Fall 12-2-2022

Deciphering the Genetic Architecture of Key Female Floral Traits for Hybrid Wheat Seed Production

Juan Jimenez
University of Nebraska-Lincoln

DECIPHERING THE GENETIC ARCHITECTURE OF KEY FEMALE FLORAL TRAITS FOR HYBRID WHEAT SEED PRODUCTION

by

Juan David Jimenez Pardo

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Agronomy

Under the Supervision of Professor Katherine Frels

Lincoln, Nebraska

November 2022

DECIPHERING THE GENETIC ARCHITECTURE OF KEY FEMALE FLORAL TRAITS FOR HYBRID WHEAT SEED PRODUCTION

Juan David Jimenez Pardo, M.S.

University of Nebraska, 2022

Advisor: Katherine Frels

Wheat (*Triticum aestivum* L.) is a staple cereal that provides 20% of the calories and proteins in human intake (Ray et al., 2013). Global population is projected to increase to 9.7 billion by 2050. Food production must increase by 70% to feed this future population. Wheat production is in crisis due to political and environmental challenges and is projected to decline by 0.8% in 2022 (FAO, 2022). To ensure food security yield genetic gain must increase by around 1.4% annually. Taking advantage of heterosis, hybrid wheat has the potential to boost grain yield. However, hybrid wheat seed production systems are not profitable due to the cleistogamy of the crop (Longin et al., 2012). Selection of parental lines with beneficial floral traits is necessary to improve outcrossing ability and thus seed set in hybrid wheat production fields. While several studies have focused on the morphological and genetic variation of male floral traits, few have studied in detail the phenotypic and genetic architecture of female floral traits and their crucial importance in hybrid wheat seed production systems. This study aims to unravel the genetic architecture of key female floral traits for hybrid wheat seed production by phenotyping key female floral traits and conducting a genome wide association study to decipher the genetic basis of the phenotyped traits. We studied a panel of winter wheat breeding lines sprayed with the Chemical Hybridizing Agent

Croisor®100. Gape Date, Gape Score, and CHA damage were measured during seven years and genotyped with 44,240 SNP markers. The phenotypic variation was very wide for all female traits in the phenotyped lines. We identified 73 significant marker-trait associations for all assessed traits. Three candidate genes coding for unknown proteins were the most promising and their specific biological function need to be explored. The understanding of the genetic architecture of the female floral traits, and the identified marker-trait associations and candidate genes in this study might serve as a foundation for future studies on developing female floral traits to enhance cross-pollination for effective hybrid wheat seed production.

Acknowledgments

All glory be to GOD. Thank the Lord for his support throughout this journey. Many thanks to my advisors, lab technicians, graduate students, faculty, and staff at UNL for their support. To my parents, thank you for your teaching and training. Thanks to you I am the professional and ethical person I have become today. To my brother in heaven, thank you for your support and for cheering me up every time I fainted. You are always in my heart and in my thoughts. Thank you very much to all my friends who accompanied me through thick and thin, especially Andrea Suazo, Luis Gomez, Kevin Lievano, Jonathan Clavijo and Jose Suazo. There are many who run by your side but few who go with you to the end. Special thanks to Dr. Amir Ibrahim for providing germplasm from Texas A&M University.

Grant Information

This project was supported by the Agriculture and Food Research Initiative Competitive Grant 2020-67013-30872 from the USDA-NIFA

Table of Contents

List of Tables

List of Figures

1. Introduction

Domesticated in the Middle East around 10,000 years ago, wheat (*Triticum aestivum* L.) is considered one of the most important crops for human civilization. Wheat is a staple cereal that provides 20% of the calories and proteins consumed by people (Ray et al., 2013), (FAOSTAT, 2019). Wheat-based foods include bread, pasta, noodles, semolina, snacks, and bakery products. The current global population is around 8.0 billion, and it is expected to reach 9.7 billion in 2050 (UN, 2022). Global human consumption of wheat will expand by 0.9% in 2022/23 reaching 536 million tons (FAO, 2022). Global food production must increase by almost 70% to feed the future population (FAO, 2019). Wheat production is in crisis due to political and environmental challenges. International conflicts, policies, and input costs have increased grain prices and reduced the global availability of grain. The FAO (2022) forecasts that global wheat production will decline by 0.8% in 2022 to 771 million tons.

Plant breeding programs contribute greatly to increasing global agricultural productivity through the development of new crop varieties. Inbred cultivars are the primary end products developed in most global wheat breeding programs (Sade et al., 2022). However, genetic gain for yield in wheat using pure-line breeding is lower than the predicted rate of gain needed to meet wheat demand. Negative recessive characteristics can be inherited through inbreeding because it increases the chance of homozygous recessive alleles in the offspring when one or both parents carry the recessive alleles, thus the harmful traits will not be eliminated. Wheat yield must increase by 1.4% each year to meet anticipated demand, but actual yield growth is less than 1%

(CGIAR, 2016). This indicates that the current systems used in wheat breeding programs need major changes to achieve higher yield genetic gain. Outcrossing crops such as maize and hybridized self-pollinated crops such as rice have higher rates of yield genetic gain.

Hybridization is a feasible method to boost wheat yield to satisfy the upcoming grain demand and guarantee food security. Several studies reported hybrid superiority compared with inbred lines in wheat. Longin et al., (2013) described 10.7% grain yield superiority of wheat hybrids compared to the mean yield of their parents. Koemel et al., (2004) reported a hybrid wheat yield advantage of 10.9% on average over wheat commercial pure lines. Hybrid wheat yield mid-parent heterosis of 24% was reported by Easterly et al., (2020). In the hybridization process, two distinct inbred lines containing desirable traits are crossed to exploit heterosis or hybrid vigor which is the expression of superior performance in hybrid offspring compared to inbred parents (Chu et al., 2012). One of the main advantages of heterosis is that it can enhance and stabilize wheat yield (Mühleisen et al., 2014).

Research for the development of hybrid wheat varieties dates back to the 1960s after the discovery of cytoplasmic male sterility (Singh et al., 2010). The investment in hybrid wheat development from the public and private sectors has been low and intermittent. It is challenging to develop hybrid wheat due to the inefficiency and lack of profitability of hybrid seed production systems (Kempe & Gils, 2011; Longin et al., 2012). Wheat is an autogamous cereal, which reduces its potential for outcrossing. This floral characteristic reduces the amount of hybrid seed produced and consequently, increases the cost of seed production, making the price of hybrid wheat seed unaffordable

for farmers. Thus, the cleistogamous nature of wheat must be optimized to improve outcrossing ability and maximize seed set (El Hanafi et al., 2021). Improving outcrossing ability and maximizing seed set can be achieved by developing elite parental lines that display favorable floral characteristics for cross-pollination and integrating them into the hybridization process.

Investigation of superior floral traits for hybrid wheat requires phenotyping, the measurement of traits expressed in the plant. Once these traits have been characterized, the selection of superior individuals could be implemented to be used as parents for the development of wheat hybrids. The ideal female parent (Figure 1a) would be short, with wide open glumes, and long and receptive stigmas at pollination time (De Vries, 1971), and the male components of the plant should be effectively sterilized (Whitford et al., 2013). It should be paired with an ideal male parent (Figure 1b) that should be taller than the female parent, with good anther extrusion (AE) and should produce viable pollen for long periods. Both parents need to be synchronized to coincide with flowering times when the male is shedding viable pollen and the female has receptive stigmas (Longin et al., 2013). Female plant structures are responsible for seed set in wheat. Understanding the morphology of traits that maximize seed set in the female parent is essential to creating a cost-effective hybrid wheat seed production system (N. Garst, 2017), thus making hybrid wheat a reality. Advanced breeding lines of hard red winter wheat from the Texas A&M University wheat breeding program were evaluated for female floral traits. The range of stigma size, stigma featheriness duration, and stigma exsertion were reported to be 1.48-4.12 (length, mm), 2-15 (days), and 1-5 (1-5 scale), respectively

(Tadlock, 2015). The range of the stigma exsertion and gape measured on a 1-5 scale was between 0.87–4 and 1–5 in a panel of hard red winter wheat in Texas A&M AgriLife wheat germplasm (Sade et al., 2022). This demonstrates that there is morphological variability in female floral traits in wheat germplasm and understanding this variation is crucial to enhance the development of female parental lines to be included in hybrid wheat development schemes.

Phenotypic selection can be supplemented by the implementation of novel technologies that are available to plant breeders who are attempting to transform wheat into a hybrid crop. Cutting-edge breeding techniques and genomic tools are advancing at an unprecedented rate and are becoming more widely available. Tools such as next-generation sequencing, genomic prediction, and high-throughput phenotyping can be used to rapidly develop hybrid breeding programs in self-pollinated crops (El Hanafi et al., 2021; Galán et al., 2020; Xu et al., 2021). Next Generation Sequencing (NGS) technology makes it possible to acquire sequence data quickly and affordably from multiple germplasm lines for use in genomics and related research. Whole genome sequencing has become widely accessible due to the steadily declining cost of DNA sequencing, encouraging the use of genome wide association studies (GWAS). Genome wide association studies are an approach used to identify associations between genome-wide sets of single-nucleotide polymorphisms (SNPs) and phenotypic traits of interest (Hee-Jong Koh et al., 2015). With a genome wide association study, it is possible to underline the genetic architecture and discover genetic variants of traits of interest. An additional benefit of GWAS is its capacity to locate quantitative trait locus (QTL) for

traits that cannot be characterized in a biparental population using interval mapping. Once the genetic basis of a trait is established, marker-assisted selection (MAS) could be used to improve breeding efficiency.

Several studies have focused on morphological differences and the genetic architecture of male floral traits, and almost all focused on anther extrusion (Adhikari et al., 2020; N. D. Garst, 2020; Langer et al., 2014; Muqaddasi et al., 2016). Elucidating the male floral traits is crucial since they are responsible for providing pollen. However, having receptive stigmas at the time of outcrossing is just as important as having sufficient and high-quality pollen. Therefore, it is essential to investigate the morphology and genetic basis of female floral structures to improve hybrid seed production. Only a small number of studies focus on the genetic architecture of female floral traits that are necessary to produce successful hybrids (El Hanafi et al., 2020, 2021; Tadlock, 2015). These studies are characterized for the evaluation of a limited number of female floral traits, the studies were replicated over a few years, and the amount of information that was provided was limited about female traits. Therefore, additional knowledge is required to comprehend female floral traits to improve hybrid seed production systems. This study aims to unravel the genetic architecture of key female floral traits for hybrid wheat seed production by i) phenotyping gape date, gape score, and chemical hybridizing agent (CHA) damage in a panel of winter wheat breeding lines from Texas and Nebraska public breeding programs, and ii) conducting a GWAS to decipher the genetic basis of the key female floral traits.

