07	WGN39	14.2	0.77	0.01	-		STP	2018
		5	CP	66.6	67.9	75.0	CP_Chr05_232392097	7.52	WGN63	25.5	0.77	0.01	-		STP	2018
		5	CP	64.2	65.8	68.0	CP_Chr05_184445451	60.18	combined	-	-	-	-		TLI	2017
		5	CP	62.6	65.9	70.4	CP_Chr05_219902957	7.50	WGN15	25.6	0.68	0.02	-		TLI	2017
		5	CP	58.8	70.4	81.6	CP_Chr05_247578147	16.83	WGN26	60.8	0.79	0.02	-		TLI	2017
		5	CP	7.1	70.4	90.6	CP_Chr05_247578147	5.45	WGN36	24.5	0.83	0	-		TLI	2017
		5	CP	56.8	60.7	71.9	-	16.84	WGN38	58.9	0.74	0.04	-		TLI	2017
		5	CP	50.7	64.5	71.9	CP_Chr05_86058515	3.86	WGN39	13.6	0.72	0	-		TLI	2017
		5	CP	56.8	59.7	71.9	-	5.60	WGN45	23	0.8	0	-		TLI	2017
		5	CP	63.0	68.6	75.0	CP_Chr05_210699121	6.53	WGN63	23.9	0.73	0	-		TLI	2017
		5	DP	43.0	62.8	69.5	DP_Chr05_196581222	10.81	combined	-	-	-	-		STP	2017
		5	DP	41.0	64.1	69.5	-	7.87	WGN15	21.2	0.89	-		0.01	STP	2017
		5	DP	49.0	62.8	69.5	DP_Chr05_196581222	10.51	combined	-	-	-	-		STP	2018
		5	DP	48.0	64.1	81.1	-	7.62	WGN15	21.5	0.85	-		0.01	STP	2018
		5	DP	45.0	67.9	80.1	-	5.53	WGN15	15.3	0.68	-		0.01	TLI	2017
		11	CP	43.0	46.8	53.3	CP_Chr11_34860853	10.77	combined	-	-	-	-		STP	2018
		11	CP	37.1	59.9	73.8	-	3.42	WGN45	14.1	0.91	0	-		STP	2018

		11	CP	44.0	51.3	55.7	-	12.06	combined	-	-		-	-		TLI	2017
		11	CP	65.2	96.6	96.6	-	4.53	WGN55	16.9	0.84		0	-		TLI	2017
		11	CP	0.0	101.4	125.8	CP_Chr11_228740879	2.05	WGN45	8.9	1		0	-		TLI	2018
		11	DP	67.3	91.3	91.3	-	3.54	WGN39	13.5	0.88	-			0	STP	2017
Floret Site Utilization	%	5	CP	61.7	67.7	86.5	CP_Chr05_231532515	3.80	WGN55	13.1	0.51	-0.02	-			STP	2017
		5	DP	73.1	139.3	146.4	-	3.51	WGN38	16.3	0.43	-			-0.03	STP	2017
		5	DP	100.9	138.6	144.4	DP_Chr05_331923141	10.90	combined	-	-		-	-		TLI	2017
		5	DP	88.1	138.6	142.3	DP_Chr05_331923141	5.12	WGN38	21.3	0.32	-			-0.04	TLI	2017
		5	DP	135.6	139.3	146.4	-	11.26	combined	-	-		-	-		TLI	2018
		5	DP	82.1	142.5	146.4	DP_Chr05_402137083	3.79	WGN26	15.4	0.03	-			-0.01	TLI	2018
		5	DP	87.1	137.6	146.4	-	4.63	WGN38	18.5	0.03	-			-0.02	TLI	2018
		14	DP	133.6	153.4	171.5	DP_Chr14_485886485	8.23	combined	-	-		-	-		STP	2018
		14	DP	119.3	145.2	170.5	DP_Chr14_270680873	3.88	WGN36	15.2	0.63	-			0.05	STP	2018
		14	DP	110.7	118.5	145.2	DP_Chr14_453518583	9.16	combined	-	-		-	-		TLI	2017
Height	cm	5	DP	67.9	88.1	125.8	DP_Chr05_360809367	9.18	combined	-	-		-	-		STP	2017
		5	DP	50.0	109.2	146.4	DP_Chr05_385257903	5.77	WGN26	59.9	120.13	-			-7.67	STP	2017
		5	DP	94.9	135.7	146.4	DP_Chr05_337482310	11.21	combined	-	-		-	-		TLI	2017
		5	DP	76.1	142.5	146.4	DP_Chr05_402137083	4.27	WGN26	16.4	112.8	-			-5.51	TLI	2017
		5	DP	38.2	109.8	146.4	-	3.76	WGN38	40.4	100.22	-			-8.78	TLI	2017
Thousand Grain Weight	g	9	CP	102.3	112.0	140.2	CP_Chr09_423583766	9.88	combined	-	-		-	-		STP	2017
		9	CP	122.5	132.8	140.2	CP_Chr09_186879709	11.82	combined	-	-		-	-		STP	2018
		9	CP	99.3	134.8	140.2	CP_Chr09_457696997	3.40	WGN26	12	8.48	0.33	-			STP	2018
		9	DP	108.1	133.5	140.2	-	4.94	WGN26	17.5	9.69	-			0.47	STP	2017
		15	CP	94.3	116.6	145.7	-	8.21	combined	-	-		-	-		STP	2018
		15	DP	29.8	31.6	37.8	DP_Chr15_103067033	10.47	combined	-	-		-	-		STP	2017
		15	DP	12.9	36.0	39.8	DP_Chr15_160752803	4.31	WGN39	15.5	9.98	-			-0.53	STP	2017
		15	DP	0.0	17.4	39.8	DP_Chr15_213770399	3.54	WGN39	14	8.87	-			-0.45	STP	2018
		17	DP	66.9	73.1	75.9	DP_Chr17_192910115	7.87	combined	-	-		-	-		STP	2017
		17	DP	19.8	22.8	29.7	DP_Chr17_27063510	9.83	combined	-	-		-	-		STP	2018

[a] LG, linkage group

[b] Map used in linkage mapping, DP for donor parent, or CP for common parent

[c-e] Genetic positions in centimorgans (cM) of the 2-LOD drop-off interval

[f] Locus at the LOD peak where "-" indicates the peak was in an interval

[g] Maximum LOD significance value

[h] Analysis whether the QTL was detected in the combined (including all families) analysis or within the analysis of a specific family. Allele effects are not reported in the combined analysis
[i] Overall mean
[j] Difference between alleles in the common parent
[k] Difference between alleles in the donor parent
[l] Loc, location
[m] Yr, year

Figure 3.1. Images depicting the forms of seed shattering in intermediate wheatgrass: A) floret shattering; B) brittle rachis. Black bars are 1 cm.

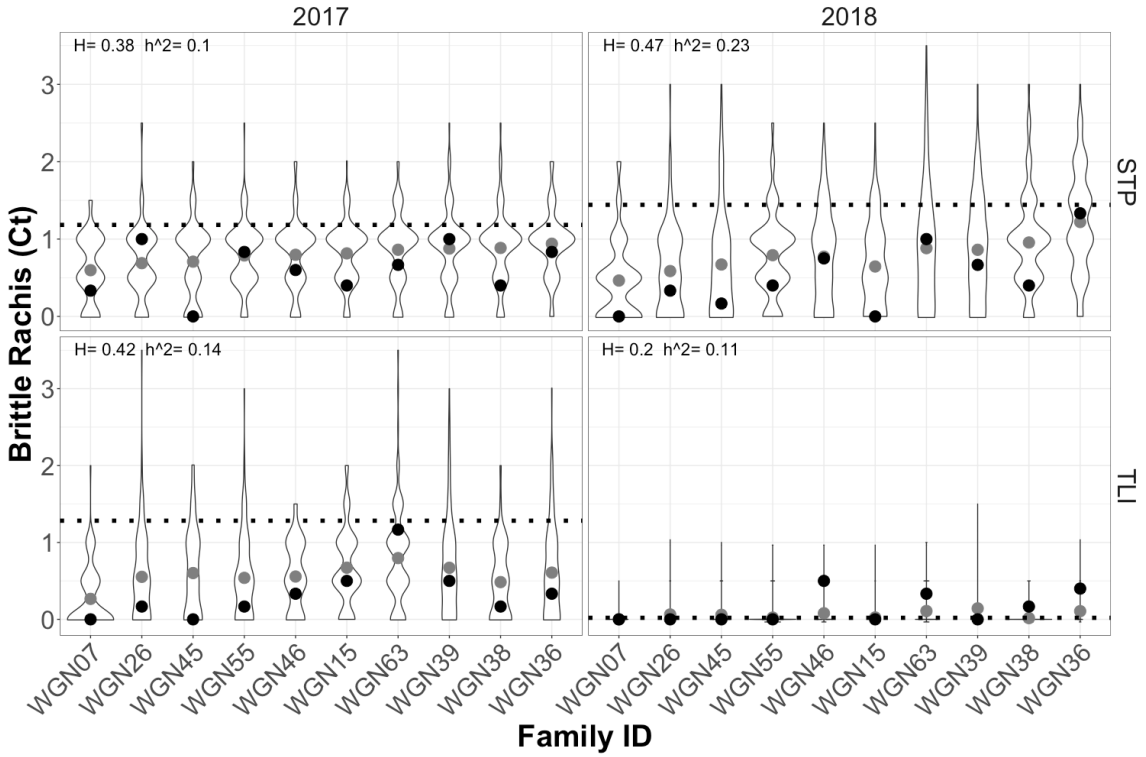
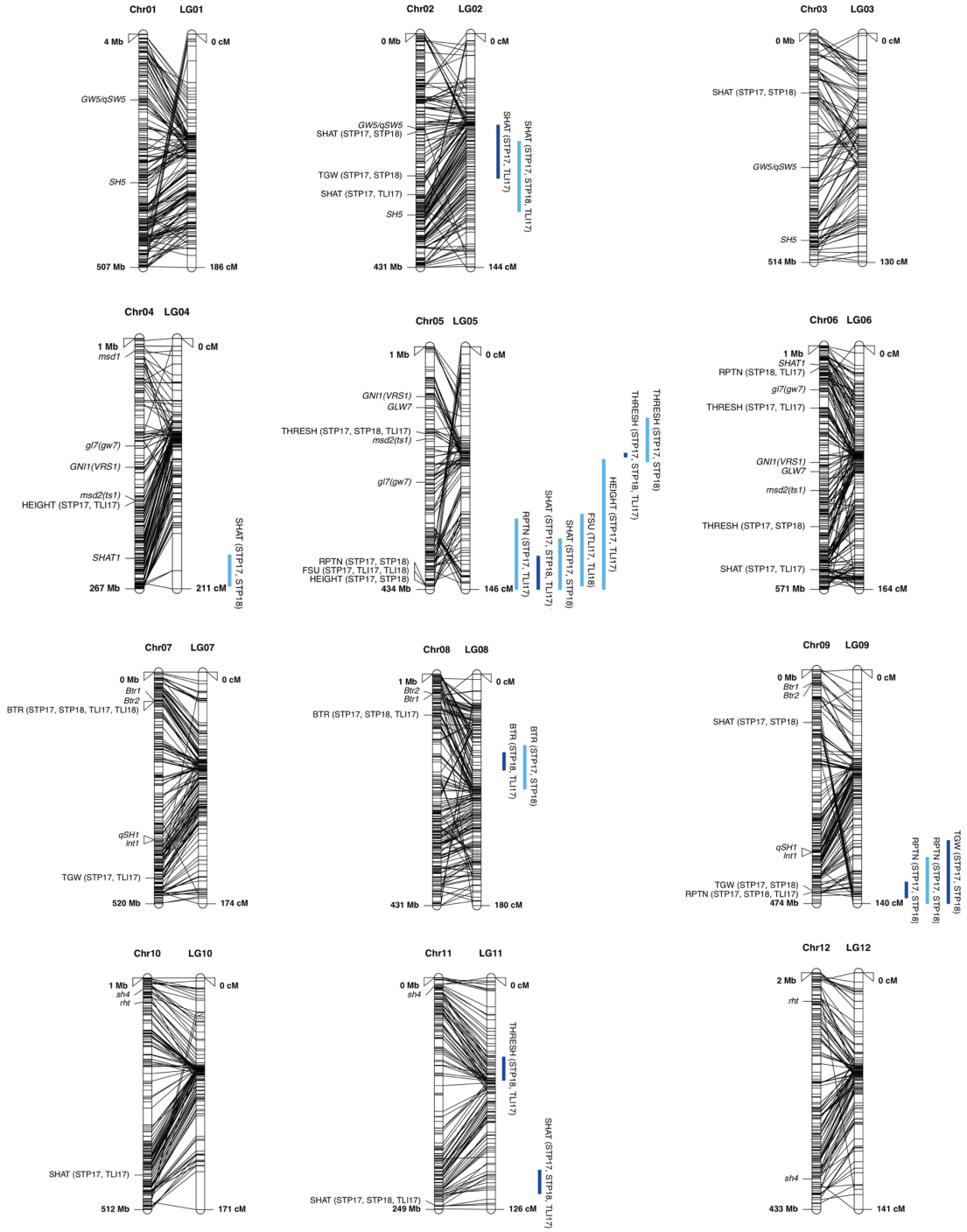


Figure 3.2. Performance of the ten intermediate wheatgrass NAM families for brittle rachis (average count per 3 spikes tested per genet) across two environments: St. Paul, MN (STP) and The Land Institute in Salina, KS (TLI) in 2017 and 2018. Black dots indicate parent means, gray dots indicate progeny means. The horizontal black dotted line indicates the mean of the common parent. Broad (H) and narrow (h^2) sense heritabilities are reported in the top left.


Chr01 LG01
4 Mb
0 cM
GW5/qSW5
SH5
507 Mb
186 cM
Chr02 LG02
0 Mb
0 cM
GW5/qSW5
SHAT (STP17, STP18)
TGW (STP17, STP18)
SHAT (STP17, TLI17)
SH5
431 Mb
144 cM
SHAT (STP17, TLI7)
SHAT (STP17, STP18, TLI17)
Chr03 LG03
0 Mb
0 cM
SHAT (STP17, STP18)
GW5/qSW5
SH5
514 Mb
130 cM
Chr04 LG04
1 Mb
msd1
0 cM
gl7(gw7)
GNI1(VRS1)
msd2(ts1)
HEIGHT (STP17, TLI17)
SHAT1
267 Mb
211 cM
SHAT (STP17, STP18)
Chr05 LG05
1 Mb
0 cM
GNI1(VRS1)
GLW7
THRESH (STP17, STP18, TLI17)
msd2(ts1)
gl7(gw7)
RPTN (STP17, STP18)
FSU (STP17, STP18, TLI17)
HEIGHT (STP17, STP18)
434 Mb
146 cM
SHAT (STP17, STP18)
FSU (TLI17, TLI18)
SHAT (STP17, STP18)
HEIGHT (STP17, TLI17)
RPTN (STP17, TLI17)
THRESH (STP17, TLI17)
THRESH (STP17, STP18)
Chr06 LG06
1 Mb
SHAT1
RPTN (STP18, TLI17)
gl7(gw7)
THRESH (STP17, TLI17)
GNI1(VRS1)
GLW7
msd2(ts1)
THRESH (STP17, STP18)
SHAT (STP17, TLI17)
571 Mb
164 cM
Chr07 LG07
0 Mb
Btr1
Btr2
BTR (STP17, STP18, TLI17, TLI18)
0 cM
qSH1
Int1
TGW (STP17, TLI17)
520 Mb
174 cM
Chr08 LG08
1 Mb
Btr2
Btr1
0 cM
BTR (STP17, STP18, TLI17)
BTR (STP18, TLI17)
BTR (STP17, STP18)
431 Mb
180 cM
Chr09 LG09
0 Mb
Btr1
Btr2
0 cM
SHAT (STP17, STP18)
qSH1
Int1
TGW (STP17, STP18)
RPTN (STP17, STP18, TLI17)
474 Mb
140 cM
RPTN (STP17, STP18)
RPTN (STP17, STP18)
TGW (STP17, STP18)
Chr10 LG10
1 Mb
sh4
rht
0 cM
SHAT (STP17, TLI17)
512 Mb
171 cM
Chr11 LG11
0 Mb
sh4
0 cM
THRESH (STP18, TLI17)
SHAT (STP17, STP18, TLI17)
249 Mb
126 cM
SHAT (STP17, STP18, TLI17)
Chr12 LG12
2 Mb
rht
0 cM
sh4
433 Mb
141 cM

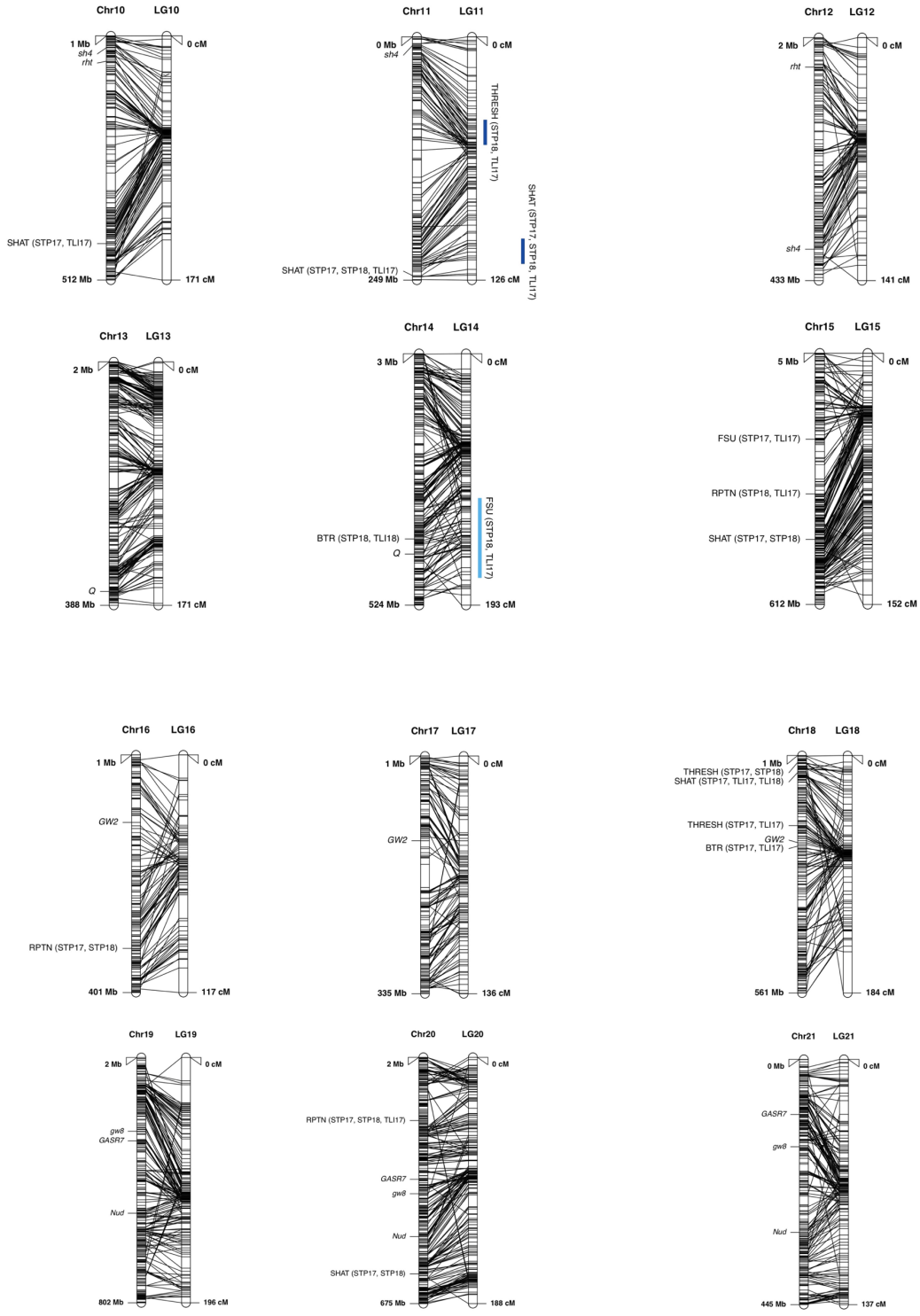


Figure 3.3. Significant markers detected in at least two environments from GWAS (black, left), possible candidate orthologous genes (italic black, left), and 2-LOD drop off intervals

for the combined analysis across populations (blue bars on right). Dark blue indicates a QTL detected in the common parent, light blue indicates a QTL detected in the donor parents. Traits are designated using the following acronyms: SHAT = floret shattering; BTR = brittle rachis; RPTN = reproductive tiller number; THRESH = threshability; FSU = floret site utilization; RPTN = reproductive tiller number; HEIGHT = height, and TGW = thousand grain weight. Traits are followed by the environments in which they were detected in parentheses. Markers that were used in both GWAS and linkage mapping are connected with a black line. Physical distances in mbp and genetic distances in centimorgans (cM) are normalized to comparable lengths. Further information on candidate orthologous genes can be found in Supplementary Table 3.1.

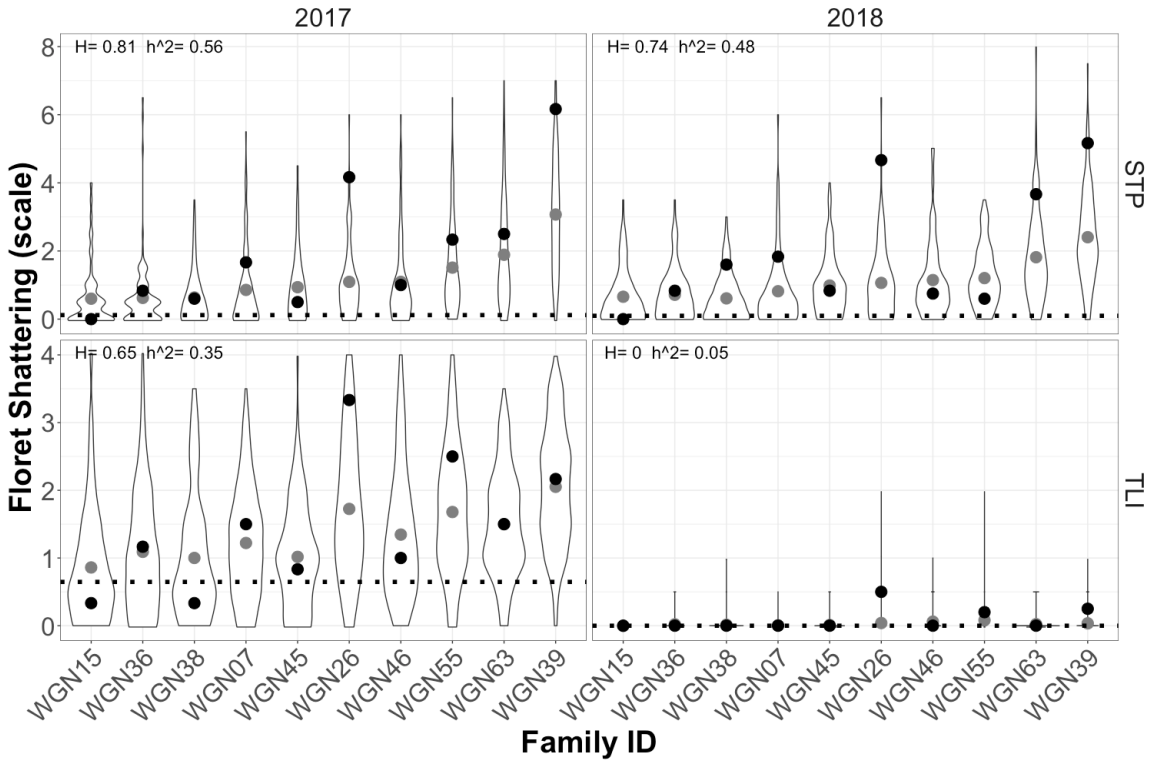


Figure 3.4. Performance of the ten intermediate wheatgrass NAM families for floret shattering (scale) across two locations: St. Paul, MN (STP) and The Land Institute in Salina, KS (TLI) in 2017 and 2018. Black dots indicate parent means, gray dots indicate progeny means. The horizontal black dotted line indicates the mean of the common parent. Broad (H) and narrow (h^2) sense heritabilities are reported in the top left.

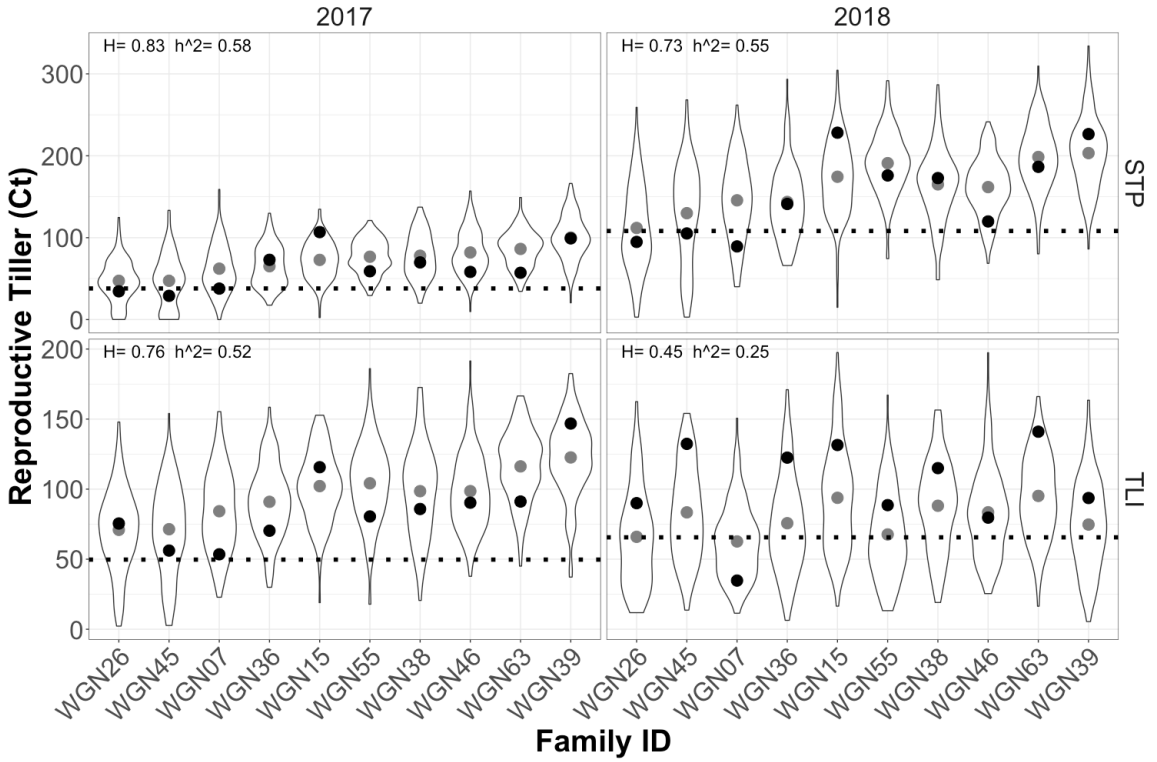


Figure 3.5. Performance of the ten intermediate wheatgrass NAM families for reproductive tiller number (ct) across two locations: St. Paul, MN (STP) and The Land Institute in Salina, KS (TLI) in 2017 and 2018. Black dots indicate parent means, gray dots indicate progeny means. The horizontal black dotted line indicates the mean of the common parent. Broad (H) and narrow (h^2) sense heritabilities are reported in the top left.

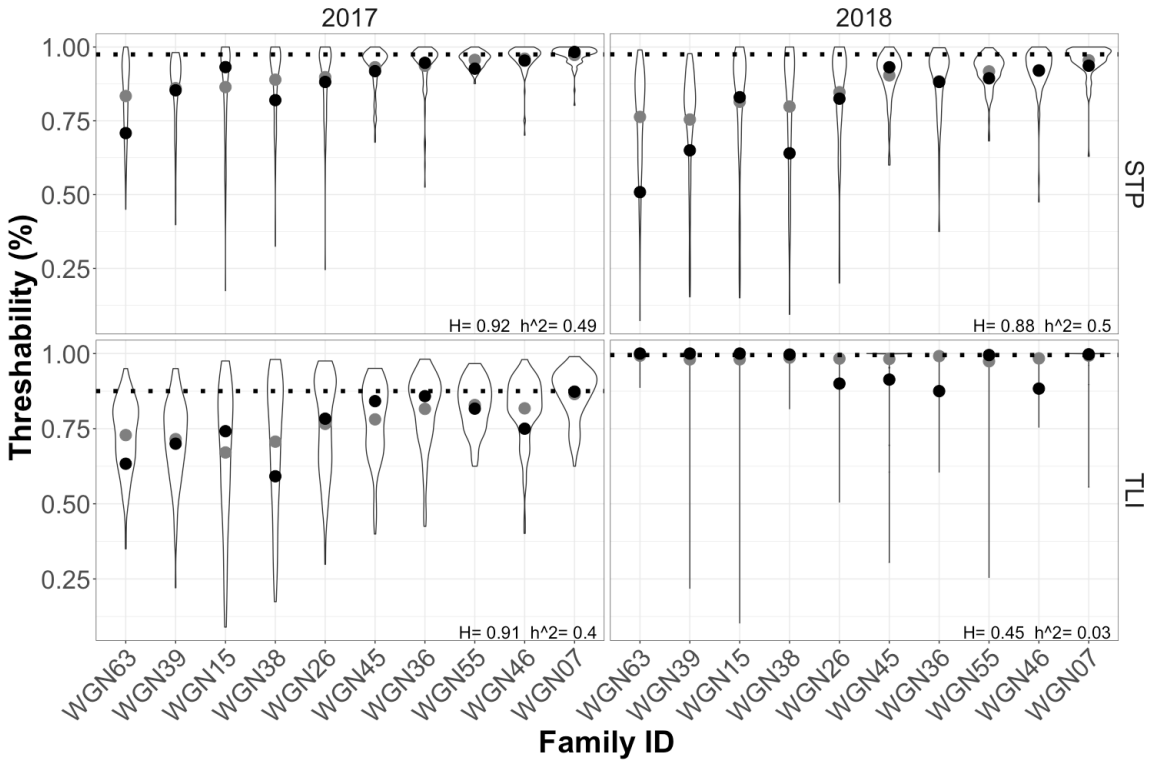


Figure 3.6. Performance of the ten intermediate wheatgrass NAM families for threshability (%) across two locations: St. Paul, MN (STP) and The Land Institute in Salina, KS (TLI) in 2017 and 2018. Black dots indicate parent means, gray dots indicate progeny means. The horizontal black dotted line indicates the mean of the common parent. Broad (H) and narrow (h^2) sense heritabilities are reported in the bottom right.

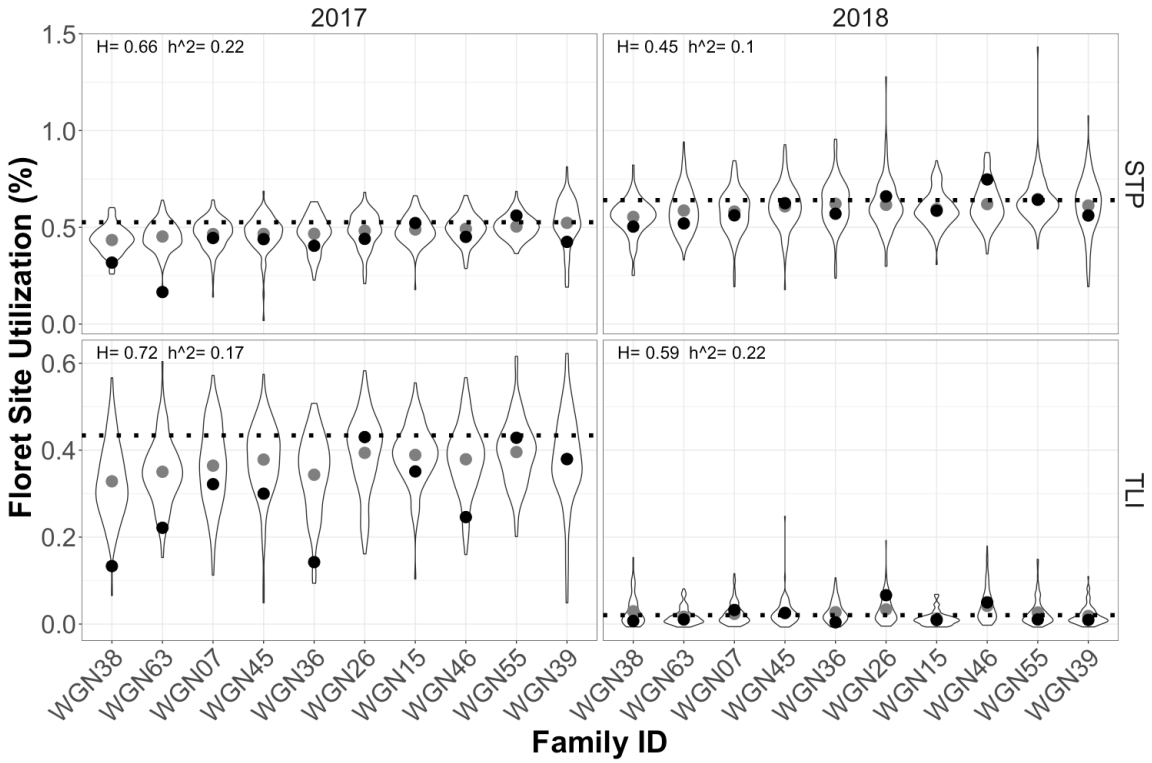


Figure 3.7. Performance of the ten intermediate wheatgrass NAM families for floret site utilization (%) across two locations: St. Paul, MN (STP) and The Land Institute in Salina, KS (TLI) in 2017 and 2018. Black dots indicate parent means, gray dots indicate progeny means. The horizontal black dotted line indicates the mean of the common parent. Broad (H) and narrow (h^2) sense heritabilities are reported in the top left.

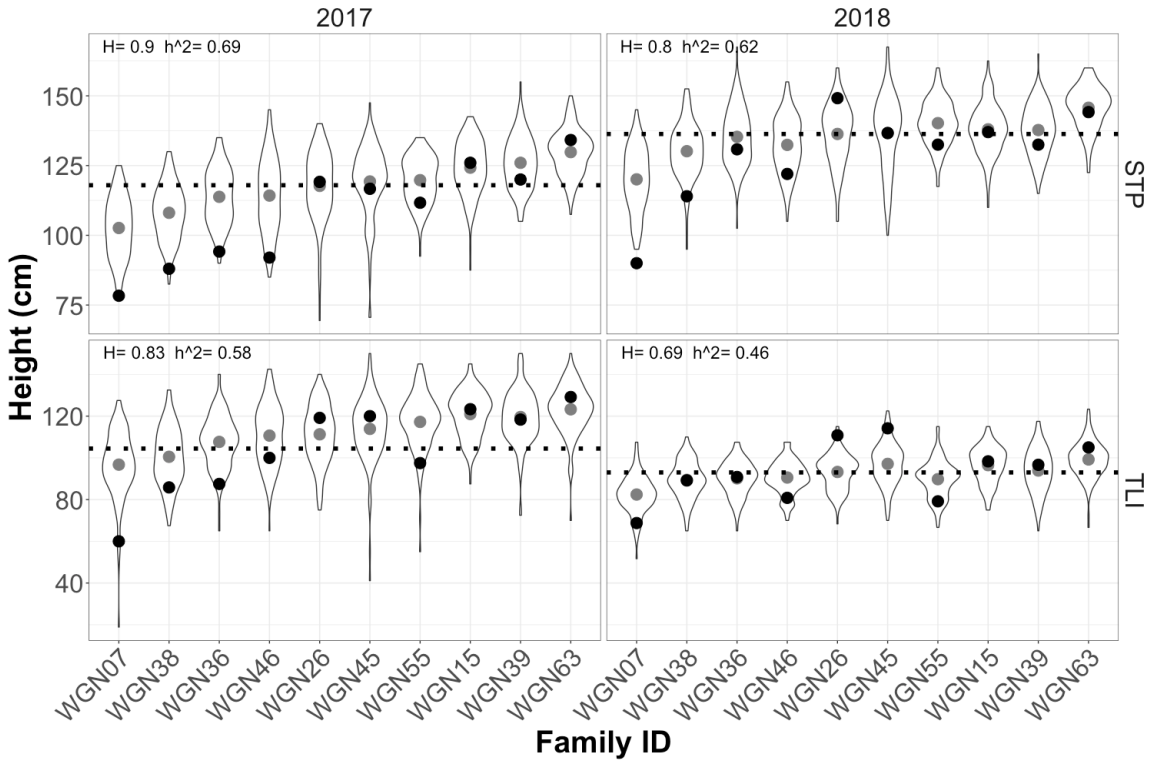


Figure 3.8. Performance of the ten intermediate wheatgrass NAM families for height (cm) across two locations: St. Paul, MN (STP) and The Land Institute in Salina, KS (TLI) in 2017 and 2018. Black dots indicate parent means, gray dots indicate progeny means. The horizontal black dotted line indicates the mean of the common parent. Broad (H) and narrow (h^2) sense heritabilities are reported in the top left.