2. Materials and Methods

2.1 Plant Materials

This study evaluated a panel of 129 winter wheat breeding lines (Table 1) from the University of Nebraska-Lincoln and Texas A&M breeding programs. These lines were selected as parents for hybrid wheat based on agronomic characteristics and genetic diversity. Field experiments were conducted for eight years from 2015 to 2022 at the University of Nebraska-Lincoln Havelock Experiment Station Farm located in Lincoln, Nebraska, USA. Trials were conducted in a hybrid crossing block where male plots were placed around female (male-sterile) plots to make wheat hybrids (Figure 2). Female and male plots were 3 meters long, and 1.6 meters wide with five rows of 25 cm between them. The Chemical Hybridizing Agent (CHA) Croisor® 100 (active substance sintofen, 1-(4-chlorophenyl)-5-(2-methoxyethoxy)-4-oxo-1,4-dihydrocinnoline-3-carboxylic acid) provided by Asur Plant Breeding, Estrées-Saint-Denis, France, was applied to sterilize female plots at a rate of 10 liters per hectare when developing heads reached 12-20 mm in the main stem. Seeds were sown with a Hege 80 cone planter at a seeding rate of 125 kg per hectare for female plots and 52 kg per hectare for male plots. Variation in seed rates at planting was used to regulate tillering and thus ensure a large window of available pollen released by males at the time that females had receptible stigmas. The high seeding rates also contributed to obtaining a high homogeneity in the development of the plants in the female plots (more main stems and fewer late tillers). Consequently, sterilization in the plots was more homogeneous because, at the time the sterilizing agent was administered, the plants and tillers had similar stages of development, reducing

possible variations due to the effectiveness of the sterilizing agent in the female plots on late tillers. The number of female lines and replications used each year (Table 2) varied due to the different objectives of the hybrids produced from the crossing blocks each year. Wheat agronomic management practices for eastern Nebraska were implemented following the recommendations of the University of Nebraska-Lincoln extension program (UNL Extension, 2019). The design used from 2015 to 2021 was an incomplete randomized design. All 129 lines were included in a two-replications randomized complete block design in 2022, however, due to environmental conditions the trial was not successful. Therefore, the data was not analyzed.

2.2 Plant Phenotyping

The female plots were used to perform the phenotypic evaluation of key female floral traits for hybrid wheat seed production, gape date (Figure 3a), gape score (Figure 3b), and CHA damage (Figure 3c). The gape date was measured as the day when the florets in the whole plot reached maximum opening (gaping), counted in Julian days. The gape score was a measure used to quantify visually how widely open the florets were at the gape date in the whole plot. It was rated using a scale from 1 to 9, where 1 denoted the least gaping, and 9 denoted widely open florets. The phytotoxicity effects of the CHA, burned, yellow-red colored florets, and wrinkled, discolored, and damaged glumes and awns were quantified as CHA damage a few days after flowering stage, when the plants started to show symptoms of phytotoxicity in the whole female plot approximately at

Zadoks 70 stage (Conley, 2018). A scale from 1 to 9 was utilized, where 1 denoted no visible head damage and 9 denoted heads significant injuries (Stoll., 2019).

2.3 Statistical Analysis

Statistical analysis of phenotypic data was conducted by using a two steps approach. In the first step, phenotypic data from 2015 to 2021 was analyzed in a year-by-year basis and the Best Linear Unbiased Estimators (BLUE) per each genotype were estimated by fitting the linear mixed model:

$$\text{Model (1) } y_{ij} = \mu + G_i + R_j + e_{ij}$$

Where y_{ij} represent the phenotypic observation (gape date, gape score, or CHA damage) of the i^{th} genotype and the j^{th} replication, μ is the intercept term, G_i is the effect of the i^{th} genotype, R_j is the effect of the j^{th} replication e_{ij} is the residual term. The genotype term was treated as fixed and the replication term as random.

In the second step, the BLUE values from the first step were used to perform statistical analysis across environments. The Best Linear Unbiased Predictor estimates (BLUP) per each genotype were estimated by fitting the mixed linear model:

$$\text{Model (2) } y_{ij} = \mu + G_i + E_j + (GxE)_{ij} + e_{ij}$$

Where y_{ijk} represent the BLUE value of the phenotypic observation (gape date, gape score, or CHA damage) of the i^{th} genotype in the j^{th} environment, μ is the intercept term, G_i is the effect of the i^{th} genotype, E_j is the effect of the j^{th} environment, $(GxE)_{ij}$ is the effect of the interaction between the i^{th} genotype and the j^{th} environment, and e_{ij} is the residual term. Years were treated as environments as all field trials were carried out

at the same location (Havelock Experiment Station Farm, Lincoln, Nebraska, USA). Environment and the intercept terms were considered fixed and genotype and genotype by environment interaction effects were considered random. All the estimations were performed using the *lmer* function in the *Lme4* R package (Bates et al., 2015).

2.4 Phenotypic-Based broad-sense heritability

Broad-sense heritability (H^2) using the phenotypic data of each measured trait was calculated using the variance components estimated in the model (2) by adjusting the formula implemented by Sade et al., (2022):

Formula (1) $$H^2 = \frac{\sigma_G^2}{\sigma_G^2 + \frac{\sigma^2_{(GxE)}}{y} + \frac{\sigma_e^2}{y*r}}$$

Where σ_G^2 represent the genotypic variance, $\sigma^2_{(GxE)}$ is the genotype by environment (year) interaction variance, σ_e^2 is the error variance, y is the number of environments (7 years), and r is the number of replications.

2.5 Genotypic data

A detailed description of the DNA extraction, sequencing, and SNP calling of the 129 lines used in this study can be found in Belamkar et al, (2018). Briefly, each line was genotyped by Genotyping-by-Sequencing (GBS) using the protocol established by Poland et al (2012) in the Wheat Genetics and Germplasm Improvement Laboratory (WGGIL) at Kansas State University. SNP calling was performed using the TASSEL GBS Pipeline (Glaubitz et al., 2014) through the high-performance computing core, the Holland Computing Center (https://hcc.unl.edu/) at the University of Nebraska. 352,712

SNP markers were identified for the 129 lines evaluated in this study. Quality control was performed to remove SNPs with minor allele frequency (MAF) <5%, and missing values with >20%. Lines with >20% of missing marker data were removed. The total number of SNPs was reduced to 44,240 and the number of lines to 123 to be used in GWAS. Quality control was performed in TASSEL 5.2.86 (Bradbury et al., 2007).

2.6 Genomic broad-sense heritability

A total of 44,240 SNP markers were available for analysis after applying quality control. These were converted to numerical format, and the missing values were imputed with the naïve method (using the mean) with the numerical impute function in TASSEL 5.2.86 (Bradbury et al., 2007). The genotype's scale was changed to 2,0,1 for homozygous major, homozygous minor, and heterozygous, respectively. Then, the matrix of SNPs markers X was standardized by columns using "scale" function in R. The genomic relationship matrix (GRM) was computed with the formula suggested by VanRaden., (2008):

$$G = \frac{XX'}{ncol(X)}$$

Where the entries of *G* describe the genomic similarities between pairs of individuals, $ncol(X)$ is the number of SNP markers. The *BGLR* function in the BGL*R* R package (Pérez & De Los Campos, 2014) was used to fit the GBLUP model:

$$y_i = \mu + G_i + e_{ij}$$

Where y_i represent the phenotypic response (BLUP values of the genotypes generated by model (2)) of the i^{th} genotype, G_i is the corresponding genomic value

(derived from the GRM), and e_{ij} is the residual term capturing the non-explained variability. Since BGLR is a Bayesian implementation, for the GIBS sampler a total of 12,000 number of iterations were considered with 2,000 for burn-in.

The broad-sense heritability (H^2) using the markers data was calculated using the variance components estimated in the GBLUP model by using the formula:

$$\text{Formula (2) } H^2 = \frac{\sigma_G^2}{\sigma_G^2+\sigma_e^2}$$

Where σ_G^2 represent the variance of the genotypes, and σ_e^2 is the variance of the residual term.

2.7 Genome-Wide Association Analysis

BLUP values of each female floral trait from formula (2) and 44,240 SNP markers from 123 female lines were used to perform GWAS. Initially, the "model selection" option in the Genome Association and Prediction Integrated Tool (GAPIT) V3.0. (Lipka et al., 2012) was used to define the model that better fits the analysis. The selection criteria to choose the model were the effectiveness to control for relatedness and population structure, and the reduced presence of spurious associations. Based on the graphic representation of the expected and observed -$\log_{10}$ P values described in Quantile-Quantile (QQ) plots, the Mixed Linear Model (MLM) (Yu et al., 2006) was used to execute GWAS because it was efficient in controlling for suspicious associations, relatedness and population structure by included kinship and population structure matrices as covariates. The principal components analysis (PCA) was performed in GAPIT. Three principal components (PC) were selected for GWAS based on the

variation explained by each PC in the Eigenvalues plot (Figure 8). The PC and kinship matrices were produced and incorporated as covariates by default in the MLM model to account for population structure and individual relatedness. QQ plots of each evaluated trait were made by plotting the $-\log_{10}$ observed and expected p-values of the SNPs to check the suitability of the model. A threshold of P-value <0.001 (LOD score ≥3.00) was used to identify significant marker-trait associations (MTA) for the female floral traits evaluated in this study (Chaurasia et al., 2020; Sheoran et al., 2022; Soleimani et al., 2022). The $-\log_{10}$ (p) values were presented in Manhattan plots generated using the *qqman* R package (Turner S., 2018).

2.8 Candidate genes

SnpEff (Cingolani et al., 2012) was used for functional gene annotations to identify potential candidate genes using significant SNPs detected in GWAS for the female floral traits evaluated in this study. The biological function of the gene in relation to the floral traits in wheat was searched in prior literature. The most promising candidate genes based on the location of the gene in the genome, and on the availability and relevance of the biological function to the trait were selected for further exploration.