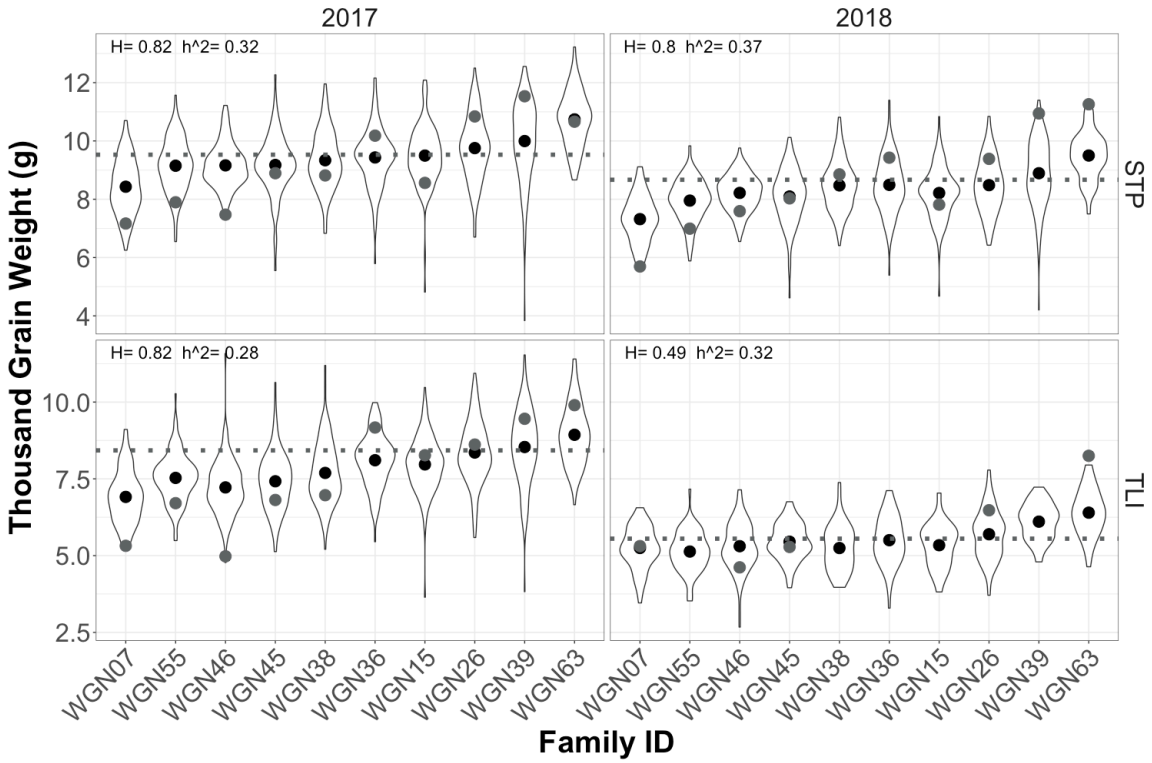


Figure 3.9. Performance of the ten intermediate wheatgrass NAM families for thousand grain weight (g) across two locations: St. Paul, MN (STP) and The Land Institute in Salina, KS (TLI) in 2017 and 2018. Black dots indicate parent means, gray dots indicate progeny means. The horizontal black dotted line indicates the mean of the common parent. Broad (H) and narrow (h^2) sense heritabilities are reported in the top left.

Supplementary Materials

Supplementary Table 3.1. Candidate domestication orthogenes, their marker name as shown in Figure 3.3 with the chromosome (Chr) and position (Pos) of significant BLAST hits in intermediate wheatgrass version 2 draft genome.

Candidate Gene	Marker Name	Chr	Position	Reference	Source (Chr)	Trait
Brittle rachis	Btr1	7	64785511	Civan and Brown 2017	barley (3)	Brittle rachis
Brittle rachis	Btr1	7	64714619	Civan and Brown 2017	barley (3)	Brittle rachis
Brittle rachis	Btr1	7	64705466	Civan and Brown 2017	barley (3)	Brittle rachis
Brittle rachis	Btr1	7	66343485	Civan and Brown 2017	barley (3)	Brittle rachis
Brittle rachis	Btr1	7	65260037	Civan and Brown 2017	barley (3)	Brittle rachis
Brittle rachis	Btr1	8	37809930	Civan and Brown 2017	barley (3)	Brittle rachis
Brittle rachis	Btr1	9	26175732	Civan and Brown 2017	barley (3)	Brittle rachis
Brittle rachis	Btr1	9	25661113	Civan and Brown 2017	barley (3)	Brittle rachis
Brittle rachis	Btr1	9	25969063	Civan and Brown 2017	barley (3)	Brittle rachis
Brittle rachis	Btr2	7	66141197	Civan and Brown 2017	barley (3)	Brittle rachis
Brittle rachis	Btr2	8	37755337	Civan and Brown 2017	barley (3)	Brittle rachis
Brittle rachis	Btr2	9	26060998	Civan and Brown 2017	barley (3)	Brittle rachis
naked	Nud	19	510120134	Yu et al. 2016	barley (7)	Free threshing
naked	Nud	20	491744614	Yu et al. 2016	barley (7)	Free threshing
naked	Nud	21	313566208	Yu et al. 2016	barley (7)	Free threshing
spelt factor (free threshing)	Q	13	365961283	Simons et al. 2006	wheat (5A)	Free threshing
spelt factor (free threshing)	Q	14	418100272	Simons et al. 2006	wheat (5A)	Free threshing
QTL of seed shattering in chr1	qSH1	7	376399378	Konishi et al. 2006	rice (1)	Seed shattering
QTL of seed shattering in chr1	qSH1	7	387186179	Konishi et al. 2006	rice (1)	Seed shattering
QTL of seed shattering in chr1	qSH1	9	369731903	Konishi et al. 2006	rice (1)	Seed shattering
shattering QTL chr4	sh4	10	26172505	Li et al. 2006	rice (4)	Seed shattering
shattering QTL chr4	sh4	11	8879818	Li et al. 2006	rice (4)	Seed shattering

shattering QTL chr4	sh4	12	377628308	Li et al. 2006	rice (4)	Seed shattering
SH1-like gene on chr5	SH5	1	324225555	Yoon et al. 2014	rice (5)	Seed shattering
SH1-like gene on chr5	SH5	2	334447331	Yoon et al. 2014	rice (5)	Seed shattering
SH1-like gene on chr5	SH5	3	464140241	Yoon et al. 2014	rice (5)	Seed shattering
Shattering abortion 1	SHAT1	4	234390930	Zhou et al. 2012	rice (4)	Seed shattering
Shattering abortion 1	SHAT1	6	46055017	Zhou et al. 2012	rice (4)	Seed shattering
Gibberellic Acid-Stimulated Regulator	GASR7	19	274140227	Dong et al. 2014	wheat (7B)	Grain size
Gibberellic Acid-Stimulated Regulator	GASR7	20	335180034	Dong et al. 2014	wheat (7A)	Grain size
Gibberellic Acid-Stimulated Regulator	GASR7	21	99929331	Dong et al. 2014	wheat (7D)	Grain size
grain length QTL chr7	gl7(gw7)	4	115252094	Wang et al. 2015b	rice (7)	Grain size
grain length QTL chr7	gl7(gw7)	4	115252094	Wang et al. 2015a	rice (7)	Grain size
grain length QTL chr7	gl7(gw7)	5	243130888	Wang et al. 2015a	rice (7)	Grain size
grain length QTL chr7	gl7(gw7)	5	243130888	Wang et al. 2015b	rice (7)	Grain size
grain length QTL chr7	gl7(gw7)	6	104589204	Wang et al. 2015a	rice (7)	Grain size
grain length QTL chr7	gl7(gw7)	6	104589204	Wang et al. 2015b	rice (7)	Grain size
grain length and width QTL chr7	GLW7	5	109765347	Si et al. 2016	rice (7)	Grain size
grain length and width QTL chr7	GLW7	6	295560462	Si et al. 2016	rice (7)	Grain size
Grain width and weight QTL chr2	GW2	16	114558207	Hong et al. 2014; also Zhang et al. 2018	wheat (6D), rice (2)	Grain size
Grain width and weight QTL chr2	GW2	17	120189638	Hong et al. 2014; also Zhang et al. 2018	wheat (6D), rice (2)	Grain size
Grain width and weight QTL chr2	GW2	18	206112778	Hong et al. 2014; also Zhang et al. 2018	wheat (6D), rice (2)	Grain size
grain width QTL chr5	GW5/qSW5	1	145568887	Liu et al. 2017	rice (5)	Grain size
grain width QTL chr5	GW5/qSW5	2	175033564	Liu et al. 2017	rice (5)	Grain size
grain width QTL chr5	GW5/qSW5	3	301247434	Liu et al. 2017	rice (5)	Grain size

grain width QTL chr8	gw8	19	242770631	Wang et al. 2012	rice (8)	Grain size
grain width QTL chr8	gw8	20	374584699	Wang et al. 2012	rice (8)	Grain size
grain width QTL chr8	gw8	21	158538535	Wang et al. 2012	rice (8)	Grain size
Reduced height	rht	10	54529961	Peng et al. 1999	wheat (5D)	Plant height
Reduced height	rht	12	54660877	Peng et al. 1999	wheat (5D)	Plant height
Grain number increase 1 (six-rowed spike 1)	GNI1(VRS 1)	4	137847048	Sakuma et al. 2019, Komatsuda et al. 2007	wheat (2), barley (2)	Seed set
Grain number increase 1 (six-rowed spike 1)	GNI1(VRS 1)	5	90182458	Sakuma et al. 2019, Komatsuda et al. 2007	wheat (2), barley (2)	Seed set
Grain number increase 1 (six-rowed spike 1)	GNI1(VRS 1)	5	285940466	Sakuma et al. 2019, Komatsuda et al. 2007	wheat (2), barley (2)	Seed set
Grain number increase 1 (six-rowed spike 1)	GNI1(VRS 1)	6	274400487	Sakuma et al. 2019, Komatsuda et al. 2007	wheat (2), barley (2)	Seed set
Grain number increase 1 (six-rowed spike 1)	GNI1(VRS 1)	6	397252383	Sakuma et al. 2019, Komatsuda et al. 2007	wheat (2), barley (2)	Seed set
multiseeded 1	msd1	4	14231438	Jiao et al. 2018	sorghum (7)	Seed set
multiseeded 1	msd1	4	14269905	Jiao et al. 2018	sorghum (7)	Seed set
multiseeded 2 (tasselseed 1)	msd2(ts1)	4	172880508	Gladman et al. 2019	sorghum (6)	Seed set
multiseeded 2 (tasselseed 1)	msd2(ts1)	5	166007365	Gladman et al. 2019	sorghum (6)	Seed set
multiseeded 2 (tasselseed 1)	msd2(ts1)	6	339906809	Gladman et al. 2019	sorghum (6)	Seed set
Low number of tillers-1	lnt1	9	369736470	_Dabbert et al. 2010	barley (3)	Tillering
Low number of tillers-1	lnt1	7	376403726	_Dabbert et al. 2010	barley (3)	Tillering

Supplementary Table 3.2. Results from analysis of variance for a combined analysis across all environments for each trait using linear mixed effects modeling. A combined analysis of variance was not conducted for brittle rachis and floret shattering due to the Poisson, zero-inflated nature of the data where a generalized linear model was required.

Reproductive Tillers (ct)

	Sum Sq[a]	Mean Sq[b]	NumDF[c]	DenDF[d]	F value	Pr(>F)	PVE[e]
famID	2882.78	320.31	9	4441.62	126.90	3.34E-213	0.11
loc	15.03	15.03	1	2.00	5.95	0.13488767	0.00
year	6441.37	6441.37	1	4488.12	2552.02	0	0.24
famID:loc	414.58	46.06	9	4409.87	18.25	3.81E-30	0.02
famID:year	492.31	54.70	9	4487.72	21.67	2.45E-36	0.02
loc:year	16141.91	16141.91	1	4488.20	6395.31	0	0.60
famID:loc:year	593.22	65.91	9	4487.74	26.11	2.36E-44	0.02

Threshability

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)	PVE
famID	10.06	1.12	9	4563.17	80.48	4.96E-139	0.05
loc	0.25	0.25	1	2.05	18.25	0.048631399	0.00
year	58.94	58.94	1	4199.50	4241.77	0	0.28
famID:loc	2.20	0.24	9	4537.88	17.56	6.43E-29	0.01
famID:year	1.80	0.20	9	4203.32	14.38	3.74E-23	0.01
loc:year	134.04	134.04	1	4199.66	9646.90	0	0.63
famID:loc:year	3.95	0.44	9	4203.26	31.62	3.93E-54	0.02

Height

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)	PVE
famID	156098.97	17344.33	9	4397.96	216.28	0	0.15
loc	38790.05	38790.05	1	2.00	483.71	0.002055383	0.04
year	3628.46	3628.46	1	4464.19	45.25	1.96E-11	0.00
famID:loc	2472.92	274.77	9	4365.08	3.43	0.000326297	0.00
famID:year	28340.21	3148.91	9	4463.44	39.27	6.93E-68	0.03
loc:year	825193.25	825193.25	1	4464.19	10290.10	0	0.77

	famID:loc:year	10489.33	1165.48	9	4463.44	14.53	1.88E-23	0.01

Floret Site Utilization

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)	PVE
famID	1.24	0.14	9	4437.79	18.88	2.70E-31	0.01
loc	5.78	5.78	1	2.01	792.01	0.001234213	0.04
year	24.77	24.77	1	4315.73	3393.05	0	0.17
famID:loc	0.30	0.03	9	4404.99	4.64	3.91E-06	0.00
famID:year	0.58	0.06	9	4318.69	8.85	2.61E-13	0.00
loc:year	115.96	115.96	1	4315.71	15883.60	0	0.78
famID:loc:year	0.27	0.03	9	4318.71	4.08	3.15E-05	0.00

Thousand Grain Weight

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)	PVE
famID	637.99	70.89	9	5411.85	133.97	9.85E-229	0.13
loc	394.12	394.12	1	2.12	744.83	0.000974785	0.08
year	3218.17	3218.17	1	3965.66	6081.92	0	0.68
famID:loc	21.50	2.39	9	5377.81	4.51	6.16E-06	0.00
famID:year	41.74	4.64	9	3908.49	8.77	3.72E-13	0.01
loc:year	384.69	384.69	1	3966.25	727.02	3.55E-147	0.08
famID:loc:year	31.47	3.50	9	3908.20	6.61	2.01E-09	0.01

[a] Sums of squares
[b] Mean squares
[c] Numerator degrees of freedom
[d] Denominator degrees of freedom
[e] PVE, proportion of variance explained by each term

Supplementary Table 3.3. Results from analysis of variance for mixed effects linear models within each environment for each trait. Brittle rachis and floret shattering were analyzed as generalized mixed effects linear models.

Reproductive Tillers (ct)

STP 2017

Source	Df[a]	Sum Sq[b]	Mean Sq[c]	F value	Pr(>F)	PVE[d]
famID	9	2604.81	289.42	215.96	3.00E-239	0.23
rep	1	234.14	234.14	174.71	3.24E-37	0.02
famID:plantID3	1190	6802.37	5.72	4.27	3.76E-124	0.61
Residuals	1141	1529.14	1.34	NA	NA	

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	3033.25	337.03	130.65	6.96E-168	0.20
rep	1	1062.40	1062.40	411.86	2.97E-78	0.07
famID:plantID3	1188	8162.30	6.87	2.66	1.06E-59	0.54
Residuals	1123	2896.83	2.58	NA	NA	

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	1789.43	198.83	133.47	1.25E-170	0.20
rep	1	121.97	121.97	81.88	6.21E-19	0.01
famID:plantID3	1193	5465.54	4.58	3.08	4.39E-77	0.60
Residuals	1125	1675.88	1.49	NA	NA	

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	987.67	109.74	27.23	9.20E-43	0.07
rep	1	215.90	215.90	53.57	4.84E-13	0.02
famID:plantID3	1188	7700.56	6.48	1.61	1.02E-15	0.58
Residuals	1087	4380.83	4.03	NA	NA	

Threshability

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	13.83	1.54	311.17	1.85E-296	0.19
rep	1	0.07	0.07	14.82	0.00012498	0.00
famID:plantID3	1173	53.14	0.05	9.17	1.23E-255	0.73
Residuals	1112	5.49	0.00	NA	NA	

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	22.02	2.45	231.87	7.86E-249	0.19
rep	1	0.38	0.38	35.79	2.95E-09	0.00
famID:plantID3	1188	82.60	0.07	6.59	1.68E-195	0.71
Residuals	1122	11.84	0.01	NA	NA	

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	11.87	1.32	220.92	1.34E-240	0.14
rep	1	0.00	0.00	0.51	0.47377859	0.00
famID:plantID3	1191	67.32	0.06	9.47	5.46E-263	0.78
Residuals	1114	6.65	0.01	NA	NA	

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	0.35	0.04	3.40	0.00042666	0.01
rep	1	0.05	0.05	4.00	0.04598931	0.00
famID:plantID3	1032	21.86	0.02	1.84	1.32E-16	0.74
Residuals	619	7.14	0.01	NA	NA	

Floret Shattering

STP 2017

Source	Df	Deviance	Resid. Df	Resid. Dev
NULL	NA	NA	2266.00	5351.31
famID	9	1253.85	2257.00	4097.46
rep	1	5.68	2256.00	4091.78

famID:plantID3	1165	3270.82	1091.00	820.95

STP 2018

Source	Df	Deviance	Resid. Df	Resid. Dev
NULL	NA	NA	2303.00	3727.53
famID	9	676.16	2294.00	3051.36
rep	1	0.64	2293.00	3050.72
famID:plantID3	1184	2320.55	1109.00	730.17

TLI 2017

Source	Df	Deviance	Resid. Df	Resid. Dev
NULL	NA	NA	2301.00	2769.90
famID	9	310.83	2292.00	2459.07
rep	1	1.47	2291.00	2457.60
famID:plantID3	1187	1772.47	1104.00	685.12

TLI 2018

Source	Df	Deviance	Resid. Df	Resid. Dev
NULL	NA	NA	2096.00	68.05
famID	9	1.10	2087.00	66.94
rep	1	0.84	2086.00	66.11
famID:plantID3	1173	38.90	913.00	27.21

Brittle Rachis

STP 2017

Source	Df	Deviance	Resid. Df	Resid. Dev
NULL	NA	NA	2285.00	739.23
famID	9	21.75	2276.00	717.49
rep	1	0.20	2275.00	717.28
famID:plantID3	1172	444.50	1103.00	272.78

STP 2018

Source	Df	Deviance	Resid. Df	Resid. Dev
NULL	NA	NA	2311.00	1460.73
famID	9	94.52	2302.00	1366.20

rep	1	0.64	2301.00	1365.56
famID:plantID3	1187	874.96	1114.00	490.60

TLI 2017

Source	Df	Deviance	Resid. Df	Resid. Dev
NULL	NA	NA	2314.00	1109.89
famID	9	42.55	2305.00	1067.34
rep	1	0.14	2304.00	1067.20
famID:plantID3	1191	673.32	1113.00	393.89

TLI 2018

Source	Df	Deviance	Resid. Df	Resid. Dev
NULL	NA	NA	2151.00	146.09
famID	9	4.65	2142.00	141.45
rep	1	1.96	2141.00	139.48
famID:plantID3	1182	80.75	959.00	58.73

Height

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	143618.95	15957.66	483.94	0	0.34
rep	1	846.26	846.26	25.66	4.75E-07	0.00
famID:plantID3	1187	247084.38	208.16	6.31	9.03E-188	0.58
Residuals	1122	36997.55	32.97	NA	NA	

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	101036.63	11226.29	201.27	2.43E-227	0.26
rep	1	0.51	0.51	0.01	0.924025919	0.00
famID:plantID3	1183	221174.58	186.96	3.35	3.90E-88	0.57
Residuals	1126	62805.49	55.78	NA	NA	

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	156498.60	17388.73	212.77	4.33E-236	0.23

rep	1	3040.00	3040.00	37.20	1.46E-09	0.00
famID:plantID3	1194	414697.92	347.32	4.25	7.55E-123	0.62
Residuals	1130	92349.09	81.72	NA	NA	

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	50017.48	5557.50	92.25	1.75E-128	0.15
rep	1	25367.08	25367.08	421.06	9.79E-80	0.08
famID:plantID3	1191	180342.52	151.42	2.51	1.99E-53	0.56
Residuals	1125	67776.03	60.25	NA	NA	

Floret Site Utilization

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	1.34	0.15	31.37	3.83E-49	0.06
rep	1	0.21	0.21	44.88	3.34E-11	0.01
famID:plantID3	1172	14.84	0.01	2.66	2.13E-58	0.69
Residuals	1095	5.21	0.00	NA	NA	

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
famID	9	1.24	0.14	10.01	7.24E-15	0.03
rep	1	0.50	0.50	36.67	1.93E-09	0.01
famID:plantID3	1186	28.24	0.02	1.74	2.46E-20	0.63
Residuals	1077	14.77	0.01	NA	NA	

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	0.97	0.11	26.24	2.75E-41	0.04
rep	1	0.04	0.04	9.61	0.001982893	0.00
famID:plantID3	1190	16.09	0.01	3.31	2.63E-85	0.74
Residuals	1103	4.51	0.00	NA	NA	

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE

famID	9	0.12	0.01	21.78	8.79E-33	0.06
rep	1	0.07	0.07	115.27	4.90E-25	0.04
famID:plantID3	1084	1.33	0.00	1.99	8.73E-23	0.68
Residuals	716	0.44	0.00	NA	NA	

Thousand Grain Weight

STP 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	796.35	88.48	185.29	6.68E-214	0.22
rep	1	0.23	0.23	0.48	0.488344611	0.00
famID:plantID3	1173	2321.87	1.98	4.15	8.04E-117	0.64
Residuals	1108	529.11	0.48	NA	NA	

STP 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	714.62	79.40	190.05	2.27E-219	0.24
rep	1	32.50	32.50	77.80	4.23E-18	0.01
famID:plantID3	1188	1753.12	1.48	3.53	7.63E-96	0.59
Residuals	1134	473.78	0.42	NA	NA	

TLI 2017

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	834.47	92.72	190.35	3.18E-218	0.24
rep	1	12.20	12.20	25.05	6.49E-07	0.00
famID:plantID3	1190	2049.46	1.72	3.54	7.24E-95	0.60
Residuals	1114	542.61	0.49	NA	NA	

TLI 2018

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)	PVE
famID	9	104.71	11.63	35.61	2.19E-36	0.20
rep	1	2.25	2.25	6.87	0.009450345	0.00
famID:plantID3	545	358.94	0.66	2.02	1.67E-08	0.68
Residuals	191	62.40	0.33	NA	NA	

[a] Df, degrees of freedom

[b] Sums of squares

[c] Mean squares
[d] PVE, proportion of variance explained by each term

Supplementary Table 3.4. Min, mean, max and standard deviation for progeny phenotypic values within families for each trait and each environment. This is the same data as presented in Figures 3.2 and 3.4-3.9 of the main text, but raw values are presented here for reference.

		STP 2017				STP 2018				TLI 2017				TLI 2018			
Trait (units)	**Family**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**
Brittle Rachis (ct)	WGN07	-0.01	0.60	1.50	0.43	-0.01	0.46	2.00	0.50	-0.01	0.27	2.00	0.39	-0.03	0.00	0.50	0.05
	WGN15	-0.01	0.81	2.01	0.39	0.00	0.65	2.50	0.56	0.00	0.67	2.00	0.52	-0.03	0.02	0.97	0.13
	WGN26	-0.01	0.69	2.50	0.47	-0.01	0.59	3.00	0.56	-0.01	0.55	3.50	0.57	-0.03	0.06	1.03	0.20
	WGN36	0.00	0.94	2.00	0.41	0.00	1.22	3.00	0.66	-0.01	0.61	3.01	0.60	-0.03	0.11	1.03	0.25
	WGN38	-0.01	0.89	2.50	0.47	-0.01	0.96	3.00	0.60	-0.01	0.48	2.00	0.48	-0.03	0.02	0.50	0.10
	WGN39	-0.01	0.88	2.50	0.48	-0.01	0.86	3.00	0.66	-0.01	0.67	3.00	0.64	-0.03	0.14	1.50	0.30
	WGN45	-0.01	0.71	2.00	0.50	-0.01	0.67	3.00	0.69	-0.01	0.60	2.01	0.58	-0.03	0.06	1.00	0.19
	WGN46	-0.01	0.80	2.00	0.45	-0.01	0.77	2.50	0.61	-0.01	0.56	1.50	0.44	-0.03	0.08	0.97	0.21
	WGN55	-0.01	0.79	2.50	0.41	0.00	0.79	2.50	0.54	-0.01	0.54	3.00	0.55	-0.03	0.02	0.97	0.13
	WGN63	-0.01	0.86	2.00	0.46	-0.01	0.88	3.50	0.78	-0.01	0.80	3.50	0.64	-0.03	0.11	1.00	0.25

		STP 2017				STP 2018				TLI 2017				TLI 2018			
	Family	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**
Floret Shattering (scale)	WGN07	-0.04	0.86	5.50	1.12	-0.01	0.82	6.00	1.02	-0.02	1.22	3.50	0.82	-0.02	0.01	0.50	0.07
	WGN15	-0.04	0.60	4.00	0.90	-0.01	0.66	3.50	0.77	0.00	0.86	4.02	0.84	-0.02	0.00	0.02	0.01
	WGN26	-0.04	1.09	6.00	1.06	-0.01	1.06	6.50	1.07	-0.02	1.73	4.00	1.04	-0.02	0.04	1.98	0.22
	WGN36	-0.04	0.62	6.50	1.00	-0.01	0.71	3.50	0.76	-0.02	1.09	4.02	0.90	-0.02	0.02	0.50	0.10
	WGN38	-0.04	0.63	3.50	0.78	-0.01	0.61	3.00	0.69	0.00	1.00	3.50	0.92	-0.02	0.01	0.98	0.11
	WGN39	-0.04	3.07	7.00	1.68	-0.01	2.40	7.50	1.39	0.00	2.05	3.98	0.87	-0.02	0.03	0.98	0.15
	WGN45	-0.04	0.94	4.50	1.03	-0.01	0.99	4.00	0.89	-0.02	1.02	3.98	0.78	-0.02	0.01	0.50	0.08
	WGN46	-0.04	1.09	6.00	1.26	-0.01	1.15	5.01	1.12	0.00	1.35	4.00	0.96	-0.02	0.06	1.00	0.18
	WGN55	0.00	1.51	6.50	1.31	0.00	1.20	3.50	0.87	-0.02	1.68	4.00	0.93	-0.02	0.08	1.98	0.27
	WGN63	-0.04	1.89	7.00	1.53	-0.01	1.82	7.99	1.34	0.00	1.50	3.50	0.78	-0.02	0.02	0.50	0.11

		STP 2017				STP 2018				TLI 2017				TLI 2018			
	Family	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**
Threshability (%)	WGN07	0.80	0.97	1.00	0.03	0.63	0.95	1.00	0.05	0.63	0.87	0.99	0.08	0.56	0.99	1.00	0.04
	WGN15	0.17	0.86	1.00	0.16	0.15	0.81	1.00	0.20	0.09	0.67	0.98	0.22	0.10	0.98	1.00	0.11
	WGN26	0.25	0.90	1.00	0.12	0.20	0.85	1.00	0.16	0.30	0.77	0.98	0.16	0.51	0.98	1.00	0.06
	WGN36	0.53	0.94	1.00	0.08	0.37	0.88	1.00	0.12	0.42	0.82	0.98	0.12	0.61	0.99	1.00	0.04
	WGN38	0.32	0.89	1.00	0.13	0.09	0.80	1.00	0.21	0.17	0.71	0.98	0.22	0.82	0.99	1.00	0.04
	WGN39	0.40	0.86	0.98	0.13	0.15	0.75	0.98	0.19	0.22	0.72	0.95	0.13	0.22	0.98	1.00	0.10
	WGN45	0.68	0.93	1.00	0.07	0.60	0.90	1.00	0.08	0.40	0.78	0.95	0.12	0.31	0.98	1.00	0.08
	WGN46	0.70	0.96	1.00	0.04	0.47	0.92	1.00	0.08	0.40	0.82	0.98	0.10	0.75	0.98	1.00	0.04
	WGN55	0.88	0.96	1.00	0.03	0.68	0.92	1.00	0.06	0.63	0.83	0.97	0.08	0.25	0.97	1.00	0.10
	WGN63	0.45	0.83	1.00	0.13	0.07	0.76	0.99	0.17	0.35	0.73	0.95	0.12	0.89	0.99	1.00	0.02

		STP 2017				STP 2018				TLI 2017				TLI 2018			
	Family	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**
Reproductive Tiller (ct)	WGN07	0.0	62.2	158.9	29.2	40.1	145.6	261.9	48.0	22.9	84.3	155.3	28.7	11.5	62.7	150.6	26.3
	WGN15	2.5	72.9	134.8	22.8	15.0	174.4	304.4	45.9	19.0	102.2	152.8	24.8	16.5	93.8	197.5	36.8
	WGN26	0.1	47.4	124.7	25.2	3.0	111.9	259.2	51.7	2.3	71.0	147.9	29.0	11.8	66.0	162.4	34.2
	WGN36	17.6	65.2	129.8	24.1	66.0	143.6	293.4	43.2	30.0	91.0	158.5	25.6	6.4	75.7	171.0	34.4
	WGN38	20.0	78.0	137.3	26.8	48.6	165.1	286.7	44.4	20.6	98.6	172.6	31.8	19.1	88.2	156.5	31.5
	WGN39	20.5	100.1	166.2	26.9	86.2	203.3	334.0	43.3	37.3	122.6	182.5	27.5	5.5	74.7	163.5	31.7
	WGN45	0.1	47.4	133.4	31.7	3.0	129.9	268.3	59.4	2.7	71.4	154.0	31.5	13.7	83.4	154.1	33.4
	WGN46	9.8	82.0	157.0	26.4	68.7	161.8	241.4	34.7	37.8	98.6	191.5	26.8	25.4	83.5	197.4	31.4
	WGN55	29.4	76.7	121.0	20.3	74.4	191.0	291.7	37.2	17.9	104.2	186.0	28.1	13.3	67.7	167.1	30.8
	WGN63	34.5	86.3	149.0	21.6	80.3	198.6	309.5	39.7	45.1	116.3	166.5	24.7	16.4	95.2	166.1	29.6

		STP 2017				STP 2018				TLI 2017				TLI 2018			
	Family	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**
Floret Site Utilization (%)	WGN07	0.14	0.47	0.64	0.08	0.19	0.58	0.84	0.11	0.11	0.36	0.57	0.09	-0.01	0.02	0.12	0.02
	WGN15	0.18	0.49	0.66	0.07	0.31	0.59	0.84	0.09	0.10	0.39	0.55	0.07	-0.01	0.01	0.07	0.02
	WGN26	0.21	0.48	0.68	0.09	0.30	0.62	1.28	0.13	0.16	0.39	0.58	0.09	0.00	0.03	0.19	0.03
	WGN36	0.23	0.47	0.63	0.09	0.24	0.62	0.96	0.12	0.09	0.34	0.51	0.09	-0.01	0.03	0.11	0.02
	WGN38	0.26	0.43	0.60	0.07	0.25	0.55	0.82	0.09	0.07	0.33	0.57	0.09	-0.01	0.03	0.15	0.03
	WGN39	0.19	0.52	0.81	0.12	0.19	0.61	1.08	0.14	0.05	0.38	0.62	0.11	-0.01	0.02	0.11	0.02
	WGN45	0.02	0.47	0.69	0.09	0.18	0.61	0.93	0.12	0.05	0.38	0.57	0.08	-0.01	0.03	0.25	0.03
	WGN46	0.29	0.49	0.66	0.07	0.36	0.62	0.89	0.11	0.16	0.38	0.57	0.08	0.00	0.04	0.18	0.04
	WGN55	0.36	0.50	0.69	0.06	0.39	0.65	1.43	0.11	0.20	0.40	0.62	0.07	-0.01	0.03	0.15	0.03
	WGN63	0.18	0.45	0.64	0.07	0.33	0.59	0.94	0.11	0.15	0.35	0.60	0.08	-0.01	0.02	0.08	0.02

		STP 2017				STP 2018				TLI 2017				TLI 2018			
	Family	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**
Height (cm)	WGN07	77.5	102.6	125.0	10.9	95.0	120.1	145.0	11.3	18.9	96.8	127.5	16.4	51.7	82.4	107.5	9.3
	WGN15	87.5	124.3	142.5	9.7	110.0	138.0	162.5	9.0	87.5	121.0	145.0	10.8	75.0	96.5	115.0	8.5
	WGN26	69.4	117.7	140.0	12.4	105.0	136.3	160.0	10.9	75.0	111.3	140.0	13.9	68.4	93.2	115.0	8.5
	WGN36	90.0	113.8	135.0	10.1	102.5	135.3	167.5	10.2	65.0	107.7	140.0	12.5	65.0	90.2	107.5	8.2
	WGN38	82.5	108.1	130.0	9.8	95.0	130.2	152.5	9.9	67.5	100.5	132.5	13.8	65.0	89.5	110.0	9.0
	WGN39	105.0	126.0	155.0	9.4	115.0	137.7	165.0	8.9	72.5	119.6	150.0	12.6	65.0	93.9	117.5	10.5
	WGN45	70.6	119.3	147.5	13.1	100.0	136.7	167.5	13.5	41.1	113.9	150.0	15.4	70.0	97.1	122.5	10.2
	WGN46	85.0	114.2	145.0	13.1	105.0	132.4	155.0	9.9	65.0	110.7	142.5	14.4	70.0	90.6	107.5	7.8
	WGN55	92.5	119.8	135.0	8.4	117.5	140.2	160.0	8.1	55.0	117.2	145.0	13.7	66.7	89.8	115.0	9.0
	WGN63	107.5	129.8	150.0	8.1	122.5	145.7	160.0	7.7	70.0	123.3	150.0	11.8	66.7	99.2	123.4	8.7

		STP 2017				STP 2018				TLI 2017				TLI 2018			
	Family	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**	**Min**	**Mean**	**Max**	**SD**
Thousand Grain Weight (g)	WGN07	6.25	8.43	10.70	0.95	5.54	7.32	9.11	0.80	5.23	6.91	9.11	0.86	3.46	5.25	6.56	0.69
	WGN15	4.82	9.50	12.09	1.02	4.68	8.21	10.83	0.88	3.65	7.97	10.47	0.91	3.82	5.34	7.04	0.72
	WGN26	6.70	9.75	12.50	1.11	6.43	8.48	10.84	0.93	5.59	8.36	10.94	1.02	3.71	5.70	7.79	0.82
	WGN36	5.79	9.43	12.16	1.01	5.39	8.49	11.40	0.96	5.45	8.11	9.98	0.86	3.29	5.50	7.12	0.79
	WGN38	6.83	9.34	11.96	0.99	6.41	8.47	10.81	0.85	5.21	7.69	11.19	1.01	3.97	5.25	7.38	0.75
	WGN39	3.84	10.00	12.56	1.39	4.20	8.89	11.40	1.17	3.83	8.54	11.53	1.21	4.80	6.11	7.23	0.59
	WGN45	5.55	9.18	12.27	1.04	4.62	8.09	10.12	0.93	5.13	7.42	10.64	0.93	3.95	5.45	6.75	0.60
	WGN46	7.37	9.16	11.22	0.76	6.55	8.22	9.76	0.61	4.77	7.22	11.58	0.89	2.67	5.31	7.14	0.79
	WGN55	6.55	9.15	11.57	0.90	5.88	7.95	9.83	0.72	5.49	7.53	10.27	0.78	3.53	5.13	7.16	0.66
	WGN63	8.66	10.74	13.22	0.95	7.50	9.50	11.32	0.81	6.66	8.93	11.40	0.97	4.64	6.40	7.95	0.75

Supplementary Table 3.5. Results from the Pearson correlation analysis of the difference between parental phenotype values and the standard deviation among their progeny.

Trait	Estimate	P-value	Estimate	P-value	Estimate	P-value	Estimate	P-value
Brittle Rachis	0.11	0.76	-0.62	0.05	-0.53	0.11	0.36	0.31
Floret Shattering	0.73	0.02	0.75	0.01	0.58	0.08	0.66	0.04
Threshability	0.57	0.09	0.75	0.01	0.55	0.10	-0.25	0.49
Reproductive Tillers	-0.30	0.40	-0.33	0.35	-0.40	0.25	0.30	0.40
Floret Site Utilization	-0.08	0.83	-0.35	0.32	0.00	0.99	0.45	0.19
Height	-0.02	0.97	0.23	0.52	0.28	0.43	0.15	0.69
Thousand Grain Weight	-0.02	0.95	-0.01	0.98	-0.33	0.35	0.46	0.44

Supplementary Table 3.6. Pearson correlation analyses between phenotypic emmeans for St. Paul in 2017.