3. Results

3.1 Phenotypic analysis and heritability

Data curation was performed after the phenotypic evaluation of key female floral traits gape date, gape score, and CHA damage was conducted. Data for the following

traits was not collected: Gape date 2015; Gape Score 2016 and 2017; and CHA damage 2016. We checked for outliers by generating box plots, we explored further any tentative outlier visualized in box plots, but the suspected data points were relevant for our study, and we decided to use all phenotypic data that was collected. Statistical analysis was performed and the resulting best linear unbiased predictors (BLUPs) from all year's data approximated a normal distribution for all female traits (Figure 4-6). Summary statistics were reported in Table 3. Gape date had a 6-day span from 141 to 147 with a mean of 144. We observed variation for gape score ranging from 6.7 to 9.6 with a mean value of 8.4. CHA damage ranged from 6.5 to 7.2 with a mean value of 6.8. Estimates of genotypic variance were significantly different than zero with a p-value of 0.001 for gape data and gape score, and a p-value of 0.05 for CHA damage in type 3 Anova test of random effects (Table 4). The estimated phenotypic broad-sense heritability for gape date, gape score, and CHA damage were 0.88, 0.73, and 0.36, and the genomic broad-sense heritability was 0.69, 0.56, and 0.50, respectively (Table 4).

3.2 Marker-Trait Associations for Female Floral Traits

Principal components analysis (PCA) showed that the three first principal components explained most of the variation with 4.11%, 3.81%, and 3.23% respectively (Figure 7). QQ plots showed a thin tail in the distribution of the $-\log_{10}$ p-values deviating from the straight line (Figure 12). Because the Bonferroni and FDR corrections were very conservative for our data, we used a threshold of significance of $P<0.001$ ($-\log_{10}(P) \geq 3.0$) to declare MTA, which is commonly used in previous studies with small sample size

(Chaurasia et al., 2020; Sheoran et al., 2022; Soleimani et al., 2022). GWAS detected in total 73 MTA for the three measured female floral traits. The MTAs were located on all chromosomes except for 1D, 3D, and 4D. Of the total, 20 MTA were detected for gape date, 33 MTA for gape score, and 20 MTA for CHA damage (Figure 9-11). The genomic regions with the most significantly associated markers based on the highest LOD values for gape date were chromosomes 4A (3), and 5A (3), for gape score chromosomes 1A (4), 5A (4), and 6D (4), and for CHA damage chromosome 7D (4). The most significant SNP detected among all female traits with a LOD of 4.55 (P-value 0.00003) was S6B_549563085 located on chromosome 6B for gape score. The phenotypic variance explained (PVE) by all detected SNPs ranges from 0-9.56% with a mean value of 1.17% (Table 5).

3.3 Candidate genes

Based on 200 kb window around the significant SNPs, we identified in total 111 candidate genes using *SnpEff* (Cingolani et al., 2012) for functional gene annotations. We found 34, 51, and 29 candidate genes for the gape date, gape score, and CHA damage respectively. TraesCS4B02G242300, TraesCS4B02G242400, and TraesCS7D02G039500 genes were identified for both gape score and CHA damage (Table 6). Of the total 111 candidate genes, 3 genes were the most promising based on their location in the missense variant genomic region (Table 7).

4. Discussion

4.1 Phenotypic evaluation of female traits

This work describes one of the first attempts to use a diverse set of germplasm to identify genomic markers associated with female floral traits in wheat. Our efforts to characterize female traits in CHA sterilized wheat germplasm from TX and NE continues to the work of Stoll (2019) and Sade et al. (2022). The genetic variance for all female floral traits in all evaluated lines used in this study was significantly different than zero (Table 4). This result shows that there is significant variation among the lines for all female floral traits, indicating that selection for these traits could improve outcrossing ability and therefore reduce the cost of hybrid wheat seed production systems (Langer et al., 2014).

Gape date play a significant role in the synchronization of flowering times in parental lines (N. Garst, 2017; Stoll, 2019). The gape date in the female parent should occur on the same day or a few days before the peak of pollen shed occurs in the male parent. A very late female line would be extremely hard to synchronize with an early male line. The wheat lines in this study exhibited a wide range for gape dates (141-147) suggesting that parent selection should be considered carefully to optimize overlap between pollen shed and floral gaping.

Wider floral gape is critical for allowing pollen into the floret for successful hybridization and improved seed set in hybrid wheat (Tadlock, 2015). The greater the gape, the better chance pollen can infiltrate the floret to pollinate the stigma. Lines with >6 of gape score are considered superior for outcrossing ability. The range of gape scores

(6.7-9.6) indicate that most of the lines have superior ability to cross pollinated in the evaluated panel. In a similar study that evaluated the angle of glume separation using a 1-5 scale in a panel of hard red winter wheat breeding lines at Texas A&M University, a wide range of variation (2-5) for this trait was found (Tadlock, 2015). Wide gape occurs frequently in both Texas and Nebraska lines suggesting that the trait may be commonly present in wheat germplasm. The high gape scores suggest that certain lines have a superior ability for outcrossing, and these lines are recommended for use to develop female heterotic pools for wheat.

The damage caused by CHAs has an enormous impact on seed development causing decreased seed formation, which reduces the efficiency of hybrid seed production systems. Most of the lines had moderate CHA damage (<7) in the evaluated panel which could cause reduced number of seeds per spike and smaller seeds with low test weight. Our results confirmed a preliminary analysis made by Stoll (2019) that reported moderate tolerance to CHA in a panel of winter wheat lines tested in Nebraska. These results show that some of the evaluated lines tolerate the hybridizing agent well, and this system could be an economically effective method to perform initial evaluation of lines as potential hybrid parents before integrating them into more sophisticated systems like cytoplasmic male sterility to create wheat hybrids. However, it is important to evaluate different systems that do not rely on CHA for testing or to develop CHAs that cause less damage to expand the germplasm available for hybrid breeding programs (Pickett, 1993).

Broad-sense heritability was estimated using phenotypic and genotypic data. The estimates of phenotypic broad sense heritability ranged from high (0.88 and 0.73) in the

case of gape date and gape score to low (0.36) in the case of CHA damage. Due to the availability of high-density marker data, genomic broad-sense heritability was estimated to better understand the effects of genetics on female floral traits (de los Campos et al., 2015). Genomic heritability was moderate (0.50-0.69) for all traits which is in accordance with the results reported by Stoll (2019). This demonstrates that female floral traits are heritable, and it is possible to breed for these traits to develop hybrid wheat.

4.2 Marker-Trait Associations and candidate genes

This study is the first to apply GWAS to female floral traits in hard winter wheat. To the best of our knowledge, the genetic mechanisms that lead to gape date and CHA damage for hybrid wheat have not yet been reported. Principal components analysis (PCA) was performed to investigate the population structure of the evaluated panel of breeding lines. The first three principal components explained limited genetic variation and identified two clusters corresponding to Nebraska and Texas public wheat breeding programs (Figure 7). However, some overlap between the clusters is present. The pedigrees of these overlapping lines confirmed that they had parents in common or were exchanged between the Nebraska and Texas breeding programs (Table 1). QQ plots revealed a good fit of the MLM for GWAS to detect naïve MTA for female floral traits, and to effectively account for population structure and kinship relatedness between lines (Figure 12).

We detected 20 MTA for gape date, 33 for gape score, and 20 for CHA damage distributed in all chromosomes except for 1D, 3D, and 4D. The D genome is the least

diverse compared with A and B genomes because it was the most recent hybridization event that occurred to create in *T. aestivum*. Therefore, the number of polymorphisms present is less compared with the other two genomes. Also, the polymorphisms present in the D genome are not well distributed among the chromosomes (Mirzaghaderi & Mason, 2019). S4B_501710513 and S7D_20057921 were identified as common SNPs for gape score and CHA damage, suggesting that these SNPs could control multiple female traits (pleiotropy). Previously, El Hanafi et al., (2021) studied the openness of the flower as the angle of separation between the glumes of the first two florets of a spikelet. They identified 11 MTAs associated with this trait but none of them were common with the ones detected in our study. The MTAs with the highest PVE for gape date in our study was S4A_731097948 located on chromosome 4A, for gape score was S2B_742526639 located on chromosome 2B, and for CHA damage was S7B_622525086 located on chromosome 7B. The MTAs with the highest PVE for all traits were located in different genomic regions. We found an lower percentage of PVE for all detected SNPs for gape score compared with the PVE reported by El Hanafi et al., (2021) for the openness of the flower in a panel of spring wheat. The divergence in the PVE results in both studies could be due to the differences in experimental design, sample size, number of years evaluated, location, and type of wheat used (spring and winter wheat). PVE varies and is very specific from experiment to experiment. Minor allele frequency of the SNP S2D_10870835 and S2D_10870844 linked with candidate gene TraesCS2D02G024600 for gape score was 0.38. Minor allele frequency of SNP S3A_705230542 linked with candidate gene TraesCS3A02G474000 for gape score was 0.07. The MAF of SNP

S3B_815282615 linked with candidate gene TraesCS3B02G590100 for CHA damage was 0.11. In each case, the minor allele is beneficial in the evaluated population and selecting for this allele will be beneficial for breeding for gape score and CHA damage.

Candidate genes TraesCS2D02G024600 and TraesCS3A02G474000 were associated with gape score, and TraesCS3B02G590100 gene was associated with CHA damage. The most promising candidate genes were missense variants with moderate impact meaning loss of function of these candidate genes could be contributing to the trait of interest. These candidate genes code for unknown proteins and the specific biological function of the candidate genes need to be explored. Gene validation approaches such as gene editing and transgenics could be deployed to dissect these candidate genes for gene function similar to prior reports in wheat for stem rust, and seed morphology (Saintenac et al., 2013; Wang et al., 2018; Zhang et al., 2021). If they are successfully validated, these candidate genes can be further used via marker assisted selection for improving female floral traits in hybrid wheat breeding programs.