	Brittle Rachis	Floret Shattering	Threshability	Reproductive Tillers	Floret Site Utilization	Height	Mass per Spike	Spike Length	Seed Area	Thousand Grain Weight	Yield per Spike	Emergence Percent	Anthesis Score
Brittle Rachis	1	-	-	-	-	-	-	-	-	-	-	-	-
Floret Shattering	0.02	1	-	-	-	-	-	-	-	-	-	-	-
Threshability	0.03	-0.2***	1	-	-	-	-	-	-	-	-	-	-
Reproductive Tillers	0.09**	0.29***	-0.14***	1	-	-	-	-	-	-	-	-	-
Floret Site Utilization	0.21***	0.15***	0.03	0.2***	1	-	-	-	-	-	-	-	-
Height	0.18***	0.3***	-0.24***	0.46***	0.26***	1	-	-	-	-	-	-	-
Mass per Spike	0.23***	0.24***	0.07*	0.35***	0.46***	0.35***	1	-	-	-	-	-	-
Spike Length	0.11***	0.26***	-0.02	0.36***	0.16***	0.54***	0.62***	1	-	-	-	-	-
Seed Area	0.09**	0.34***	-0.3***	0.2***	0.06*	0.42***	0.42***	0.39***	1	-	-	-	-
Thousand Grain Weight	0.11***	0.33***	-0.16***	0.24***	0.05.	0.42***	0.47***	0.37***	0.87***	1	-	-	-
Yield per Spike	0.25***	0.2***	0.15***	0.36***	0.64***	0.36***	0.94***	0.54***	0.31***	0.39***	1	-	-
Emergence Percent	-0.02	-0.02	0.15***	0.17***	0.09**	-0.05.	0.09**	0.09**	-0.05.	-0.03	0.14***	1	-
Anthesis Score	-0.03	0.09**	0.15***	0.22***	0.14***	0.17***	0.13***	0.16***	-0.04	-0.01	0.21***	0.65***	1

. Significant at the 0.1 probability level

* Significant at the .05 probability level

** Significant at the .01 probability level

*** Significant at the .001 probability level

Supplementary Table 3.7. Pearson correlation analyses between phenotypic emmeans for St. Paul in 2018.

	Brittle Rachis	Floret Shattering	Threshability	Reproductive Tillers	Floret Site Utilization	Height	Mass per Spike	Spike Length	Seed Area	Thousand Grain Weight	Yield per Spike	Emergence Percent	Anthesis Score
Brittle Rachis	1	-	-	-	-	-	-	-	-	-	-	-	-
Floret Shattering	-0.07*	1	-	-	-	-	-	-	-	-	-	-	-
Threshability	0.02	-0.31***	1	-	-	-	-	-	-	-	-	-	-
Reproductive Tillers	0.09**	0.22***	-0.2***	1	-	-	-	-	-	-	-	-	-
Floret Site Utilization	0.16***	0.12***	0.09**	0.06*	1	-	-	-	-	-	-	-	-
Height	0.14***	0.24***	-0.21***	0.4***	0.18***	1	-	-	-	-	-	-	-
Mass per Spike	0.22***	0.24***	-0.03	0.07*	0.42***	0.39***	1	-	-	-	-	-	-
Spike Length	0.14***	0.15***	-0.16***	0.11***	0.12***	0.47***	0.54***	1	-	-	-	-	-
Seed Area	0.1***	0.36***	-0.35***	0.1***	0.07*	0.44***	0.43***	0.42***	1	-	-	-	-
Thousand Grain Weight	0.11***	0.32***	-0.19***	0.11***	0.06*	0.43***	0.5***	0.38***	0.86***	1	-	-	-
Yield per Spike	0.21***	0.23***	0.11***	0.09**	0.54***	0.35***	0.93***	0.42***	0.33***	0.43***	1	-	-
Emergence Percent	-0.17***	0.02	0.19***	-0.18***	0.13***	-0.2***	0.03	-0.33***	-0.22***	-0.16***	0.13***	1	-
Anthesis Score	-0.16***	0.09**	0.17***	-0.02	0.18***	-0.01	0.04	-0.25***	-0.16***	-0.15***	0.18***	0.71***	1

. Significant at the 0.1 probability level

* Significant at the .05 probability level

** Significant at the .01 probability level

*** Significant at the .001 probability level

Supplementary Table 3.8. Pearson correlation analyses between phenotypic emmeans for the Land Institute in Salina, KS in 2017.

	Brittle Rachis	Floret Shattering	Threshability	Reproductive Tillers	Floret Site Utilization	Height	Mass per Spike	Spike Length	Seed Area	Thousand Grain Weight	Yield per Spike	Emergence Percent	Anthesis Score
Brittle Rachis	1	-	-	-	-	-	-	-	-	-	-	-	-
Floret Shattering	0.06*	1	-	-	-	-	-	-	-	-	-	-	-
Threshability	-0.02	-0.16***	1	-	-	-	-	-	-	-	-	-	-
Reproductive Tillers	0.1***	0.15***	-0.16***	1	-	-	-	-	-	-	-	-	-
Floret Site Utilization	0.18***	0.06*	-0.02	0.15***	1	-	-	-	-	-	-	-	-
Height	0.19***	0.21***	-0.22***	0.55***	0.24***	1	-	-	-	-	-	-	-
Mass per Spike	0.1***	0.13***	0.11***	0.09**	0.36***	0.18***	1	-	-	-	-	-	-
Spike Length	0	0.16***	-0.04	0.23***	0.11***	0.42***	0.53***	1	-	-	-	-	-
Seed Area	0.11***	0.29***	-0.3***	0.19***	-0.01	0.36***	0.23***	0.3***	1	-	-	-	-
Thousand Grain Weight	0.11***	0.28***	-0.14***	0.2***	-0.08**	0.37***	0.2***	0.25***	0.82***	1	-	-	-
Yield per Spike	0.19***	0.13***	0.12***	0.22***	0.72***	0.34***	0.76***	0.34***	0.18***	0.2***	1	-	-
Emergence Percent	-0.01	0.06*	0.05.	0.28***	0.04	0.17***	0.22***	0.04	0	0.05.	0.19***	1	-
Anthesis Score	-0.06*	0.03	0.08**	0.16***	0.06.	0.06*	0.27***	0.11***	-0.11***	-0.07*	0.22***	0.67***	1

. Significant at the 0.1 probability level

* Significant at the .05 probability level

** Significant at the .01 probability level

*** Significant at the .001 probability level

Supplementary Table 3.9. Pearson correlation analyses between phenotypic emmeans for the Land Institute in Salina, KS in 2018.

	Brittle Rachis	Floret Shattering	Threshability	Reproductive Tillers	Floret Site Utilization	Height	Mass per Spike	Spike Length	Seed Area	Thousand Grain Weight	Yield per Spike	Emergence Percent	Anthesis Score
Brittle Rachis	1	-	-	-	-	-	-	-	-	-	-	-	-
Floret Shattering	0.1***	1	-	-	-	-	-	-	-	-	-	-	-
Threshability	-0.01	-0.15***	1	-	-	-	-	-	-	-	-	-	-
Reproductive Tillers	0	-0.03	-0.06.	1	-	-	-	-	-	-	-	-	-
Floret Site Utilization	0.04	0.19***	-0.19***	0.19***	1	-	-	-	-	-	-	-	-
Height	0.15***	0.07*	-0.02	0.5***	0.22***	1	-	-	-	-	-	-	-
Mass per Spike	0.14***	0.11***	-0.01	-0.07*	0.23***	0.24***	1	-	-	-	-	-	-
Spike Length	0.13***	0.08**	-0.01	-0.07*	0.03	0.45***	0.44***	1	-	-	-	-	-
Seed Area	0.18***	0.04	-0.18***	-0.05	-0.01	0.27***	0.22***	0.36***	1	-	-	-	-
Thousand Grain Weight	0.19***	0.05	-0.07	-0.03	0.08.	0.33***	0.23***	0.36***	0.78***	1	-	-	-
Yield per Spike	0.08*	0.23***	-0.13***	0.14***	0.93***	0.25***	0.31***	0.12***	0.12**	0.25***	1	-	-
Emergence Percent	-0.01	0	0.01	0.2***	0.3***	0.13***	0.01	-0.33***	-0.12**	-0.04	0.27***	1	-
Anthesis Score	0.03	0.1***	0	0.11***	0.43***	0.15***	0.26***	0.09**	0.03	0.06	0.45***	0.62***	1

. Significant at the 0.1 probability level

* Significant at the .05 probability level

** Significant at the .01 probability level

*** Significant at the .001 probability level

Supplementary Table 3.10. Mean of progeny derived from the common parent (C) or the donor parent (D) as the mother within families across environments for each trait. P-value is the result from a t-test with unequal variance between the two progeny types.

Brittle Rachis (ct)	**STP 2017**			**STP 2018**			**TLI 2017**			**TLI 2018**		
Family	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**
WGN07	0.57	0.63	0.41	0.43	0.5	0.41	0.27	0.27	1	0	0	0.59
WGN15	0.83	0.8	0.59	0.61	0.68	0.49	0.67	0.68	0.91	0.03	0.02	0.76
WGN26	0.71	0.66	0.56	0.57	0.61	0.68	0.53	0.57	0.69	0.05	0.08	0.36
WGN36	0.89	0.98	0.29	1.24	1.2	0.77	0.71	0.52	0.1	0.12	0.09	0.58
WGN38	0.87	0.92	0.64	0.96	0.95	0.97	0.48	0.5	0.82	0	0.04	0.17
WGN39	0.89	0.86	0.71	0.95	0.78	0.14	0.74	0.61	0.26	0.14	0.15	0.91
WGN45	0.74	0.68	0.56	0.76	0.6	0.2	0.58	0.62	0.76	0.08	0.04	0.32
WGN46	0.86	0.73	0.11	0.77	0.77	0.98	0.53	0.58	0.6	0.07	0.09	0.53
WGN55	0.82	0.76	0.46	0.86	0.73	0.19	0.61	0.48	0.19	0.03	0.02	0.76
WGN63	0.78	0.94	0.05	0.75	1.02	0.05	0.69	0.9	0.06	0.06	0.15	0.06

Floret Shattering (scale)	**STP 2017**			**STP 2018**			**TLI 2017**			**TLI 2018**		
Family	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**
WGN07	0.83	0.89	0.77	0.75	0.88	0.49	1.19	1.25	0.7	0	0.02	0.1
WGN15	0.66	0.54	0.47	0.71	0.61	0.48	0.91	0.81	0.49	0	0	0.82
WGN26	1.06	1.13	0.73	1.21	0.91	0.12	1.75	1.69	0.76	0.06	0.02	0.41
WGN36	0.69	0.57	0.54	0.74	0.69	0.75	1.09	1.09	0.98	0	0.04	0.02
WGN38	0.58	0.76	0.33	0.59	0.64	0.74	0.96	1.08	0.57	0.01	0.03	0.44
WGN39	2.79	3.34	0.06	2.28	2.52	0.33	1.94	2.15	0.18	0.02	0.05	0.38
WGN45	0.87	0.99	0.53	0.94	1.02	0.62	1.04	1	0.8	0.01	0.01	0.73
WGN46	1.08	1.11	0.88	1.19	1.11	0.71	1.32	1.37	0.78	0.04	0.08	0.28
WGN55	1.34	1.67	0.16	1.12	1.28	0.27	1.5	1.85	0.04	0.06	0.1	0.47
WGN63	1.83	1.94	0.69	2	1.64	0.13	1.54	1.46	0.57	0.03	0.01	0.38

Threshability (%)	STP 2017			STP 2018			TLI 2017			TLI 2018		
Family	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**
WGN07	0.97	0.97	0.51	0.95	0.96	0.21	0.85	0.88	0.05	1	0.99	0.28
WGN15	0.87	0.86	0.58	0.83	0.8	0.32	0.69	0.65	0.42	0.97	0.99	0.26
WGN26	0.89	0.91	0.25	0.83	0.87	0.21	0.73	0.81	0.01	0.98	0.99	0.13
WGN36	0.94	0.93	0.79	0.89	0.88	0.77	0.82	0.81	0.78	0.99	0.99	0.47
WGN38	0.9	0.87	0.39	0.81	0.76	0.27	0.71	0.7	0.82	0.98	0.99	0.48
WGN39	0.87	0.85	0.38	0.77	0.74	0.25	0.72	0.71	0.86	0.98	0.98	0.91
WGN45	0.93	0.93	0.88	0.9	0.91	0.68	0.77	0.79	0.61	0.97	0.99	0.21
WGN46	0.95	0.97	0.02	0.91	0.93	0.17	0.81	0.82	0.75	0.99	0.98	0.43
WGN55	0.96	0.95	0.14	0.93	0.91	0.03	0.85	0.81	0.02	0.98	0.97	0.31
WGN63	0.84	0.83	0.76	0.76	0.77	0.68	0.73	0.72	0.68	0.99	0.99	0.93

Reproductive Tiller (Ct)	STP 2017			STP 2018			TLI 2017			TLI 2018		
Family	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**
WGN07	70.12	54.48	0	155.21	136.2	0.03	90.75	77.98	0.01	65.28	60.11	0.28
WGN15	78.81	66.76	0	182.5	166.15	0.05	102.22	102.12	0.98	91.81	95.85	0.54
WGN26	50.65	43.88	0.14	123.75	99.4	0.01	78.22	63.58	0	73.48	58.22	0.01
WGN36	70.02	61.19	0.05	153.2	135.7	0.03	90.35	91.46	0.82	75.07	76.2	0.86
WGN38	81.2	70.75	0.08	171.22	151.29	0.02	103.48	87.78	0.03	90.91	82.09	0.18
WGN39	100.83	99.46	0.78	202.44	204.1	0.83	123.07	122.24	0.87	75.2	74.3	0.88
WGN45	42.46	51.48	0.12	120.4	137.45	0.12	67.45	74.75	0.21	80.92	85.62	0.44
WGN46	81.21	82.63	0.77	165.54	158.31	0.25	104	93.78	0.04	88.14	79.32	0.12
WGN55	75.22	78.16	0.42	193.66	188.47	0.44	101.59	106.61	0.32	67.52	67.79	0.96
WGN63	88.02	84.54	0.37	201.1	196.03	0.48	113.43	119.02	0.21	94.11	96.22	0.69

Floret Site Utilization (%)	STP 2017			STP 2018			TLI 2017			TLI 2018		
Family	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**
WGN07	0.47	0.46	0.17	0.59	0.57	0.44	0.39	0.34	0	0.02	0.02	0.69
WGN15	0.49	0.49	0.71	0.61	0.58	0.15	0.4	0.38	0.15	0.01	0.01	0.43
WGN26	0.49	0.47	0.27	0.61	0.62	0.49	0.4	0.38	0.16	0.04	0.03	0.02
WGN36	0.47	0.47	0.82	0.61	0.63	0.32	0.36	0.33	0.16	0.03	0.03	0.96
WGN38	0.44	0.42	0.25	0.55	0.56	0.93	0.35	0.29	0	0.03	0.02	0.15
WGN39	0.52	0.52	0.88	0.64	0.59	0.06	0.38	0.38	0.7	0.02	0.02	0.85
WGN45	0.47	0.47	0.97	0.6	0.62	0.43	0.38	0.38	0.99	0.03	0.02	0.1
WGN46	0.5	0.48	0.17	0.62	0.62	0.67	0.39	0.37	0.12	0.05	0.04	0.09
WGN55	0.51	0.5	0.22	0.67	0.63	0.06	0.4	0.39	0.67	0.02	0.03	0.4
WGN63	0.45	0.46	0.31	0.57	0.6	0.12	0.34	0.36	0.44	0.02	0.01	0.26

Height (cm)	STP 2017			STP 2018			TLI 2017			TLI 2018		
Family	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**
WGN07	107.05	98.27	0	123.93	116.31	0	101.83	91.93	0	83.86	81.03	0.09
WGN15	125.54	122.93	0.13	138.19	137.79	0.8	121.03	120.99	0.98	96.67	96.32	0.82
WGN26	119.88	115.4	0.05	137.42	135.22	0.27	114.12	108.4	0.02	94.53	91.92	0.09
WGN36	113.87	113.73	0.94	136.06	134.69	0.47	108.67	106.84	0.43	90.26	90.1	0.92
WGN38	109.82	104.15	0.01	131.64	126.77	0.03	103.06	94.92	0.01	90.71	86.69	0.04
WGN39	124.11	127.75	0.03	136.89	138.54	0.3	119.59	119.54	0.98	93.03	94.76	0.36
WGN45	117.82	120.56	0.26	135.86	137.3	0.57	111.41	115.98	0.1	97.35	96.91	0.82
WGN46	117.12	111.59	0.02	134.49	130.51	0.03	113.38	108.25	0.05	92.27	89.04	0.02
WGN55	120.79	118.82	0.19	141.31	139.19	0.14	119.09	115.5	0.14	90.19	89.4	0.62
WGN63	131.32	128.36	0.04	147.38	144.11	0.02	122.09	124.43	0.27	100.38	98.08	0.14

Thousand Grain Weight (g)	STP 2017			STP 2018			TLI 2017			TLI 2018		
Family	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**	**C**	**D**	**P-value**
WGN07	8.43	8.43	1	7.29	7.35	0.7	6.77	7.05	0.07	5.3	5.2	0.6
WGN15	9.7	9.29	0.03	8.2	8.23	0.83	7.9	8.04	0.37	5.38	5.22	0.69
WGN26	9.54	10	0.03	8.26	8.71	0.01	8.15	8.57	0.02	5.49	5.99	0.02
WGN36	9.57	9.32	0.18	8.64	8.37	0.15	8.33	7.92	0.01	5.64	5.4	0.23
WGN38	9.33	9.35	0.95	8.39	8.68	0.13	7.69	7.71	0.93	5.18	5.46	0.16
WGN39	10.04	9.96	0.76	8.96	8.83	0.52	8.56	8.52	0.85	6.07	6.14	0.7
WGN45	9.18	9.17	0.97	8.06	8.12	0.73	7.28	7.54	0.13	5.55	5.38	0.27
WGN46	9.22	9.11	0.44	8.25	8.19	0.61	7.29	7.16	0.43	5.4	5.22	0.29
WGN55	9.27	9.03	0.14	7.97	7.94	0.85	7.6	7.47	0.35	5.01	5.28	0.11
WGN63	10.84	10.64	0.24	9.6	9.4	0.17	8.86	9.01	0.4	6.36	6.44	0.72

Supplementary Table 3.11. Mean performance of the parents across traits and environments. Raw data as presented in Figures 3.2 and 3.4-3.9 of the main text.