5. Conclusions

The observed large phenotypic variation for the female floral traits evaluated in the present study suggests the potential for selecting superior female parents for hybrid wheat. GWAS demonstrated that female floral traits are complex and controlled by multiple genetic loci (quantitative traits). Increasing the population size in GWAS would be beneficial to further enhance the power to detect more significant marker-trait associations. Identified candidate genes need to be further validated for their direct

implication in improving female floral traits in hybrid wheat breeding programs. The understanding of the genetics of the female floral traits, and the identified MTAs and candidate genes in this study could serve as a foundation for future research in designing female floral traits to improve cross-pollination for efficient hybrid wheat seed production.

6. References

Adhikari, A., Basnet, B. R., Crossa, J., Dreisigacker, S., Camarillo, F., Bhati, P. K., Jarquin, D., Manes, Y., & Ibrahim, A. M. H. (2020). Genome-Wide Association Mapping and Genomic Prediction of Anther Extrusion in CIMMYT Hybrid Wheat Breeding Program via Modeling Pedigree, Genomic Relationship, and Interaction With the Environment. *Frontiers in Genetics*, *11*(December), 1–13. https://doi.org/10.3389/fgene.2020.586687

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Belamkar, V., Guttieri, M. J., Hussain, W., Jarquín, D., El-basyoni, I., Poland, J., Lorenz, A. J., & Baenziger, P. S. (2018). Genomic selection in preliminary yield trials in a winter wheat breeding program. *G3: Genes, Genomes, Genetics*, *8*(8), 2735–2747. https://doi.org/10.1534/g3.118.200415

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

CGIAR. (2016). Wheat, Agri-Food Systems 2017-2022. https://cgspace.cgiar.org/bitstream/handle/10947/4419/2. WHEAT - CRP and FP Narratives Proposal 2017-2022.pdf?sequence=1&isAllowed=y

Chaurasia, S., Singh, A. K., Songachan, L. S., Sharma, A. D., Bhardwaj, R., & Singh, K. (2020). Multi-locus genome-wide association studies reveal novel genomic regions associated with vegetative stage salt tolerance in bread wheat (Triticum aestivum L.). *Genomics*, *112*(6), 4608–4621. https://doi.org/10.1016/j.ygeno.2020.08.006

Chu, S. H., Jiang, W., Lee, J., Chin, J. H., & Koh, H. J. (2012). QTL analyses of heterosis for grain yield and yield-related traits in indica-japonica crosses of rice (Oryza sativa L.). *Genes and Genomics*, *34*(4), 367–377. https://doi.org/10.1007/s13258-011-0223-x

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly*, *6*(2), 80–92. https://doi.org/10.4161/fly.19695

Conley, S. P. (2018). A Visual Guide to Winter Wheat Development and Growth Staging. *WCM Newsletter*. https://ipcm.wisc.edu/blog/2018/04/a-visual-guide-to-winter-wheat-development-and-growth-staging/%0Ahttp://coolbean.info/wp-content/uploads/sites/3/2018/04/2018_WheatGrowthStages_FINAL.pdf

de los Campos, G., Sorensen, D., & Gianola, D. (2015). Genomic Heritability: What Is It? *PLoS Genetics*, *11*(5), 1–21. https://doi.org/10.1371/journal.pgen.1005048

De Vries, A. P. (1971). Flowering biology of wheat, particularly in view of hybrid seed production — A review. *Euphytica*, *20*(2), 152–170. https://doi.org/10.1007/BF00056076

Easterly, A. C., Garst, N., Belamkar, V., Ibrahim, A. M. H., Rudd, J. C., Sarazin, J. B., & Baenziger, P. S. (2020). Evaluation of hybrid wheat yield in Nebraska. *Crop Science*, *60*(3), 1210–1222. https://doi.org/10.1002/csc2.20019

El Hanafi, S., Bendaou, N., Kehel, Z., Sanchez-Garcia, M., & Tadesse, W. (2020). Phenotypic evaluation of elite spring bread wheat genotypes for hybrid potential traits. *Euphytica*, *216*(11), 1–16. https://doi.org/10.1007/s10681-020-02695-7

El Hanafi, S., Cherkaoui, S., Kehel, Z., Al-Abdallat, A., & Tadesse, W. (2021). Genome-Wide Association and Prediction of Male and Female Floral Hybrid Potential Traits in Elite Spring Bread Wheat Genotypes. *Plants*, *10*(5), 895. https://doi.org/10.3390/plants10050895

FAO. (2019). Global agriculture towards 2050. High level experts forum. Rome. https://www.fao.org/fileadmin/templates/wsfs/docs/Issues_papers/HLEF2050_Global_Agriculture.pdf

FAOSTAT. (2019). Production domain. Crops and livestock products. http://faostat.fao.org

FAO. (2022). Food Outlook. Biannual Report on Global Food Markets. Rome. https://doi.org/10.4060/cb9427en

Galán, R. J., Bernal-Vasquez, A. M., Jebsen, C., Piepho, H. P., Thorwarth, P., Steffan, P., Gordillo, A., & Miedaner, T. (2020). Integration of genotypic, hyperspectral, and phenotypic data to improve biomass yield prediction in hybrid rye. *Theoretical and Applied Genetics*, *133*(11), 3001–3015. https://doi.org/10.1007/s00122-020-03651-8

Garst, N. (2017). *Assessing anther extrusion and its effect on US hard winter wheat (Triticum aestivum L.) hybrid seed production*. 1–37. http://digitalcommons.unl.edu/agronhortdiss%0Ahttp://digitalcommons.unl.edu/agronhortdiss/121

Garst, N. D. (2020). *EVALUATION AND GENETIC ARCHITECTURE OF POLLINATION TRAITS FOR HYBRID WHEAT SEED PRODUCTION IN UNITED STATES GREAT PLAINS GERMPLASM by A DISSSERTATION Presented to the Faculty of The Graduate College at the University of Nebraska In Partial Fulfillment of*. https://www.proquest.com/docview/2437839157?parentSessionId=ymVtteKpkmVEM6NhmtkKiE0coRMk3IQbCMhODGpm8S0%3D

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler, E. S. (2014). TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline. *PLoS ONE*, *9*(2), e90346. https://doi.org/10.1371/journal.pone.0090346

Hee-Jong Koh, Suk-Yoon Kwon, & Michael Thomson. (2015). Current Technologies in Plant Molecular Breeding. In H.-J. Koh, S.-Y. Kwon, & M. Thomson (Eds.), *Current Technologies in Plant Molecular Breeding*. Springer Netherlands. https://doi.org/10.1007/978-94-017-9996-6

Kempe, K., & Gils, M. (2011). Pollination control technologies for hybrid breeding. *Molecular Breeding*, *27*(4), 417–437. https://doi.org/10.1007/s11032-011-9555-0

Koemel, J. E., Guenzi, A. C., Carver, B. F., Payton, M. E., Morgan, G. H., & Smith, E. L. (2004). Hybrid and Pureline Hard Winter Wheat Yield and Stability. *Crop Science*, *44*(1), 107–113. https://doi.org/10.2135/cropsci2004.1070

Langer, S. M., Longin, C. F. H., & Würschum, T. (2014). Phenotypic evaluation of floral and flowering traits with relevance for hybrid breeding in wheat (Triticum aestivum L.). *Plant Breeding*, *133*(4), 433–441. https://doi.org/10.1111/pbr.12192

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

Longin, C. F. H., Gowda, M., Mühleisen, J., Ebmeyer, E., Kazman, E., Schachschneider, R., Schacht, J., Kirchhoff, M., Zhao, Y., & Reif, J. C. (2013). Hybrid wheat: Quantitative genetic parameters and consequences for the design of breeding programs. *Theoretical and Applied Genetics*, *126*(11), 2791–2801. https://doi.org/10.1007/s00122-013-2172-z

Longin, C. F. H., Mühleisen, J., Maurer, H. P., Zhang, H., Gowda, M., & Reif, J. C. (2012). Hybrid breeding in autogamous cereals. *Theoretical and Applied Genetics*, *125*(6), 1087–1096. https://doi.org/10.1007/s00122-012-1967-7

Mirzaghaderi, G., & Mason, A. S. (2019). Broadening the bread wheat D genome. *Theoretical and Applied Genetics*, *132*(5), 1295–1307. https://doi.org/10.1007/s00122-019-03299-z

Mühleisen, J., Piepho, H. P., Maurer, H. P., Longin, C. F. H., & Reif, J. C. (2014). Yield stability of hybrids versus lines in wheat, barley, and triticale. *Theoretical and Applied Genetics*, *127*(2), 309–316. https://doi.org/10.1007/s00122-013-2219-1

Muqaddasi, Q. H., Lohwasser, U., Nagel, M., Börner, A., Pillen, K., & Röder, M. S. (2016). Genome-Wide Association Mapping of Anther Extrusion in Hexaploid Spring Wheat. *PLOS ONE*, *11*(5), e0155494. https://doi.org/10.1371/journal.pone.0155494

Pérez, P., & De Los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics*, *198*(2), 483–495. https://doi.org/10.1534/genetics.114.164442

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE*, *7*(2). https://doi.org/10.1371/journal.pone.0032253

Sade, B., Ibrahim, A. M. H., Subramanian, N., Rudd, J., Liu, S., Opena, G., & Baenziger, S. (2022). Assessment of floral characteristics for hybrid wheat (Triticum aestivum L.) production in Texas. *Agrosystems, Geosciences & Environment*, *5*(1), 1–11. https://doi.org/10.1002/agg2.20228

Saintenac, C., Zhang, W., Salcedo, A., Rouse, M. N., Trick, H. N., Akhunov, E., & Dubcovsky, J. (2013). Identification of wheat gene Sr35 that confers resistance to Ug99 stem rust race group. *Science*, *341*(6147), 783–786. https://doi.org/10.1126/science.1239022

Sheoran, S., Jaiswal, S., Raghav, N., Sharma, R., Sabhyata, Gaur, A., Jaisri, J., Tandon, G., Singh, S., Sharma, P., Singh, R., Iquebal, M. A., Angadi, U. B., Gupta, A., Singh, G., Singh, G. P., Rai, A., Kumar, D., & Tiwari, R. (2022). Genome-Wide Association Study and Post-genome-Wide Association Study Analysis for Spike

Fertility and Yield Related Traits in Bread Wheat. *Frontiers in Plant Science*, *12*(February), 1–17. https://doi.org/10.3389/fpls.2021.820761

Singh, S. K., Chatrath, R., & Mishra, B. (2010). Perspective of hybrid wheat research: A review. *Indian Journal of Agricultural Sciences*, *80*(12), 1013–1027.

Soleimani, B., Lehnert, H., Babben, S., Keilwagen, J., Koch, M., Arana-Ceballos, F. A., Chesnokov, Y., Pshenichnikova, T., Schondelmaier, J., Ordon, F., Börner, A., & Perovic, D. (2022). Genome wide association study of frost tolerance in wheat. *Scientific Reports*, *12*(1), 1–11. https://doi.org/10.1038/s41598-022-08706-y

Stoll, H. K. (2019). *UNRAVELING THE INFLUENCE OF FEMALE FLOWERING CHARACTERISTICS AND*.

Tadlock, J. K. (2015). *Floral characteristics and hybrid performance of potential candidates of a hybrid wheat (Triticum aestivum L.) program in Texas*. *December*. https://oaktrust.library.tamu.edu/handle/1969.1/156541

United Nations Department of Economic and Social Affairs, Population Division (2022). World Population Prospects 2022: Summary of Results. UN DESA/POP/2022/TR/ NO. 3

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91*(11), 4414–4423. https://doi.org/10.3168/jds.2007-0980

Wang, W., Pan, Q., He, F., Akhunova, A., Chao, S., Trick, H., & Akhunov, E. (2018). Transgenerational CRISPR-Cas9 Activity Facilitates Multiplex Gene Editing in Allopolyploid Wheat. *The CRISPR Journal*, *1*(1), 65–74. https://doi.org/10.1089/crispr.2017.0010

Whitford, R., Fleury, D., Reif, J. C., Garcia, M., Okada, T., Korzun, V., & Langridge, P. (2013). Hybrid breeding in wheat: Technologies to improve hybrid wheat seed production. *Journal of Experimental Botany*, *64*(18), 5411–5428. https://doi.org/10.1093/jxb/ert333

Xu, Y., Zhao, Y., Wang, X., Ma, Y., Li, P., Yang, Z., Zhang, X., Xu, C., & Xu, S. (2021). Incorporation of parental phenotypic data into multi-omic models improves prediction of yield-related traits in hybrid rice. *Plant Biotechnology Journal*, *19*(2), 261–272. https://doi.org/10.1111/pbi.13458

Yu, J., Pressoir, G., Briggs, W. H., Bi, I. V., Yamasaki, M., Doebley, J. F., McMullen, M. D., Gaut, B. S., Nielsen, D. M., Holland, J. B., Kresovich, S., & Buckler, E. S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, *38*(2), 203–208. https://doi.org/10.1038/ng1702

Zhang, J., Hewitt, T. C., Boshoff, W. H. P., Dundas, I., Upadhyaya, N., Li, J., Patpour, M., Chandramohan, S., Pretorius, Z. A., Hovmøller, M., Schnippenkoetter, W., Park, R. F., Mago, R., Periyannan, S., Bhatt, D., Hoxha, S., Chakraborty, S., Luo, M., Dodds, P., … Lagudah, E. S. (2021). A recombined Sr26 and Sr61 disease resistance gene stack in wheat encodes unrelated NLR genes. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-23738-0

Table 1: List of the 129 breeding lines used to evaluate female floral traits for hybrid wheat seed production

Line	Pedigree	Source
FREEMAN	ABI86*3414/Jagger//Karl 92 (KS92-946-B-15-1)/3/ALLIANCE	Nebraska
GOODSTREAK	SD3055/KS88H164//NE89646	Nebraska
HARRY	NE90614 (=BRL/4/PKR*4/AGT//BEL.198/LCR/3/NWT/BRL) (=BRL/4/PKR*4/AGT//BEL.198/LCR/3/NWT/BRL)/NE87612 (=NWT//WRR*5/AGT/3/NE69441)	Nebraska
LCH13NEDH_11_24	NE06469/Pronghorn	Nebraska
NE05496	KS87H325/RIO BLANCO (KS95HW62-6)//HALLAM	Nebraska
NE07486_2	Lakin/Ok102 F3	Nebraska
NE07531	HBA142A/HBZ623A//ALE (HBK0630-4-5)/3/(NE98574) CO850267/RAWHIDE/4/HALLAM	Nebraska
NE09517_1	W96x1080-21(Jagger/Thunderbolt)/Jagalene	Nebraska
NE09521	OK96717-99-6756=(Abilene/2180//Chisholm)/NI01824=(INTENSIVNAJA/NE92458 (=OK83201/REDLAND)//VBF0168)//NE005564 = (T81/NE91635 (=NE82761/NE82599))	Nebraska
NE10478	NI03418=(W91-248/NE95544 (=MCVEY 78015/NE88521)//THUNDERBIRD)/NE01604	Nebraska
NE10683	NE01481=(OK83201/REDLAND//IKE)/Harry=(NE90614 (=BRL/4/PKR*4/AGT//BEL.198/LCR/3/NWT/BRL)/NE87612 (=NWT//WRR*5/AGT/3/NE69441))	Nebraska
NE12589	CO97547-7/OVERLAND//CAMELOT	Nebraska
NE13434	TX00D1390/NE02495//McGill	Nebraska
NE13515	HV9W00-B267/NI04421//NI04427	Nebraska
NE13604	SD00258-1/McGill	Nebraska
NE13625	W04-417/CAMELOT//Goodstreak	Nebraska
NE13672	MO980829/NE01604//NE03490	Nebraska
NE14419	Cassiopea/NW03681//NE03490	Nebraska
NE14421	NE05426/Overland	Nebraska
NE14434	SD98W175-1/NW03666//Freeman	Nebraska
NE14448	NI06737/1(ND2928/Wesley//Wesley)F3/Wesley F3	Nebraska
NE14494	OK06822W/HV9W96-1383W//NW03681	Nebraska
NE14531	NI06737/1(ND2928/Wesley//Wesley)F3/Wesley F3	Nebraska
NE14538	SD98W175-1/NW03666//Freeman	Nebraska
NE14606	KS04HW101-3/NW03670//NW06655	Nebraska
NE14663	NI06737/1(ND2928/Wesley//Wesley)F3/Wesley F3	Nebraska
NE14691	SD05W138/NE01604	Nebraska
NE14696	NE05537/Overland	Nebraska
NE15405	TX99A0153-1/Santa Fe//NE04424	Nebraska

Line	Pedigree	Source
NE15406	BIG RED/MACE//NI04420	Nebraska
NE15410	Wesley * Madsen /TAM111	Nebraska
NE15417	Wahoo * Dn4/NE01643	Nebraska
NE15434	NW03681/NE97465 = Goodstreak//NW07534	Nebraska
NE15440	NE05569/NW07534	Nebraska
NE15445	Wesley * Madsen /BILLBROWN	Nebraska
NE15468	KS05HW15-2/NW03681	Nebraska
NE15475	NW03666/KS05HW15-2	Nebraska
NE15545	NE07569/NE04424	Nebraska
NE15571	NE05569/NW07534	Nebraska
NE15605	NI04436/Agripro Art//NE04490	Nebraska
NE15624	NE05537/KS05HW15-2	Nebraska
NE15689	HV9W02-942R/NI04421	Nebraska
NE16443	NE96644(=ODESSKAYA P./CODY)//PAVON/*3SCOUT66/3/Wahoo (sib)	Nebraska
NI10718W	SD97W609=(Abilene/Karl)/NW98S097=(WA691213-27/N86L177//AP-WI89-163)	Nebraska
NI12702W	N03Y2014/NW03681//NuHills 10005	Nebraska
NI13706	NI02425/HV9W99-558//Robidoux	Nebraska
NI14733	MV-Regiment/NE04550	Nebraska
NW07534	KS920709-B-5-2/NW98S061	Nebraska
NW13493	SD98W175-1/NW03666	Nebraska
NW13669	SD98W175-1-14/NW03666	Nebraska
NW15404	KS05HW15-2/NW03681	Nebraska
NW15443	OR 2060108/NW03681//NW03666	Nebraska
NW15564	KS05HW15-2/NW03681	Nebraska
NW15573	KS05HW15-2/NW03681	Nebraska
NW15677	KS05HW15-2/NW03681	Nebraska
OVERLAND	Millennium sib//(ND8974) Seward/Archer	Nebraska
PANHANDLE	BRIGANTINA/2*ARAPAHOE (NE97426)//NE98574	Nebraska
PSB13NEDH_14_71W	NW03681 / SD07W084	Nebraska
PSB13NEDH_14_83W	NW03681 / SD07W084	Nebraska
PSB13NEDH_15_58W	NW03681 / SD07W084	Nebraska
ROBIDOUX	Odesskaya P / Cody // Pavon 76 /3* Scout 66 (NE96644)/3/ Wahoo sib	Nebraska
RUTH	OK98697/Jagalene//Camelot	Nebraska
SETTLER_CL	Wesley sib//Millennium sib/Above sib	Nebraska
SIEGE	NI04420/NE00403	Nebraska