Brittle Rachis (ct)

Parent	STP 2017	STP 2018	TLI 2017	TLI 2018
WGN07	0.33	0	0	0
WGN15	0.4	0	0.5	0
WGN26	1	0.33	0.17	0
WGN36	0.83	1.33	0.33	0.4
WGN38	0.4	0.4	0.17	0.17
WGN39	1	0.67	0.5	0
WGN45	0	0.17	0	0
WGN46	0.6	0.75	0.33	0.5
WGN55	0.83	0.4	0.17	0
WGN59	1.18	1.44	1.28	0.02
WGN63	0.67	1	1.17	0.33

Floret Shattering (scale)

Parent	STP 2017	STP 2018	TLI 2017	TLI 2018
WGN07	1.67	1.83	1.5	0
WGN15	0	0	0.33	0
WGN26	4.17	4.67	3.33	0.5
WGN36	0.83	0.83	1.17	0
WGN38	0.6	1.6	0.33	0
WGN39	6.17	5.17	2.17	0.25
WGN45	0.5	0.83	0.83	0
WGN46	1	0.75	1	0
WGN55	2.33	0.6	2.5	0.2
WGN59	0.12	0.1	0.65	0
WGN63	2.5	3.67	1.5	0

Threshability (%)

Parent	STP 2017	STP 2018	TLI 2017	TLI 2018
WGN07	0.98	0.94	0.87	1
WGN15	0.93	0.83	0.74	1
WGN26	0.88	0.82	0.78	0.9
WGN36	0.95	0.88	0.86	0.88
WGN38	0.82	0.64	0.59	1
WGN39	0.85	0.65	0.7	1
WGN45	0.92	0.93	0.84	0.91
WGN46	0.95	0.92	0.75	0.88
WGN55	0.93	0.89	0.82	1
WGN59	0.97	0.98	0.88	1
WGN63	0.71	0.51	0.63	1

Reproductive Tiller (Ct)

Parent	STP 2017	STP 2018	TLI 2017	TLI 2018
WGN07	37.83	89.33	53.5	34.75
WGN15	106.8	228.2	115.67	131.5
WGN26	34.5	94.83	75.5	90
WGN36	73	141.17	70.33	122.5
WGN38	69.8	172.8	85.83	115
WGN39	99.17	226.5	146.83	93.67
WGN45	29	105.17	56.17	132.33
WGN46	58.2	119.83	90.33	79.67
WGN55	59	176	80.5	88.6
WGN59	37.98	108.22	49.71	65.6
WGN63	57.17	186.5	91.17	141

Floret Site Utilization (%)

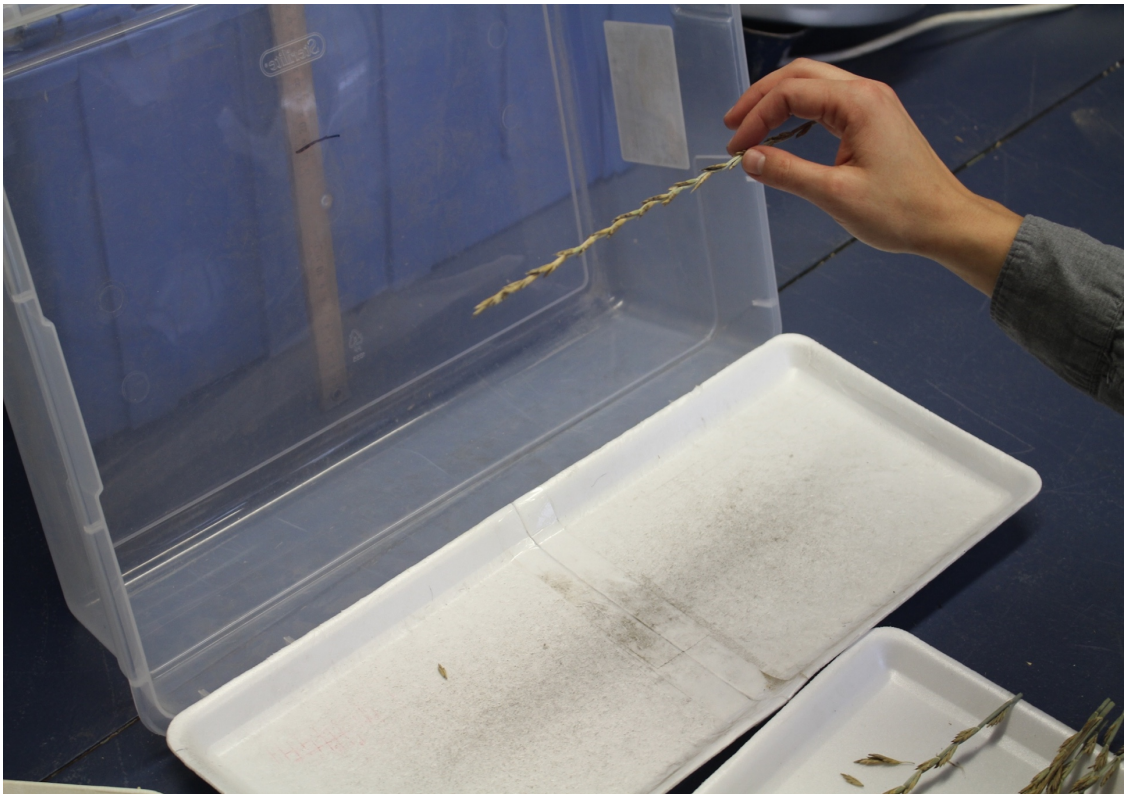
Parent	STP 2017	STP 2018	TLI 2017	TLI 2018
WGN07	0.44	0.56	0.32	0.03
WGN15	0.52	0.59	0.35	0.01
WGN26	0.44	0.66	0.43	0.07
WGN36	0.4	0.57	0.14	0
WGN38	0.32	0.5	0.13	0.01
WGN39	0.42	0.56	0.38	0.01
WGN45	0.44	0.62	0.3	0.03
WGN46	0.45	0.75	0.25	0.05
WGN55	0.56	0.64	0.43	0.01
WGN59	0.53	0.64	0.43	0.02
WGN63	0.17	0.52	0.22	0.01

Height (cm)

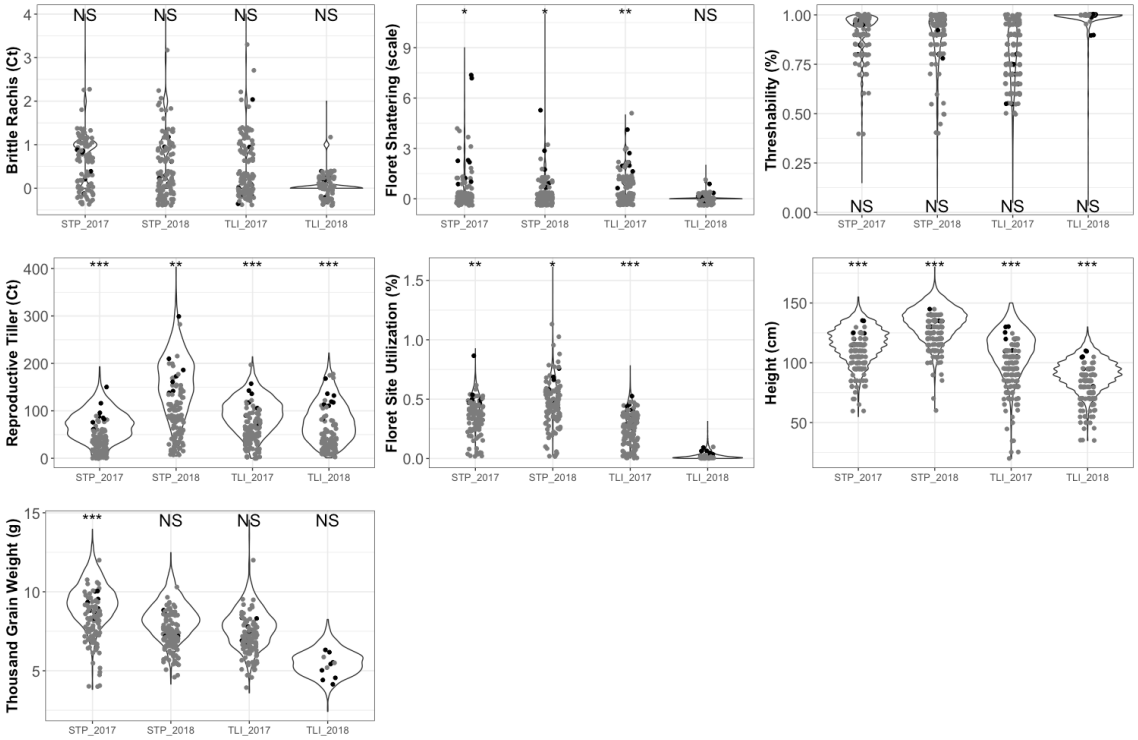
Parent	STP 2017	STP 2018	TLI 2017	TLI 2018
WGN07	78.33	90	60	68.75
WGN15	126	137	123.33	98.33
WGN26	119.17	149.17	119.17	110.83
WGN36	94.17	130.83	87.5	90.83
WGN38	88	114	85.83	89.17
WGN39	120	132.5	118.33	96.67
WGN45	116.67	136.67	120	114.17
WGN46	92	122	100	80.83
WGN55	111.67	132.5	97.5	79.17
WGN59	117.95	136.31	104.44	93
WGN63	134.17	144.17	129.17	105

Thousand Grain Weight (g)

Parent	STP 2017	STP 2018	TLI 2017	TLI 2018
WGN07	7.17	5.69	5.32	5.31
WGN15	8.56	7.82	8.27	-
WGN26	10.84	9.39	8.62	6.48
WGN36	10.18	9.43	9.17	-
WGN38	8.82	8.85	6.97	-
WGN39	11.53	10.94	9.46	-
WGN45	8.89	8.03	6.81	5.28
WGN46	7.47	7.59	4.98	4.62
WGN55	7.9	6.99	6.71	-
WGN59	9.53	8.67	8.43	5.55
WGN63	10.66	11.26	9.9	8.25



Supplementary Figure 3.1. Shattering drop test method. Three IWG spikes were dropped horizontally from 20 cm height three times onto a foam tray and disarticulation events were characterized, counted and recorded.



Supplementary Figure 3.2. Performance of selfs (gray dots) and outcrosses (black dots) relative to F1 progeny for anthesis using the raw data (not emmeans) for all traits across all environments. A "geom_jitter" function was used in ggplot2 to offset data points to increase visibility. Stars at top indicate significance of p-value resulting from a t-test for differences between selfs and F1 progeny, where 0 = ***; 0.001 = **; 0.01 = *; 0.05 = . ; and not significant = NS.

References

Abel, S., R. Gislum, and B. Boelt. 2017. Path and correlation analysis of perennial ryegrass (Lolium perenne L.) seed yield components. J. Agron. Crop Sci. (January)Available at http://doi.wiley.com/10.1111/jac.12202.

Alonso, M.P., N.E. Mirabella, J.S. Panelo, M.G. Cendoya, and A.C. Pontaroli. 2018. Selection for high spike fertility index increases genetic progress in grain yield and stability in bread wheat. Euphytica 214(7): 1–12Available at https://doi.org/10.1007/s10681-018-2193-4.

Altendorf, K. 2020. Phenotypic and Genetic Characterization of Domestication and Yield Component Traits in the Perennial Grain Crop Intermediate Wheatgrass. Doctoral Dissertation, University of Minnesota.

Altschul, S.F., T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25(4): 3389–34017.

Armstead, I.P., L.B. Turner, A.H. Marshall, M.O. Humphreys, I.P. King, and D. Thorogood. 2008. Identifying genetic components controlling fertility in the outcrossing grass species perennial ryegrass (Lolium perenne) by quantitative trait loci analysis and comparative genetics. New Phytol. 178(3): 559–571.

Asay, K.H., and D.A. Johnson. 1997. Genotype x Competition Level Interactions in Crested Wheatgrass (Agropyron desertorum Poaceae: Triticeae). Int. J. Plant Sci. 158(6): 851–855.

Bajgain, P., M.N. Rouse, T.J. Tsilo, G.K. Macharia, S. Bhavani, Y. Jin, and J.A. Anderson. 2016. Nested association mapping of stem rust resistance in wheat using genotyping by sequencing. PLoS One 11(5): 1–22.

Bajgain, P., X. Zhang, and J.A. Anderson. 2019a. Genome-wide association study of yield component traits in intermediate wheatgrass and implications in genomic selection and breeding. G3 Genes, Genomes, Genet. 9(8): 2429–2439.

Bajgain, P., X. Zhang, and J.A. Anderson. 2020. Dominance and G×E interaction effects improve genomic prediction and genetic gain in intermediate wheatgrass (Thinopyrum intermedium). Plant Genome (January): 1–13.

Bajgain, P., X. Zhang, M.K. Turner, R.D. Curland, B. Heim, R. Dill-Macky, C.A. Ishimaru, and J.A. Anderson. 2019b. Characterization of Genetic Resistance to Fusarium Head Blight and Bacterial Leaf Streak in Intermediate Wheatgrass (Thinopyrum intermedium). Agronomy 9(8): 429.

Barrett, J.C., B. Fry, J. Maller, and M.J. Daly. 2005. Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21(2): 263–265.

Bates, D., M. Maechler, B. Bolker, and S. Walker. 2014. Fitting Linear Mixed-Effects Models Using lme4. J. Stat. Softw. 67(1): 1–48.

Berdahl, J.D., and A.B. Frank. 1998. Seed maturity in four cool-season forage grasses. Agron. J. 90(4): 483–488.

Blake, N.K., S.P. Lanning, J.M. Martin, J.D. Sherman, and L.E. Talbert. 2007. Relationship of flag leaf characteristics to economically important traits in two spring wheat crosses. Crop Sci. 47(2): 491–496.

Boelt, B., and B. Studer. 2010. Fodder Crops and Amenity Grasses.

Bouchet, S., M.O. Olatoye, S.R. Marla, R. Perumal, and T. Tesso. 2017. Increased Power To Dissect Adaptive Traits in Global. 206(June): 573–585.
Brettell, R.I.S., P.M. Banks, Y. Cauderon, X. Chen, Z.M. Cheng, P.J. Larkin, and P.M. Waterhouse. 1988. A single wheatgrass chromosome reduces the concentration of barley yellow dwarf virus in wheat. Ann. appl. Biol. 113: 559–603.
Broman, K.W., D.M. Gatti, P. Simecek, N.A. Furlotte, P. Prins, Ś. Sen, B.S. Yandell, and G.A. Churchill. 2019. R/qtl2: Software for mapping quantitative trait loci with high-dimensional data and multiparent populations. Genetics 211(2): 495–502.
Buckler, E.S., J.B. Holland, P.J. Bradbury, C.B. Acharya, P.J. Brown, C. Browne, E. Ersoz, S. Flint-Garcia, A. Garcia, J.C. Glaubitz, M.M. Goodman, C. Harjes, K. Guill, D.E. Kroon, S. Larsson, N.K. Lepak, H. Li, S.E. Mitchell, G. Pressoir, J.A. Peiffer, M.O. Rosas, T.R. Rocheford, M.C. Romay, S. Romero, S. Salvo, H.S. Villeda, H.S. Da Silva, Q. Sun, F. Tian, N. Upadyayula, D. Ware, H. Yates, J. Yu, Z. Zhang, S. Kresovich, and M.D. McMullen. 2009. The genetic architecture of maize flowering time. Science (80-.). 325(5941): 714–718.
Cattani, D. 2016. Selection of a perennial grain for seed productivity across years: Intermediate wheatgrass as a test species. Can. J. Plant Sci. 97(3): CJPS-2016-0280Available at http://www.nrcresearchpress.com/doi/10.1139/CJPS-2016-0280.
Cattani, D.J., and S.R. Asselin. 2018. Has selection for grain yield altered intermediate wheatgrass? Sustain. 10(3).
Charlesworth, D. 1989. Why do plants produce so many more ovules than seeds ? Nature 338: 0–1.
Charlesworth, B., and D. Charlesworth. 1998. Some evolutionary consequences of deleterious mutations. Genetica 102/103: 3–19.
Civáň, P., and T.A. Brown. 2017. A novel mutation conferring the nonbrittle phenotype of cultivated barley. New Phytol. 214(1): 468–472.
Cockram, J., H. Jones, F.J. Leigh, D. O'Sullivan, W. Powell, D.A. Laurie, and A.J. Greenland. 2007. Control of flowering time in temperate cereals: Genes, domestication, and sustainable productivity. J. Exp. Bot. 58(6): 1231–1244.
Colvill, K.E., and C. Marshall. 1984. Tiller dynamics and assimilate partitioning in Lolium perenne with particular reference to flowering. Ann. Appl. Biol. 104(3): 543–557.
Cook, J.P., M.D. McMullen, J.B. Holland, F. Tian, P. Bradbury, J. Ross-Ibarra, E.S. Buckler, and S.A. Flint-Garcia. 2012. Genetic architecture of maize kernel composition in the nested association mapping and inbred association panels. Plant Physiol. 158(2): 824–834.
Covarrubias-Pazaran, G. 2016. Genome-Assisted prediction of quantitative traits using the r package sommer. PLoS One 11(6): 1–15.
Cox, T.S., J.D. Glover, D.L. Van Tassel, C.M. Cox, and L.R. DeHAAN. 2006. Prospects for Developing Perennial Grain Crops. Bioscience 56(8): 649–659.
Cox, T.S., D.L. Van Tassel, C.M. Cox, and L.R. Dehaan. 2010. Progress in breeding perennial grains. Crop Pasture Sci. 61(7): 513–521.
Crain, J., P. Bajgain, J. Anderson, X. Zhang, L. DeHaan, and J. Poland. 2020. Enhancing Crop Domestication Through Genomic Selection, a Case Study of Intermediate Wheatgrass. Front. Plant Sci. 11(March): 1–15.