Line	Pedigree	Source
STURDY2K	Selection from Sturdy (Citr 13684=Sinvalocho / Wichita // Hope / Cheyenne /3/2* Wichita /4/ Seu Seun 27) released in 1966	Nebraska
TAM111	TAM-107//TX78V3630/CTK78/3/TX87V1233	Texas
TAM112	TAM 112 (=TX98V9628=U1254-7-9-2-1/TXGH10440)	Texas
TAM113	TAM113 (=TX02A0252=TX90V6313//TX94V3724(TAM-200 BC41254-1-8-1-1/TX86V1405	Texas
TAM114	TX07A001505=T107//TX98V3620/Ctk78/3/TX87V1233/4/N87V106//TX86V1540/T200	Texas
TAM204	TX06V7266=TAM 112/TX01M5009	Texas
TAM304	TX01D3232 =TX92U3060/TX91D6564 (=X95U104-P66)	Texas
TAM305	TAM 305 (=TX06A001263=TX97V3006/TX98V62390)	Texas
TAM401	Mason/Jagger (=TX03M1096)	Texas
TX09D1172	TAM303/TAM112	Texas
TX10D2063	OK99610/TX00V1131//TX02D5868	Texas
TX10D2230	NW01L2019/TX96D1073//TX01D3215	Texas
TX10D2363	OK99610/TAM 109//TAM 304	Texas
TX11A001295	TAM 112/TX02U2508	Texas
TX11D3008	TX03M1004/TX02V7930	Texas
TX11D3026	TX01V5425/KS03HW155-2//TX03M1004	Texas
TX11D3049	TX96D1073/KSS9011-1-45 IP76//KS00F5-14-7	Texas
TX11D3112	TX98V9628/TX02U2508	Texas
TX11D3129	WBLL 1*2/TUKURU//OK BULLET	Texas
TX12A001041	RonL/TX04V072075	Texas
TX12A001638	TAM 112/TX02U2508	Texas
TX12M4004	KS980478-3-~5/FULLER	Texas
TX12M4063	AP04TW9819/O3A-B3//KS980512-11-22	Texas
TX12M4065	AP04TW1318/KS980512-11-9//KS06O3A~49	Texas
TX12M4068	AP04TW1318/KS980512-11-9//KS06O3A~49	Texas
TX13M5604	O3A-89-2/KS990159-3-7//Aspen	Texas
TX13M5625	O3A-B7/HV9W96-1270R-1//KS980512-11-24	Texas
TX13M5652	AP04T9029/W04-417//KS010525-1-1	Texas
TX14A001035	OK03522 (=N566/OK94P597)/TX03A0563 (=X96V107/OGALLALA)	Texas
TX14A001112	Duster (=OK93P656H3299-2C04=WO405D/HGF112//W7469C/HCF012)/TX01V5134WC-2 (=TAM-200/JAGGER)	Texas
TX14A001336	X07A457S [04AKF3B-106 (=FALCIN/AE.SQ (312)// RB/KSWGRC10 (OK98G508W)/3/GK ARON/AGSECO 7846//2180 (OK99711))/TX05A001838 (=X920709-B-5-2-2/X940786-6-7)]/Endurance (=OK94P549-11=HBY756A/Siouxland//2180)	Texas
TX14M7013	TX04A001246 (=TX95V4339/TX94VT938-6)/Jackpot (=AP04T8211=W98-232/KS96WGRC38)	Texas

Line	Pedigree	Source
TX14M7034	TX01V5134RC-3 (=TAM-200/JAGGER)/TX01V5134WC-2 (=TAM-200/JAGGER)	Texas
TX14M7051	TX04M410164 (=MIT/TX93V5722//W95-301)/TX05A001846 (=TX99V2437/Ventor)	Texas
TX14M7054	TX04M410164 (=MIT/TX93V5722//W95-301)/TX05A001846 (=TX99V2437/Ventor)	Texas
TX14M7057	TX04M410164 (=MIT/TX93V5722//W95-301)/TX05A001846 (=TX99V2437/Ventor)	Texas
TX14M7061	TX02A0252 (=TX90V6313//TX94V3724(TAM-200 BC41254-1-8-1-1/TX86V1405)/TX03A0148 (=TX89A7137/TIPACNA)	Texas
TX14M7088	TX02A0252 (=TX90V6313//TX94V3724(TAM-200 BC41254-1-8-1-1/TX86V1405)/OK02522W (=KS96WGRC39/Jagger)	Texas
TX14M7153	TX03A0563 (=X96V107/OGALLALA)/TX01V5134WC-2 (=TAM-200/JAGGER)	Texas
TX14M7174	TAM 401 (=TX03M1096=MASON/JAGGER)/TX01V5134WC-2 (=TAM-200/JAGGER)	Texas
TX14M7177	TAM 401 (=TX03M1096=MASON/JAGGER)/TX06A001236 (=OGALLALA/KS94U275)	Texas
TX14M7290	TAM 304 (=TX01D3232=TX92U3060/TX91D6564 (=X95U104-P66))/TX05A001419 (=HBG0358/4/T107//TX78V3620/Ctk78/3/TX87V1233)	Texas
TX14M7306	OK03522 (=N566/OK94P597)/TX04A001246 (=TX95V4339/TX94VT938-6)	Texas
TX14M7320	OK03522 (=N566/OK94P597)/TX03A0563 (=X96V107/OGALLALA)	Texas
TX14M7327	BC98331-03$-2W (=KS920709B-5-2/2137//KS920709B5-2)/TX03A0148 (=TX89A7137/TIPACNA)	Texas
TX14M7333	BC98331-03$-2W (=KS920709B-5-2/2137//KS920709B5-2)/TX03A0148 (=TX89A7137/TIPACNA)	Texas
TX14M7334	BC98331-03$-2W (=KS920709B-5-2/2137//KS920709B5-2)/TX03A0148 (=TX89A7137/TIPACNA)	Texas
TX14M7347	KS980512-2-2 (=T67/X84W063-9-45//K92/3/SNF/4/X86509-1-1/X84W063-9-39-2//K92)/TX04A001246 (=TX95V4339/TX94VT938-6)	Texas
TX14M7373	TX05V071029 (=L90239B19-2-4/JAGGER//THUNDERBOLT)/Deliver (=OK98690=OK91724/Karl)	Texas
TX14M7384	TX06A001263 (=TX97V3006/TX98V6239)/Duster (=OK93P656H3299-2C04=WO405D/HGF112//W7469C/HCF012)	Texas
TX14M7391	TX06A001263 (=TX97V3006/TX98V6239)/TAM 303Resel (=TX98D1170-05AHR15819=TX89D1253*2/TTCC404 (=WX93D208-9-1-2))	Texas
TX14M7475	X07A495S [Cas03OSUF3-258 (=VORONA//PRL/VEE#6/3/HBY756A/SXL//2180 (OK94P549-6611))/TX02A0252 (=TX90V6313//TX94V3724(TAM-200 BC41254-1-8-1-1/TX86V1405)]/OK02522W (=KS96WGRC39/Jagger)	Texas
TX14M7500	X07A578S [STEELE-ND (=PARSHALL/ND706)/TX04A001268 (=TX96V2627/TX94D7091)]/TX04A001246 (=TX95V4339/TX94VT938-6)	Texas

Line	Pedigree	Source
TX14M7549	X07A615S [Wheatear (=Sr2+Sr25)/TAM 112 (=TX98V9628=U1254-7-9-2-1/TXGH10440)]/TX01V5134WC-2 (=TAM-200/JAGGER)	Texas
TX14M7607	Art (=98x0338-13=Jagger/W94-244-132)/TAM 401 (=TX03M1096=MASON/JAGGER)	Texas
TX14M7626	X07A543S [U5109-106-1-4-1-3m (=Jagger*2/Kakatsi)/TAM 304 (=TX01D3232 =TX92U3060/TX91D6564 (=X95U104-P66))]/TX05A001639 (=X96V079/KS84W063-9-39)	Texas
TX14M7645	Duster/TX03A0148	Texas
TX15M8018	TAM 203 (=TX01V5314=TX89V4132/704 L I-2221)/TX06A001186W (=L92283C64-1/JAGGER//OGALLALA)	Texas
TX15M8023	TAM 203/Duster	Texas
TX15M8206	TX05A001188 (=T107//TX98V3620/Ctk78/3/TX87V1233/4/N87V106//TX86V1540/T200)/TX03A0148 (=TX89A7137/TIPACNA)	Texas
TX15M8239	TX04CS00189 (=Seri82/5009)/Doans (=AP02T4342=Coronado//1174-27-46/X960210)	Texas
TX15M8456	X08A586S [=ME-1-1/TX04M410164)]/TX01V5134RC-3	Texas
TX15M8558	TX03CS00115/Deliver	Texas
TX15M8596	TX04V075080/TX02A0252	Texas
WESLEY	KS831936-3 / NE86501 = Sumner sib (Plainsman V / Odesskaya 51)// Colt / Cody	Nebraska

Table 2. List of the number of replications and lines per year

Years	Replications per line		Lines
	Min	Max	
2015	25	25	26
2016	18	18	26
2017	3	4	100
2018	3	4	100
2019	3	4	100
2020	1	29	71
2021	1	27	65

Table 3: Summary statistics of female traits where N is the number of lines evaluated, Min is the minimum value, Max is the maximum value, and the median and mean were arithmetically derived.

Trait	N	Min	Max	Median	Mean
Gape Date	123	141	147	143	144
Gape Score	123	6.7	9.6	8.4	8.4
CHA Damage	123	6.5	7.2	6.8	6.8

Table 4: Phenotypic and genomic broad sense heritability of female floral traits, and variance components extracted from model 2 and GBLUP model

Trait	σ^2_G	σ^2_{GxE}	σ^2_{error}	H^2	Marker σ^2_G	Marker σ^2_{error}	Marker H^2
Gape Date	1.77***	0.55	0.87	0.88	0.58	0.26	0.69
Gape Score	0.59***	0.68	0.55	0.73	0.17	0.13	0.56
CHA Damage	0.08*	0.70	0.17	0.36	0.03	0.03	0.50

****. **, *** Significantly different at the p=0.05, 0.01, 0.001 probability level.***

Table 5: Marker-trait associations detected from GWAS for gape date, gape score and CHA damage

Trait	SNP	Chr	Physical position (bp)	P-value	LOD	MAF	R^2 of model with SNP	Effect	Phenotype variance explained (%)
gape date	S1A_590521793	1A	590521793	0.0010	3.01	0.21	0.37	-0.67	1.5E-08
gape date	S1B_291897879	1B	291897879	0.0008	3.09	0.15	0.37	0.44	8.3E-01
gape date	S1B_530957854	1B	530957854	0.0009	3.05	0.15	0.37	-0.45	2.9E+00
gape date	S2A_771152522	2A	771152522	0.0002	3.62	0.25	0.38	0.43	7.2E-01
gape date	S2B_58748696	2B	58748696	0.0008	3.11	0.32	0.37	-0.43	9.0E-01
gape date	S4A_690430825	4A	690430825	0.0002	3.73	0.23	0.39	0.47	0.0E+00
gape date	S4A_731097948	4A	731097948	0.0004	3.46	0.25	0.38	-0.81	9.6E+00
gape date	S4A_11319290	4A	11319290	0.0007	3.15	0.30	0.37	-0.96	5.6E+00
gape date	S5A_18653383	5A	18653383	0.0002	3.74	0.31	0.39	-0.56	2.3E-06
gape date	S5A_18653391	5A	18653391	0.0002	3.74	0.13	0.39	0.56	8.2E-02
gape date	S5A_18653369	5A	18653369	0.0007	3.14	0.33	0.37	0.56	0.0E+00
gape date	S6A_3505092	6A	3505092	0.0007	3.13	0.11	0.37	-0.63	5.7E-08
gape date	S6A_3505120	6A	3505120	0.0007	3.13	0.16	0.37	-0.63	1.5E-08
gape date	S6B_137355745	6B	137355745	0.0004	3.35	0.14	0.38	0.45	2.7E+00
gape date	S6B_720193972	6B	720193972	0.0005	3.28	0.14	0.37	0.58	3.4E+00
gape date	S6D_430047660	6D	430047660	0.0008	3.09	0.38	0.37	-0.54	1.8E+00
gape date	S7B_11703747	7A	11703747	0.0004	3.43	0.38	0.38	-0.39	0.0E+00
gape date	S7B_719857174	7B	719857174	0.0005	3.27	0.13	0.37	-0.38	4.7E+00
gape date	S7B_719857169	7B	719857169	0.0007	3.13	0.24	0.37	-0.73	0.0E+00
gape date	SUN_36058948	UN	36058948	0.0003	3.50	0.07	0.38	-0.58	7.2E-07
gape score	S1A_69763425	1A	69763425	0.0007	3.13	0.44	0.14	-0.37	1.3E-01
gape score	S1A_69763451	1A	69763451	0.0007	3.13	0.11	0.14	-0.37	5.4E-01
gape score	S1A_309699031	1A	309699031	0.0008	3.12	0.26	0.14	0.33	2.0E+00
gape score	S1A_520433427	1A	520433427	0.0009	3.06	0.11	0.14	0.22	5.4E-01

Trait	SNP	Chr	Physical position (bp)	P-value	LOD	MAF	R^2 of model with SNP	Effect	Phenotype variance explained (%)
gape score	S2A_750604838	2A	750604838	0.0001	4.17	0.05	0.18	-0.32	3.4E-01
gape score	S2B_742526639	2B	742526639	0.0003	3.48	0.23	0.15	-0.47	5.4E+00
gape score	S2B_10777749	2B	10777749	0.0009	3.05	0.49	0.14	-0.28	3.5E+00
gape score	S2D_10870835	2D	10870835	0.0003	3.53	0.49	0.16	0.34	3.9E-02
gape score	S2D_10870844	2D	10870844	0.0003	3.53	0.21	0.16	-0.34	6.7E-01
gape score	S3A_705230542	3A	705230542	0.0005	3.31	0.21	0.15	0.58	1.3E+00
gape score	S4A_599647431	4A	599647431	0.0007	3.17	0.17	0.14	0.28	1.7E+00
gape score	S4B_501710513	4B	501710513	0.0004	3.44	0.05	0.15	0.60	3.3E-07
gape score	S5A_707037627	5A	707037627	0.0004	3.40	0.05	0.15	-0.61	1.8E+00
gape score	S5A_707037628	5A	707037628	0.0004	3.40	0.05	0.15	-0.61	2.7E+00
gape score	S5A_707037653	5A	707037653	0.0004	3.40	0.45	0.15	-0.61	1.9E+00
gape score	S5A_678232514	5A	678232514	0.0007	3.13	0.49	0.14	-0.20	1.2E+00
gape score	S5B_485574635	5B	485574635	0.0006	3.22	0.20	0.14	0.30	1.5E-01
gape score	S6A_189384369	6A	189384369	0.0006	3.21	0.15	0.14	-0.28	0.0E+00
gape score	S6A_189384379	6A	189384379	0.0006	3.21	0.15	0.14	-0.28	0.0E+00
gape score	S6B_549563085	6B	549563085	0.0000	4.55	0.15	0.20	-0.37	1.2E+00
gape score	S6B_549563098	6B	549563098	0.0000	4.55	0.15	0.20	-0.37	5.7E-01
gape score	S6B_699005350	6B	699005350	0.0003	3.59	0.44	0.16	0.24	3.9E+00
gape score	S6D_241449380	6D	241449380	0.0010	3.01	0.44	0.14	0.25	4.8E-01
gape score	S6D_241449412	6D	241449412	0.0010	3.01	0.30	0.14	0.25	3.3E-01
gape score	S6D_241449394	6D	241449394	0.0010	3.01	0.18	0.14	-0.25	1.7E-01
gape score	S6D_241449401	6D	241449401	0.0010	3.01	0.49	0.14	-0.25	1.8E-01
gape score	S7A_9825402	7A	9825402	0.0010	3.01	0.49	0.14	-0.36	1.0E+00
gape score	S7B_639403457	7B	639403457	0.0002	3.73	0.44	0.16	0.38	0.0E+00
gape score	S7B_639403443	7B	639403443	0.0008	3.09	0.16	0.14	0.36	8.5E-02
gape score	S7D_20057921	7D	20057921	0.0008	3.10	0.29	0.14	0.53	0.0E+00

Trait	SNP	Chr	Physical position (bp)	P-value	LOD	MAF	R^2 of model with SNP	Effect	Phenotype variance explained (%)
gape score	SUN_398210579	UN	398210579	0.0001	4.15	0.29	0.18	0.45	8.8E-01
gape score	SUN_137679181	UN	137679181	0.0001	4.04	0.29	0.18	0.41	1.3E+00
gape score	SUN_301100107	UN	301100107	0.0007	3.13	0.29	0.14	0.45	2.4E-01
CHA damage	S2B_2416524	2B	2416524	0.0005	3.27	0.39	0.15	0.09	2.8E+00
CHA damage	S2D_22571312	2D	22571312	0.0005	3.30	0.12	0.15	-0.10	0.0E+00
CHA damage	S2D_22571267	2D	22571267	0.0005	3.30	0.41	0.15	0.10	0.0E+00
CHA damage	S3A_734184147	3A	734184147	0.0009	3.05	0.21	0.14	-0.08	4.5E+00
CHA damage	S3A_694947647	3A	694947647	0.0009	3.02	0.07	0.14	-0.06	8.3E-02
CHA damage	S3B_594745123	3B	594745123	0.0006	3.25	0.40	0.15	-0.06	7.4E-01
CHA damage	S3B_815282615	3B	815282615	0.0010	3.01	0.20	0.14	-0.08	1.1E+00
CHA damage	S4B_501710513	4B	501710513	0.0006	3.21	0.17	0.14	-0.15	0.0E+00
CHA damage	S5B_109210554	5B	109210554	0.0006	3.19	0.15	0.14	0.16	0.0E+00
CHA damage	S5D_527956626	5D	527956626	0.0006	3.23	0.17	0.15	0.09	3.6E-02
CHA damage	S5D_527956636	5D	527956636	0.0006	3.23	0.17	0.15	-0.09	2.6E+00
CHA damage	S7A_694790146	7A	694790146	0.0003	3.50	0.17	0.16	-0.12	6.7E-01
CHA damage	S7A_94602690	7A	94602690	0.0007	3.14	0.46	0.14	0.08	2.8E-06
CHA damage	S7B_622525086	7B	622525086	0.0010	3.01	0.46	0.14	0.10	4.9E+00
CHA damage	S7D_511050625	7D	511050625	0.0004	3.45	0.37	0.15	-0.10	0.0E+00
CHA damage	S7D_511050628	7D	511050628	0.0004	3.45	0.07	0.15	-0.10	0.0E+00
CHA damage	S7D_511050597	7D	511050597	0.0004	3.45	0.16	0.15	0.10	0.0E+00
CHA damage	S7D_20057921	7D	20057921	0.0009	3.07	0.39	0.14	-0.13	1.9E-09
CHA damage	SUN_288183582	UN	288183582	0.0002	3.76	0.27	0.17	-0.07	8.1E-01
CHA damage	SUN_43080472	UN	43080472	0.0007	3.14	0.10	0.14	0.13	0.0E+00

Table 6: Annotated candidate genes identified in *SnpEff* using the MTAs found for all female floral traits

SNP	Chr	Position bp	Candidate gene	Trait	Ref	Alt	Genomic region	Effect	Gene
S1A_590521793	1A	590521793	S1A_590521793	gape date	C	T	intergenic_region	MODIFIER	TraesCS1A02G441600-TraesCS1A02G441700
S1B_291897879	1B	291897879	S1B_291897879	gape date	T	C	intergenic_region	MODIFIER	TraesCS1B02G166200-TraesCS1B02G166300
S1B_530957854	1B	530957854	S1B_530957854	gape date	A	G	downstream_gene_variant	MODIFIER	No gene identified in the 200 kb window
S2A_771152522	2A	771152522	S2A_771152522	gape date	T	C	intergenic_region	MODIFIER	TraesCS2A02G576900-TraesCS2A02G577000
S2B_58748696	2B	58748696	S2B_58748696	gape date	T	C	intergenic_region	MODIFIER	TraesCS2B02G098900-TraesCS2B02G099100
S4A_11319290	4A	11319290	S4A_11319290	gape date	C	G	downstream_gene_variant	MODIFIER	No gene identified in the 200 kb window
S4A_690430825	4A	690430825	S4A_690430825	gape date	T	C	intergenic_region	MODIFIER	TraesCS4A02G419600-TraesCS4A02G419700
S4A_731097948	4A	731097948	S4A_731097948	gape date	C	T	synonymous_variant	LOW	No gene identified in the 200 kb window
S5A_18653369	5A	18653369	S5A_18653369	gape date	G	A	intergenic_region	MODIFIER	TraesCS5A02G022700-TraesCS5A02G022800
S5A_18653383	5A	18653383	S5A_18653383	gape date	A	G	intergenic_region	MODIFIER	TraesCS5A02G022700-TraesCS5A02G022800
S5A_18653391	5A	18653391	S5A_18653391	gape date	T	C	intergenic_region	MODIFIER	TraesCS5A02G022700-TraesCS5A02G022800
S6A_3505092	6A	3505092	S6A_3505092	gape date	G	A	intergenic_region	MODIFIER	TraesCS6A02G008400-TraesCS6A02G008500
S6A_3505120	6A	3505120	S6A_3505120	gape date	G	A	intergenic_region	MODIFIER	TraesCS6A02G008400-TraesCS6A02G008500
S6B_137355745	6B	137355745	S6B_137355745	gape date	A	G	intergenic_region	MODIFIER	TraesCS6B02G139900-TraesCS6B02G140000
S6B_720193972	6B	720193972	S6B_720193972	gape date	T	A	upstream_gene_variant	MODIFIER	No gene identified in the 200 kb window