Culman, S.W., S.S. Snapp, M. Ollenburger, B. Basso, and L.R. DeHaan. 2013. Soil and water quality rapidly responds to the perennial grain Kernza wheatgrass. Agron. J. 105(3): 735–744.
Dabbert, T., R.J. Okagaki, S. Cho, S. Heinen, J. Boddu, and G.J. Muehlbauer. 2010. The genetics of barley low-tillering mutants: Low number of tillers-1 (lnt1). Theor. Appl. Genet. 121(4): 705–715.
Danecek, P., A. Auton, G. Abecasis, C.A. Albers, E. Banks, M.A. DePristo, R.E. Handsaker, G. Lunter, G.T. Marth, S.T. Sherry, G. McVean, and R. Durbin. 2011. The variant call format and VCFtools. Bioinformatics 27(15): 2156–2158.
DeHaan, L., M. Christians, J. Crain, and J. Poland. 2018. Development and Evolution of an Intermediate Wheatgrass Domestication Program. Sustain. 10(1499): 1–19.
DeHaan, L.R., and B.P. Ismail. 2017. Perennial Cereals Provide Ecosystem Benefits. AACC Int. 62(6): 278–281.
DeHaan, L., S. Larson, R.L. López-Marqués, S. Wenkel, C. Gao, and M. Palmgren. 2020. Roadmap for Accelerated Domestication of an Emerging Perennial Grain Crop. Trends Plant Sci. xx(xx): 1–13Available at https://doi.org/10.1016/j.tplants.2020.02.004.
DeHaan, L.R., D.L. Van Tassel, and T.S. Cox. 2004. Perennial grain crops : A synthesis of ecology and plant breeding. Renew. Agric. Food Syst. 20(1): 5–14.
DeHaan, L.R., and D.L. Van Tassel. 2014. Useful insights from evolutionary biology for developing perennial grain crops. Am. J. Bot. 101(10): 1801–1819.
DeHaan, L.R., D.L. Van Tassel, J.A. Anderson, S.R. Asselin, R. Barnes, G.J. Baute, D.J. Cattani, S.W. Culman, K.M. Dorn, B.S. Hulke, M. Kantar, S. Larson, M. David Marks, A.J. Miller, J. Poland, D.A. Ravetta, E. Rude, M.R. Ryan, D. Wyse, and X. Zhang. 2016. A pipeline strategy for grain crop domestication. Crop Sci. 56(3): 917–930.
DeHaan, L.R., S. Wang, S.R. Larson, D.J. Cattani, X. Zhang, and T. Kantarski. 2014. Current efforts to develop perennial wheat and domesticate Thinopyrum intermedium as a perennial grain. Perenn. Crop. food Secur. Proc. FAO Expert Work.: 72–89.
Dewey, D.R. 1962. The genome structure of intermediate wheatgrass. J. Hered. 53(6): 282–290.
Dewey, D.R., and K.H. Lu. 1959. A Correlation and Path-Coefficient Analysis of Components of Crested Wheatgrass Seed Production1. Agron. J. 51(9): 515Available at https://www.agronomy.org/publications/aj/abstracts/51/9/AJ0510090515.
Doebley, J.F., B.S. Gaut, and B.D. Smith. 2006. Review The Molecular Genetics of Crop Domestication. : 1309–1321.
Dofing, S.M., and C.W. Knight. 1992. Alternative Model for Path Analysis of Small-Grain Yield. Crop Sci. 32(2): 487.
Donaldson, E., W.F. Schillinger, and S.M. Dofing. 2001. Straw Production and Grain Yield Relationships in Winter Wheat. 41(February): 362–368.
Dong, L., F. Wang, T. Liu, Z. Dong, A. Li, R. Jing, L. Mao, Y. Li, X. Liu, K. Zhang, and D. Wang. 2014. Natural variation of TaGASR7-A1 affects grain length in common wheat under multiple cultivation conditions. Mol. Breed. 34(3): 937–947.

Earl, D.A., and B.M. vonHoldt. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resour. 4(2): 359–361.
Elgersma, A. 1985. Floret Site Utilization in Grasses: Definitions, Breeding Perspectives and Methodology. J. Appl. Seed Prod. III.
Elgersma, A. 1990. Seed yield related to crop development and to yield components in nine cultivars of perennial ryegrass (Lolium perenne L.). Euphytica 49(2): 141–154.
Elgersma, A., and R. Sniezko. 1988. Cytology of seed development related to floret position in perennial ryegrass (Lolium perenne L.). Euphytica 39(3 Supplement): 59–68.
Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, and S.E. Mitchell. 2011. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS One 6(5): 1–10.
Endelman, J.B. 2011. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. Plant Genome J. 4(3): 250.
Falconer, D.S., and T.F.C. Mckay. 1996. Introduction to Quantitative Genetics. Fourth. Pearson Education Limited, Essex, England.
Fan, X., B. Thompson, and L. Wang. 1999. Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. Struct. Equ. Model. 6(1): 56–83.
Fang, C., T.S. Aamlid, O. Jorgensen, and O.A. Rognli. 2004. Phenotypic and genotypic variation in seed production traits within a full-sib family of meadow fescue. Plant Breed. (123): 241–246.
Fragoso, C.A., M. Moreno, Z. Wang, C. Heffelfinger, L.J. Arbelaez, J.A. Aguirre, N. Franco, L.E. Romero, K. Labadie, H. Zhao, S.L. Dellaporta, and M. Lorieux. 2017. Genetic Architecture of a Rice Nested Association Mapping Population. G3 7(June): g3.117.041608Available at http://g3journal.org/lookup/doi/10.1534/g3.117.041608.
Frahm, C.S., N.E. Tautges, J.M. Jungers, N.J. Ehlke, D.L. Wyse, and C.C. Sheaffer. 2018. Responses of intermediate wheatgrass to plant growth regulators and nitrogen fertilizer. Agron. J. 110(3): 1028–1035.
Frank, A.B., and L. Hoffman. 2017. Society for Range Management Relationship among Grazing Management , Growing Degree-Days , and Morphological Development for Native Grasses on the Northern Great Plains Published by : Society for Range Management Stable URL : http://www.jstor.org/stable/3. J. Range Manag. 42(3): 199–202.
Friebe, B., J. Jiang, W.J. Raupp, R.A. McIntosh, and B.S. Gill. 1996. Characterization of wheat-alien translocations conferring resistance to diseases and pests: current status*. Euphytica 91: 59–87.
Fu, D., P. Szűcs, L. Yan, M. Helguera, J.S. Skinner, J. Von Zitzewitz, P.M. Hayes, and J. Dubcovsky. 2005. Large deletions within the first intron in VRN-1 are associated with spring growth habit in barley and wheat. Mol. Genet. Genomics 273(1): 54–65.
Garcia del Moral, L.F., Y. Rharrabti, D. Villegas, and C. Royo. 2003. Evaluation of Grain Yield and Its Components in Durum Wheat under Medeterranean Conditions: An Ontogenic Approach. Agron. J. 95(2): 266–274.
Garin, V., V. Wimmer, S. Mezmouk, M. Malosetti, and F. van Eeuwijk. 2017. How do

the type of QTL effect and the form of the residual term influence QTL detection in multi-parent populations? A case study in the maize EU-NAM population. Theor. Appl. Genet. 130(8): 1753–1764.
Glover, J.D., J.P. Reganold, L.W. Bell, J. Borevitz, E.C. Brummer, E.S. Buckler, C.M. Cox, T.S. Cox, T.E. Crews, S.W. Culman, L.R. DeHaan, D. Eriksson, B.S. Gill, J. Holland, F. Hu, B.S. Hulke, A.M.H. Ibrahim, W. Jackson, S.S. Jones, S.C. Murray, A.H. Peterson, E. Ploschuk, E.J. Sacks, S. Snapp, D. Tao, D.L. Van Tassel, L.J. Wade, D.L. Wyse, and Y. Xu. 2010. Increased food and ecosystem security via perennial grains. Science (80-.). 328(5986): 1638–1639.
Gomez, K.A., and A.. Gomez. 1984. Statistical Procedures for Agricultural Research. 2nd ed. John Wiley and Sons, New York.
González, F.G., I.I. Terrile, and M.O. Falcón. 2011. Spike fertility and duration of stem elongation as promising traits to improve potential grain number (and yield): Variation in modern Argentinean wheats. Crop Sci. 51(4): 1693–1702.
Grace, J.B. 2006. Structural Equation Modeling and Natural Systems. Cambridge University Press.
Grace, J.B., and J.E. Keeley. 2006. A structural equation model analysis of postfire plant diversity in California shrublands. Ecol. Appl. 16(2): 503–514.
Grafius, J.E. 1956. Components of Yield in Oats: A Geometrical Interpretation. Agron. J. 48(9): 419.
Griffiths, S., R.P. Dunford, G. Coupland, and D. a Laurie. 2003. The Evolution of CONSTANS-Like Gene Families in. Society 131(April): 1855–1867.
Guo, Z., D. Chen, M. Roder, M. Ganal, and T. Schnurbusch. 2018. Genetic dissection of pre-anthesis sub-phase durations during the reproductive spike development of wheat. Plant J. (95): 909–918.
Haas, M., M. Schreiber, and M. Mascher. 2019. Domestication and crop evolution of wheat and barley: Genes, genomics, and future directions. J. Integr. Plant Biol. 61(3): 204–225.
Harlan, J.R., J.M.J. de Wet, and E.G. Price. 1973. Comparative Evolution of Cereals. Evolution (N. Y). (June): 311–325.
Harrell, F.E., and C. Dupont. 2018. Hmisc: Harrell Miscellaneous. R package version 4.1-1. Available at https://cran.r-project.org/web/packages/Hmisc/index.html (verified 8 January 2018).
Hayes, P.M., R.E. Stucker, and G.G. Wandrey. 1989. The Domestication of American Wildrice (Zizania palustris, Poaceae). Econ. Bot. 43(July 1989): 203–214.
Hemshrot, A., A.M. Poets, P. Tyagi, L. Lei, C.K. Carter, C.N. Hirsch, L. Li, G. Brown-Guedira, P.L. Morrell, G.J. Muehlbauer, and K.P. Smith. 2019. Development of a multiparent population for genetic mapping and allele discovery in six-row barley. Genetics 213(2): 595–613.
Hill, E.C., K.A. Renner, C.L. Sprague, and J.E. Fry. 2017. Structural equation modeling of cover crop effects on soil nitrogen and dry bean. Agron. J. 109(6): 2781–2788.
Holland, J.B. 2007. Genetic architecture of complex traits in plants. Curr. Opin. Plant Biol. 10(2): 156–161.
Hong, Y., L. Chen, L.P. Du, Z. Su, J. Wang, X. Ye, L. Qi, and Z. Zhang. 2014. Transcript suppression of TaGW2 increased grain width and weight in bread wheat. Funct.

Integr. Genomics 14(2): 341–349.
Hooper, D., J. Coughlan, and M. Mullen. 2008. Structural Equation Modelling: Guidelines for Determining Model Fit. Electron. J. Bus. Res. Methods 6(1): 53–60Available at https://arrow.dit.ie/cgi/viewcontent.cgi?referer=https://scholar.google.com/&httpsredir=1&article=1001&context=buschmanart.
Hu, L.T., and P.M. Bentler. 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Struct. Equ. Model. 6(1): 1–55.
Huang, X., and B. Han. 2014. Natural Variations and Genome-Wide Association Studies in Crop Plants. Annu. Rev. Plant Biol. 65(1): 531–551Available at http://www.annualreviews.org/doi/10.1146/annurev-arplant-050213-035715.
Hunter, M.C., C.C. Sheaffer, S.W. Culman, and J.M. Jungers. 2020. Effects of defoliation and row spacing on intermediate wheatgrass i: Grain production. Agron. J. (August 2019): 1–16.
Jiao, Y., Y.K. Lee, N. Gladman, R. Chopra, S.A. Christensen, M. Regulski, G. Burow, C. Hayes, J. Burke, D. Ware, and Z. Xin. 2018. MSD1 regulates pedicellate spikelet fertility in sorghum through the jasmonic acid pathway. Nat. Commun. 9(1)Available at http://dx.doi.org/10.1038/s41467-018-03238-4.
Jordan, K.W., S. Wang, F. He, S. Chao, Y. Lun, E. Paux, P. Sourdille, J. Sherman, A. Akhunova, N.K. Blake, M.O. Pumphrey, K. Glover, J. Dubcovsky, L. Talbert, and E.D. Akhunov. 2018. The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping. Plant J. 95(6): 1039–1054.
Jungers, J.M., L.R. Dehaan, K.J. Betts, C.C. Sheaffer, and D.L. Wyse. 2017a. Intermediate Wheatgrass Grain and Forage Yield Responses to Nitrogen Fertilization.
Jungers, J.M., L.R. DeHaan, K.J. Betts, C.C. Sheaffer, and D.L. Wyse. 2017b. Intermediate Wheatgrass Grain and Forage Yield Responses to Nitrogen Fertilization. Agron. J. 0(0): 0Available at https://dl.sciencesocieties.org/publications/aj/abstracts/0/0/agronj2016.07.0438.
Jungers, J.M., L.R. DeHaan, D.J. Mulla, C.C. Sheaffer, and D.L. Wyse. 2019. Reduced nitrate leaching in a perennial grain crop compared to maize in the Upper Midwest, USA. Agric. Ecosyst. Environ. 272: 63–73.
Kalinowski, S.T., M.L. Taper, and T.C. Marshall. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Mol. Ecol. 16(5): 1099–1106.
Kantar, M., C. Tyl, K. Dorn, X. Zhang, J. Jungers, J.M. Kaser, R. Schendel, J. Eckberg, B. Runck, M. Bunzel, N. Jordan, R. Stupar, M.D. Marks, J. Anderson, G. Johnson, C. Sheaffer, T. Schoenfuss, B. Ismail, G. Heimpel, and D. Wyse. 2016. Perennial grain and oilseed crops. Annu. Rev. Plant Biol. 67: 11.1-11.27.
Kantarski, T., S. Larson, X. Zhang, L. DeHaan, J. Borevitz, J. Anderson, and J. Poland. 2016. Development of the first consensus genetic map of intermediate wheatgrass (Thinopyrum intermedium) using genotyping-by-sequencing. Theor. Appl. Genet.: 1–14.

Kazan, K., and R. Lyons. 2016. The link between flowering time and stress tolerance. J. Exp. Bot. 67(1): 47–60.
Kennard, W.C., R.L. Phillips, and R.A. Porter. 2002. Genetic dissection of seed shattering, agronomic, and color traits in American wildrice (Zizania palustris var. interior L.) with a comparative map. Theor. Appl. Genet. 105(6–7): 1075–1086.
Kerber, E.R., and G.G. Rowland. 1974. Origin of the free threshing character in hexaploid wheat. Can. J. Genet. Cytol. 16(1): 145–154.
Knaus, B.J., and N.J. Grünwald. 2017. vcfr: a package to manipulate and visualize variant call format data in R. Mol. Ecol. Resour. 17(1): 44–53.
Komatsuda, T., M. Pourkheirandish, C. He, P. Azhaguvel, K. Kanamori, D. Perovic, N. Stein, A. Graner, T. Wicker, A. Tagiri, U. Lundqvist, T. Fujimura, M. Matsuoka, T. Matsumoto, and M. Yano. 2007. Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. Proc. Natl. Acad. Sci. U. S. A. 104(4): 1424–1429.
Konishi, S., T. Izawa, S.Y. Lin, K. Ebana, Y. Fukuta, T. Sasaki, and M. Yano. 2006. An SNP caused loss of seed shattering during rice domestication. Science (80-.). 312(5778): 1392–1396.
Korte, A., and A. Farlow. 2013. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9(1): 29Available at http://plantmethods.biomedcentral.com/articles/10.1186/1746-4811-9-29.
Lamb, E., S. Shirtliffe, and W. May. 2011. Structural equation modeling in the plant sciences: An example using yield components in oat. Can. J. Plant Sci. 91(4): 603–619.
Lancashire, P.D., H. Bleiholder, T. Van Den Boom, P. Langelüddeke, R. Stauss, E. Weber, and A. Witzenberger. 1991. A uniform decimal code for growth stages of crops and weeds. Ann. Appl. Biol. 119(3): 561–601.
Langmead, B., and S.L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9(4): 357–359.
Large, E.C. 1954. Illustration of the Feekes Scale. Plant Pathol. 3(4): 128–129.
Larkin, P.J., M.T. Newell, R.C. Hayes, J. Aktar, M.R. Norton, S.J. Moroni, and L.J. Wade. 2014. Progress in developing perennial wheats for grain and grazing. Crop Pasture Sci. 65(11): 1147–1164.
Larson, S., L. DeHaan, J. Poland, X. Zhang, K. Dorn, T. Kantarski, J. Anderson, J. Schmutz, J. Grimwood, J. Jenkins, S. Shu, J. Crain, M. Robbins, and K. Jensen. 2019. Genome mapping of quantitative trait loci (QTL) controlling domestication traits of intermediate wheatgrass (Thinopyrum intermedium). Theor. Appl. Genet. 132(8): 2325–2351Available at https://doi.org/10.1007/s00122-019-03357-6.
Lazenby, A., and H.H. Rogers. 1964. Selection criteria in grass breeding II. Effect, on Lolium perenne, of differences in population density, variety and available moisture. J. Agric. Sci. (62): 285–298.
Lenth, R. 2018. emmeans: Estimated Marginal Means aka Least-Squares Means. Available at https://cran.r-project.org/package=emmeans (verified 8 January 2018).
Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16): 2078–2079.

Li, R., M. Li, U. Ashraf, S. Liu, and J. Zhang. 2019. Exploring the Relationships Between Yield and Yield-Related Traits for Rice Varieties Released in China From 1978 to 2017. Front. Plant Sci. 10(May).
Li, C., A. Zhou, and T. Sang. 2006. Rice Domestication by Reducing Shattering. Science (80-.). 311: 1936–1939.
Lipka, A.E., F. Tian, Q. Wang, J. Peiffer, M. Li, P.J. Bradbury, M.A. Gore, E.S. Buckler, and Z. Zhang. 2012. GAPIT: Genome association and prediction integrated tool. Bioinformatics 28(18): 2397–2399.
Liu, J., J. Chen, X. Zheng, F. Wu, Q. Lin, Y. Heng, P. Tian, Z.J. Cheng, X. Yu, K. Zhou, X. Zhang, X. Guo, J. Wang, H. Wang, and J. Wan. 2017. GW5 acts in the brassinosteroid signalling pathway to regulate grain width and weight in rice. Nat. Plants 3(April): 1–7.
Lv, S., W. Wu, M. Wang, R.S. Meyer, M.N. Ndjiondjop, L. Tan, H. Zhou, J. Zhang, Y. Fu, H. Cai, C. Sun, R.A. Wing, and Z. Zhu. 2018. Genetic control of seed shattering during African rice domestication. Nat. Plants 4(6): 331–337Available at http://dx.doi.org/10.1038/s41477-018-0164-3.
Maliepaard, C., J. Jansen, and J.W. Van Ooijen. 1997. Linkage analysis in a full-sib family of an outbreeding plant species: Overview and consequences for applications. Genet. Res. 70(3): 237–250.
Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17(1): 10–12.
Maurer, A., V. Draba, and K. Pillen. 2016. Genomic dissection of plant development and its impact on thousand grain weight in barley through nested association mapping. J. Exp. Bot. 67(8): 2507–2518.
McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M.A. DePristo. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20: 1297–1303.
McMullen, M.D., S. Kresovich, H.S. Villeda, P. Bradbury, H. Li, Q. Sun, S. Flint-Garcia, J. Thornsberry, C. Acharya, C. Bottoms, P. Brown, C. Browne, M. Eller, K. Guill, C. Harjes, D. Kroon, N. Lepak, S.E. Mitchell, B. Peterson, G. Pressoir, S. Romero, M.O. Rosas, S. Salvo, H. Yates, M. Hanson, E. Jones, S. Smith, J.C. Glaubitz, M. Goodman, D. Ware, J.B. Holland, and E.S. Buckler. 2009. Genetic properties of the maize nested association mapping population. Science (80-.). 325(5941): 737–740.
Meyer, R.S., A.E. Duval, and H.R. Jensen. 2012. Patterns and processes in crop domestication: an historical review and quantitative analysis of 203 global food crops. New Phytol. 196: 29–48.
Money, D., K. Gardner, Z. Migicovsky, H. Schwaninger, G. Zhong, and S. Myles. 2015. LinkImpute : Fast and Accurate Genotype Imputation for Nonmodel Organisms. 5(November): 2383–2390.
Moore, K.J., and L.E. Moser. 1995. Quantifying Developmental Morphology of Perennial Grasses. Crop Sci. 35(1): 37–43.
Mortenson, J.S., B.L. Waldron, S.R. Larson, K.B. Jensen, L.R. Dehaan, M.D. Peel, P.G. Johnson, and J.E. Creech. 2019. Quantitative Trait Loci (QTL) for Forage Traits in Intermediate Wheatgrass When Grown as Spaced-Plants versus Monoculture and

Polyculture Swards. : 1–22.
Mueller-Warrant, G.W., and S. Caprice Rosato. 2002. Weed control for stand duration perennial ryegrass seed production: I. Residue removed. Agron. J. 94(5): 1181–1191.
Muggeo, V.M.R. 2008. segmented: An R package to Fit Regression Models with Broken-Line Relationships. R News 8(1): 20–25.
Neyhart, J.L., D. Sweeney, M. Sorrells, C. Kapp, K.D. Kephart, J. Sherman, E.J. Stockinger, S. Fisk, P. Hayes, S. Daba, M. Mohammadi, N. Hughes, L. Lukens, P.G. Barrios, L. Gutiérrez, and K.P. Smith. 2019. Registration of the S2MET Barley Mapping Population for Multi-Environment Genomewide Selection. J. Plant Regist. 13(2): 270–280.
Nice, L.M., B.J. Steffenson, T.K. Blake, R.D. Horsley, K.P. Smith, and G.J. Muehlbauer. 2017. Mapping agronomic traits in a wild barley advanced backcross–nested association mapping population. Crop Sci. 57(3): 1199–1210.
Nishida, H., D. Ishihara, M. Ishii, T. Kaneko, and H. Kawahigashi. 2013. Phytochrome C Is A Key Factor Controlling Long-Day. 163(October): 804–814.
Ogle, D., L. St John, D. Tober, and K. Jensen. 2011. Plant guide for Intermediate wheatgrass (Thinopyrum intermedium). USDA-Natural Resour. Conserv. Serv. Idaho North Dakota Plant Mater. Cent.Available at https://www.nrcs.usda.gov/Internet/FSE_PLANTMATERIALS/publications/idpmspg10369.pdf (verified 17 July 2019).
Oliveira, G. de, N.A. Brunsell, C.E. Sutherlin, T.E. Crews, and L.R. DeHaan. 2018. Energy, water and carbon exchange over a perennial Kernza wheatgrass crop. Agric. For. Meteorol. 249: 120–137.
Van Ooijen, M. 2009. MapQTL6: Software for the mapping of quantitative trait loci in experimental populations. Kyazma B.V., Wageningen, the Netherlands. Kyazma B.V., Wageningen, the Netherlands.
Van Ooijen, J.W. 2011. Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. Genet. Res. (Camb). 93(05): 343–349Available at http://www.journals.cambridge.org/abstract_S0016672311000279.
Ouellette, L.A., R.W. Reid, S.G. Blanchard, and C.R. Brouwer. 2018. LinkageMapView-rendering high-resolution linkage and QTL maps. Bioinformatics 34(2): 306–307.
Pedro, A., R. Savin, M.A.J. Parry, and G.A. Slafer. 2012. Selection for high grain number per unit stem length through four generations from mutants in a durum wheat population to increase yields of individual plants and crops. F. Crop. Res. 129: 59–70Available at http://dx.doi.org/10.1016/j.fcr.2012.01.016.
Philipp, N., H. Weichert, U. Bohra, W. Weschke, A.W. Schulthess, and H. Weber. 2018. Grain number and grain yield distribution along the spike remain stable despite breeding for high yield in winter wheat. PLoS One 13(10).
Pimentel, D., D. Cerasale, R.C. Stanley, R. Perlman, E.M. Newman, L.C. Brent, A. Mullan, and D.T. Chang. 2012. Annual vs. perennial grain production. Agric. Ecosyst. Environ. 161: 1–9Available at http://dx.doi.org/10.1016/j.agee.2012.05.025.
Poland, J.A., P.J. Bradbury, E.S. Buckler, and R.J. Nelson. 2011. Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize.

108(17): 6893–6898.
Poland, J.A., P.J. Brown, M.E. Sorrells, and J. Jannink. 2012. Development of High-Density Genetic Maps for Barley and Wheat Using a Novel Two-Enzyme Genotyping-by- Sequencing Approach. 7(2).
Porter, J.R., and M. Gawith. 1999. Temperatures and the growth and development of wheat: a review. Eur. J. Agron. 10: 23–36Available at %5C%5CGRAEFE%5CJournals%5CEuropean_Journal_of_Agronomy%5C_3.pdf.
Pourkheirandish, M., G. Hensel, B. Kilian, N. Senthil, G. Chen, M. Sameri, P. Azhaguvel, S. Sakuma, S. Dhanagond, R. Sharma, M. Mascher, A. Himmelbach, S. Gottwald, S.K. Nair, A. Tagiri, F. Yukuhiro, Y. Nagamura, H. Kanamori, T. Matsumoto, G. Willcox, C.P. Middleton, T. Wicker, A. Walther, R. Waugh, G.B. Fincher, N. Stein, J. Kumlehn, K. Sato, and T. Komatsuda. 2015. Evolution of the Grain Dispersal System in Barley. Cell 162(3): 527–539Available at http://dx.doi.org/10.1016/j.cell.2015.07.002.
Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155(2): 945–959.
Purugganan, M.D., and D.Q. Fuller. 2009. The nature of selection during plant domestication. Nature 457(7231): 843–848.
Rebetzke, G.J., and R.A. Richards. 2000. Gibberellic acid-sensitive dwarfing genes reduce plant height to increase kernel number and grain yield of wheat. Aust. J. Agric. Res. (51): 235–245.
Rosseel, Y. 2012. lavaan: An R package for structural equation modeling. R package version 0.5-15. J. Stat. Softw. 48(2): 1–36.
Sakuma, S., G. Golan, Z. Guo, T. Ogawa, A. Tagiri, K. Sugimoto, N. Bernhardt, J. Brassac, M. Mascher, G. Hensel, S. Ohnishi, H. Jinno, Y. Yamashita, I. Ayalon, Z. Peleg, T. Schnurbusch, and T. Komatsuda. 2019. Unleashing floret fertility in wheat through the mutation of a homeobox gene. Proc. Natl. Acad. Sci. U. S. A. 116(11): 5182–5187.
Sakuma, S., B. Salomon, and T. Komatsuda. 2011. The domestication syndrome genes responsible for the major changes in plant form in the triticeae crops. Plant Cell Physiol. 52(5): 738–749.
Schauer, A. 1989. Evaluations of intermediate wheatgrass germplasm--1988 Summary. Rodale Press, Inc, Emmaus, PA.
Schmalenbach, I., T.J. March, K. Pillen, T. Bringezu, and R. Waugh. 2011. High-resolution genotyping of wild barley introgression lines and fine-mapping of the threshability locus thresh-1 using the illumina goldengate assay. G3 Genes, Genomes, Genet. 1(3): 187–196.
Schneider, C.A., W.S. Rasband, and K.W. Eliceiri. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9(7): 671–675Available at http://dx.doi.org/10.1038/nmeth.2089.
Si, L., J. Chen, X. Huang, H. Gong, J. Luo, Q. Hou, T. Zhou, T. Lu, J. Zhu, Y. Shangguan, E. Chen, C. Gong, Q. Zhao, Y. Jing, Y. Zhao, Y. Li, L. Cui, D. Fan, Y. Lu, Q. Weng, Y. Wang, Q. Zhan, K. Liu, X. Wei, K. An, G. An, and B. Han. 2016. OsSPL13 controls grain size in cultivated rice. Nat. Genet. 48(4): 447–456.

Simons, K.J., J.P. Fellers, H.N. Trick, Z. Zhang, Y. Tai, B.S. Gill, and J.D. Faris. 2006. Molecular Characterization of the Major Wheat Domestication Gene Q. Genetics 172: 547–555.

Song, Q., L. Yan, C. Quigley, B.D. Jordan, E. Fickus, S. Schroeder, B.-H. Song, Y.-Q. Charles An, D. Hyten, R. Nelson, K. Rainey, W.D. Beavis, J. Specht, B. Diers, and P. Cregan. 2017. Genetic Characterization of the Soybean Nested Association Mapping Population. Plant Genome.

Sood, S., V. Kuraparthy, G. Bai, and B.S. Gill. 2009. The major threshability genes soft glume (sog) and tenacious glume (Tg), of diploid and polyploid wheat , trace their origin to independent mutations at non-orthologous loci. : 341–351.

Spindel, J.E., H. Begum, D. Akdemir, B. Collard, E. Redoña, J.L. Jannink, and S. McCouch. 2016. Genome-wide prediction models that incorporate de novo GWAS are a powerful new tool for tropical rice improvement. Heredity (Edinb). 116(4): 395–408.

Sprunger, C.D., S.W. Culman, G.P. Robertson, and S.S. Snapp. 2017. Perennial grain on a Midwest Alfisol shows no sign of early soil carbon gain. Renew. Agric. Food Syst.: 1–13Available at https://www.cambridge.org/core/product/identifier/S1742170517000138/type/journal_article.

Sreenivasulu, N., and T. Schnurbusch. 2012. A genetic playground for enhancing grain number in cereals. Trends Plant Sci. 17(2): 91–101Available at http://dx.doi.org/10.1016/j.tplants.2011.11.003.

Steinfort, U., S. Fukai, B. Trevaskis, D. Glassop, A. Chan, and M.F. Dreccer. 2017. Vernalisation and photoperiod sensitivity in wheat: The response of floret fertility and grain number is affected by vernalisation status. F. Crop. Res. 203: 243–255Available at http://dx.doi.org/10.1016/j.fcr.2016.10.013.

Studer, B., L.B. Jensen, S. Hentrup, G. Brazauskas, R. Kölliker, and T. Lübberstedt. 2008. Genetic characterisation of seed yield and fertility traits in perennial ryegrass (Lolium perenne L.). Theor. Appl. Genet. 117(5): 781–791.

Szucs, P., I. Karsai, J. Von Zitzewitz, K. Mészáros, L.L.D. Cooper, Y.Q. Gu, T.H.H. Chen, P.M. Hayes, and J.S. Skinner. 2006. Positional relationships between photoperiod response QTL and photoreceptor and vernalization genes in barley. Theor. Appl. Genet. 112(7): 1277–1285.

Taketa, S., S. Amano, Y. Tsujino, T. Sato, D. Saisho, K. Kakeda, M. Nomura, T. Suzuki, T. Matsumoto, K. Sato, H. Kanamori, S. Kawasaki, and K. Takeda. 2008. Barley grain with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. Proc. Natl. Acad. Sci. U. S. A. 105(10): 4062–4067.

Talukder, S.K., and M.C. Saha. 2017. Toward Genomics-Based Breeding in C3 Cool-Season Perennial Grasses. Front. Plant Sci. 8(July): 1–14Available at http://journal.frontiersin.org/article/10.3389/fpls.2017.01317/full.

Tange, O. 2011. GNU Parallel: the command-line power tool. ;login USENIX Mag. 36(1): 42–47Available at https://www.usenix.org/publications/login/february-2011-volume-36-number-1/gnu-parallel-command-line-power-tool.

Van Tassel, D.L., L.R. Dehaan, and T.S. Cox. 2010. Missing domesticated plant forms:

Can artificial selection fill the gap? Evol. Appl. 3(5–6): 434–452.
Tautges, N.E., J.M. Jungers, L.R. Dehaan, D.L. Wyse, and C.C. Sheaffer. 2018. Maintaining grain yields of the perennial cereal intermediate wheatgrass in monoculture v. bi-culture with alfalfa in the Upper Midwestern USA. J. Agric. Sci. 156(6): 758–773.
Team, R.C. 2018. R: A language and environment for statistical computing. Available at https://www.r-project.org/.
Tian, F., P.J. Bradbury, P.J. Brown, H. Hung, Q. Sun, S. Flint-garcia, T.R. Rocheford, M.D. Mcmullen, J.B. Holland, and E.S. Buckler. 2011. Genome-wide association study of leaf architecture in the maize nested association mapping population. 43(2): 6–11.
Trethewey, J., and M. Rolston. 2008. Is the flag leaf important in perennial ryegrass seed production. Agron. Soc. New Zeal. Spec. Publ. 13(Grassland Research and Practice Series No. 14): 67–73.
Turner, A., J. Beales, S. Faure, R.P. Dunford, and D.A. Laurie. 2005. The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. Science (80-.). 310(5750): 1031–1034.
Turner, M.K., L.R. Dehaan, Y. Jin, and J.A. Anderson. 2015. Wheatgrass – Wheat Partial Amphiploids as a Novel Source of Stem Rust and Fusarium Head Blight Resistance. (october 2013).
Tyl, C., and B.P. Ismail. 2019. Compositional evaluation of perennial wheatgrass (Thinopyrum intermedium) breeding populations. Int. J. Food Sci. Technol. 54(3): 660–669.
Tzarfati, R., Y. Saranga, V. Barak, A. Gopher, A.B. Korol, and S. Abbo. 2013. Threshing efficiency as an incentive for rapid domestication of emmer wheat. Ann. Bot. 112(5): 829–837.
Vico, G., S. Manzoni, L. Nkurunziza, K. Murphy, and M. Weih. 2016. Trade-offs between seed output and life span - a quantitative comparison of traits between annual and perennial congeneric species. New Phytol. 209(1): 104–114.
Vos, P.G., M.J. Paulo, R.E. Voorrips, R.G.F. Visser, H.J. van Eck, and F.A. van Eeuwijk. 2017. Evaluation of LD decay and various LD-decay estimators in simulated and SNP-array data of tetraploid potato. Theor. Appl. Genet. 130(1): 123–135.
Wagoner, P. 1990. Perennial Grain New use for intermediate wheatgrass. J. Soil Water Conserv. 45(81–82).
Waldron, B.L., J.G. Robins, M.D. Peel, and K.B. Jensen. 2008. Predicted Effi ciency of Spaced-Plant Selection to Indirectly Improve Tall Fescue Sward Yield and Quality. (April): 443–449.
Wang, X., L. Dong, J. Hu, Y. Pang, L. Hu, G. Xiao, X. Ma, X. Kong, J. Jia, H. Wang, and L. Kong. 2019. Dissecting genetic loci affecting grain morphological traits to improve grain weight via nested association mapping. Theor. Appl. Genet. 132(11): 3115–3128Available at https://doi.org/10.1007/s00122-019-03410-4.
Wang, S., K. Wu, Q. Yuan, X. Liu, Z. Liu, X. Lin, R. Zeng, H. Zhu, G. Dong, Q. Qian, G. Zhang, and X. Fu. 2012. Control of grain size, shape and quality by OsSPL16 in rice. Nat. Genet. 44(8): 950–954Available at http://dx.doi.org/10.1038/ng.2327.
Wang, Y., G. Xiong, J. Hu, L. Jiang, H. Yu, J. Xu, Y. Fang, L. Zeng, E. Xu, J. Xu, W.

Ye, X. Meng, R. Liu, H. Chen, Y. Jing, Y. Wang, X. Zhu, J. Li, and Q. Qian. 2015. Copy number variation at the GL7 locus contributes to grain size diversity in rice. Nat. Genet. 47(8): 944–948.

Warnes, G., G. Gorjanc, F. Leisch, and M. Man. 2019. Package 'genetics.' CRANAvailable at http://cran.uvigo.es/web/packages/genetics/ (verified 17 April 2020).

Watanabe, N., K. Sugiyama, Y. Yamagishi, and Y. Sakata. 2002. Comparative telosomic mapping of homoeologous genes for brittle rachis in tetraploid and hexaploid wheats. 185: 180–185.

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wright, S. 1921. Correlation and causation. J. Agric. Res. 20: 557–580.

Xavier, A., S. Xu, W.M. Muir, and K.M. Rainey. 2015. NAM: Association studies in multiple populations. Bioinformatics 31(23): 3862–3864.

Yan, L., D. Fu, C. Li, A. Blechl, G. Tranquilli, M. Bonafede, A. Sanchez, M. Valarik, S. Yasuda, and J. Dubcovsky. 2006. The wheat and barley vernalization gene VRN3 is an orthologue of FT. Proc. Natl. Acad. Sci. U. S. A. 103(51): 19581–19586.

Yoon, J., L.H. Cho, S.L. Kim, H. Choi, H.J. Koh, and G. An. 2014. The BEL1-type homeobox gene SH5 induces seed shattering by enhancing abscission-zone development and inhibiting lignin biosynthesis. Plant J. 79(5): 717–728.

Yu, J., J.B. Holland, M.D. McMullen, and E.S. Buckler. 2008. Genetic Design and Statistical Power of Nested Association Mapping in Maize. Genetics 551(January): 539–551.

Yu, S., H. Long, G. Deng, Z. Pan, J. Liang, X. Zeng, Y. Tang, N. Tashi, and M. Yu. 2016. A Single Nucleotide Polymorphism of Nud Converts the Caryopsis Type of Barley (Hordeum Vulgare L.). Plant Mol. Biol. Report. 34(1): 242–248.

Zadoks, J.C., T.T. Chang, and C.F. Konzak. 1974. A decimal code for the growth stages of cereals. Weed Res. 14(415): 421.

Zegeye, H., A. Rasheed, F. Makdis, A. Badebo, and F.C. Ogbonnaya. 2014. Genome-wide association mapping for seedling and adult plant resistance to stripe rust in synthetic hexaploid wheat. PLoS One 9(8).

Zhang, X., S.R. Larson, L. Gao, S.L. Teh, L.R. Dehaan, M. Fraser, A. Sallam, T. Kantarski, K. Frels, J. Poland, D. Wyse, and J.A. Anderson. 2017. Uncovering the Genetic Architecture of Seed Weight and Size in Intermediate Wheatgrass through Linkage and Association Mapping. (november): 1–15.

Zhang, X., A. Sallam, L. Gao, T. Kantarski, J. Poland, L.R. DeHaan, D.L. Wyse, and J.A. Anderson. 2016. Establishment and Optimization of Genomic Selection to Accelerate the Domestication and Improvement of Intermediate Wheatgrass. Plant Genome 9(1): 0Available at https://dl.sciencesocieties.org/publications/tpg/abstracts/0/0/plantgenome2015.07.0059.

Zhou, Y., D. Lu, C. Li, J. Luo, B.F. Zhu, J. Zhu, Y. Shangguan, Z. Wang, T. Sang, B. Zhou, and B. Han. 2012. Genetic control of seed shattering in rice by the APETALA2 transcription factor Shattering Abortion1. Plant Cell 24(3): 1034–1048.