SNP	Chr	Position bp	Candidate gene	Trait	Ref	Alt	Genomic region	Effect	Gene
S6D_430047660	6D	430047660	S6D_430047660	gape date	T	C	intergenic_region	MODIFIER	TraesCS6D02G323500-TraesCS6D02G323600
S7B_11703747	7B	11703747	S7B_11703747	gape date	C	T	intergenic_region	MODIFIER	TraesCS7B02G014500-TraesCS7B02G014600
S7B_719857169	7B	719857169	S7B_719857169	gape date	T	A	intergenic_region	MODIFIER	TraesCS7B02G462700-TraesCS7B02G462800
S7B_719857174	7B	719857174	S7B_719857174	gape date	G	T	intergenic_region	MODIFIER	TraesCS7B02G462700-TraesCS7B02G462800
S1A_309699031	1A	309699031	S1A_309699031	gape score	G	T	intergenic_region	MODIFIER	TraesCS1A02G173500-TraesCS1A02G173600
S1A_520433427	1A	520433427	S1A_520433427	gape score	C	T	intergenic_region	MODIFIER	TraesCS1A02G331700-TraesCS1A02G331800
S1A_69763425	1A	69763425	S1A_69763425	gape score	G	C	intergenic_region	MODIFIER	TraesCS1A02G084200-TraesCS1A02G084300
S1A_69763451	1A	69763451	S1A_69763451	gape score	G	A	intergenic_region	MODIFIER	TraesCS1A02G084200-TraesCS1A02G084300
S2A_750604838	2A	750604838	S2A_750604838	gape score	C	T	intergenic_region	MODIFIER	TraesCS2A02G536600-TraesCS2A02G536700
S2B_10777749	2B	10777749	S2B_10777749	gape score	A	G	upstream_gene_variant	MODIFIER	No gene identified in the 200 kb window
S2B_742526639	2B	742526639	S2B_742526639	gape score	C	A	intron_variant	MODIFIER	No gene identified in the 200 kb window
S2D_10870835	2D	10870835	S2D_10870835	gape score	G	A	missense_variant	MODERATE	No gene identified in the 200 kb window
S2D_10870844	2D	10870844	S2D_10870844	gape score	A	G	missense_variant	MODERATE	No gene identified in the 200 kb window
S3A_705230542	3A	705230542	S3A_705230542	gape score	A	G	missense_variant	MODERATE	No gene identified in the 200 kb window
S4A_599647431	4A	599647431	S4A_599647431	gape score	A	C	upstream_gene_variant	MODIFIER	No gene identified in the 200 kb window
S5A_678232514	5A	678232514	S5A_678232514	gape score	G	T	upstream_gene_variant	MODIFIER	No gene identified in the 200 kb window

SNP	Chr	Position bp	Candidate gene	Trait	Ref	Alt	Genomic region	Effect	Gene
S5A_707037627	5A	707037627	S5A_707037627	gape score	C	T	intergenic_region	MODIFIER	TraesCS5A02G555800-TraesCS5A02G555900
S5A_707037628	5A	707037628	S5A_707037628	gape score	G	A	intergenic_region	MODIFIER	TraesCS5A02G555800-TraesCS5A02G555900
S5A_707037653	5A	707037653	S5A_707037653	gape score	G	C	intergenic_region	MODIFIER	TraesCS5A02G555800-TraesCS5A02G555900
S5B_485574635	5B	485574635	S5B_485574635	gape score	C	A	intergenic_region	MODIFIER	TraesCS5B02G300500-TraesCS5B02G300600
S6A_189384369	6A	189384369	S6A_189384369	gape score	G	A	intergenic_region	MODIFIER	TraesCS6A02G174900-TraesCS6A02G175000
S6A_189384379	6A	189384379	S6A_189384379	gape score	T	C	intergenic_region	MODIFIER	TraesCS6A02G174900-TraesCS6A02G175000
S6B_549563085	6B	549563085	S6B_549563085	gape score	C	A	intergenic_region	MODIFIER	TraesCS6B02G306200-TraesCS6B02G306300
S6B_549563098	6B	549563098	S6B_549563098	gape score	T	A	intergenic_region	MODIFIER	TraesCS6B02G306200-TraesCS6B02G306300
S6B_699005350	6B	699005350	S6B_699005350	gape score	G	A	intron_variant	MODIFIER	No gene identified in the 200 kb window
S6D_241449380	6D	241449380	S6D_241449380	gape score	A	G	intergenic_region	MODIFIER	TraesCS6D02G185600-TraesCS6D02G185800
S6D_241449394	6D	241449394	S6D_241449394	gape score	T	C	intergenic_region	MODIFIER	TraesCS6D02G185600-TraesCS6D02G185800
S6D_241449401	6D	241449401	S6D_241449401	gape score	T	C	intergenic_region	MODIFIER	TraesCS6D02G185600-TraesCS6D02G185800
S6D_241449412	6D	241449412	S6D_241449412	gape score	C	T	intergenic_region	MODIFIER	TraesCS6D02G185600-TraesCS6D02G185800
S7A_9825402	7A	9825402	S7A_9825402	gape score	G	A	intergenic_region	MODIFIER	TraesCS7A02G025000-TraesCS7A02G025100
S7B_639403443	7B	639403443	S7B_639403443	gape score	G	A	intergenic_region	MODIFIER	TraesCS7B02G374000-TraesCS7B02G374100
S7B_639403457	7B	639403457	S7B_639403457	gape score	T	C	intergenic_region	MODIFIER	TraesCS7B02G374000-TraesCS7B02G374100

SNP	Chr	Position bp	Candidate gene	Trait	Ref	Alt	Genomic region	Effect	Gene
S2B_2416524	2B	2416524	S2B_2416524	CHA damage	A	C	intergenic_region	MODIFIER	TraesCS2B02G004500-TraesCS2B02G004600
S2D_22571267	2D	22571267	S2D_22571267	CHA damage	A	G	upstream_gene_variant	MODIFIER	No gene identified in the 200 kb window
S2D_22571312	2D	22571312	S2D_22571312	CHA damage	T	A	upstream_gene_variant	MODIFIER	No gene identified in the 200 kb window
S3A_694947647	3A	694947647	S3A_694947647	CHA damage	A	C	intergenic_region	MODIFIER	TraesCS3A02G457100-TraesCS3A02G457300
S3A_734184147	3A	734184147	S3A_734184147	CHA damage	G	C	intergenic_region	MODIFIER	PR4A-TraesCS3A02G517200
S3B_594745123	3B	594745123	S3B_594745123	CHA damage	G	C	synonymous_variant	LOW	No gene identified in the 200 kb window
S3B_815282615	3B	815282615	S3B_815282615	CHA damage	A	G	missense_variant	MODERATE	No gene identified in the 200 kb window
S4B_501710513	4B	501710513	S4B_501710513	CHA damage	A	C	intergenic_region	MODIFIER	TraesCS4B02G242300-TraesCS4B02G242400
S5B_109210554	5B	109210554	S5B_109210554	CHA damage	G	T	upstream_gene_variant	MODIFIER	No gene identified in the 200 kb window
S5D_527956626	5D	527956626	S5D_527956626	CHA damage	A	G	intergenic_region	MODIFIER	TraesCS5D02G497900-TraesCS5D02G498000
S5D_527956636	5D	527956636	S5D_527956636	CHA damage	T	G	intergenic_region	MODIFIER	TraesCS5D02G497900-TraesCS5D02G498000
S7A_694790146	7A	694790146	S7A_694790146	CHA damage	G	A	intergenic_region	MODIFIER	TraesCS7A02G508000-TraesCS7A02G508100
S7A_94602690	7A	94602690	S7A_94602690	CHA damage	G	A	intergenic_region	MODIFIER	TraesCS7A02G143600-TraesCS7A02G143700
S7B_622525086	7B	622525086	S7B_622525086	CHA damage	A	G	downstream_gene_variant	MODIFIER	No gene identified in the 200 kb window
S7D_20057921	7D	20057921	S7D_20057921	CHA damage	A	G	synonymous_variant	LOW	No gene identified in the 200 kb window
S7D_511050597	7D	511050597	S7D_511050597	CHA damage	A	G	intergenic_region	MODIFIER	TraesCS7D02G396300-TraesCS7D02G396400

SNP	Chr	Position bp	Candidate gene	Trait	Ref	Alt	Genomic region	Effect	Gene
S7D_511050625	7D	511050625	S7D_511050625	CHA damage	C	A	intergenic_region	MODIFIER	TraesCS7D02G396300-TraesCS7D02G396400
S7D_511050628	7D	511050628	S7D_511050628	CHA damage	G	A	intergenic_region	MODIFIER	TraesCS7D02G396300-TraesCS7D02G396400
SUN_137679181	UN	137679181	SUN_137679181	Unknown	A	T	Unknown	MODIFIER	No gene identified in the 200 kb window
SUN_288183582	UN	288183582	SUN_288183582	Unknown	C	T	Unknown	MODIFIER	No gene identified in the 200 kb window
SUN_301100107	UN	301100107	SUN_301100107	Unknown	G	A	Unknown	MODIFIER	No gene identified in the 200 kb window
SUN_36058948	UN	36058948	SUN_36058948	Unknown	T	C	Unknown	MODIFIER	No gene identified in the 200 kb window
SUN_398210579	UN	398210579	SUN_398210579	Unknown	A	G	Unknown	MODIFIER	No gene identified in the 200 kb window
SUN_43080472	UN	43080472	SUN_43080472	Unknown	A	G	Unknown	MODIFIER	No gene identified in the 200 kb window

Table 7: Most promising candidate genes for female floral traits determined by their location in the missense genomic region

SNP	Chr	Position bp	P.value	LOD	MAF	Rsquare of model with SNP	PVE (%)	Gene	Trait
S2D_10870835	2D	10870835	0.00029	3.53	0.38	0.16	0.04	TraesCS2D02G024600	gape score
S2D_10870844	2D	10870844	0.00029	3.53	0.38	0.16	0.67	TraesCS2D02G024600	gape score
S3A_705230542	3A	705230542	0.00049	3.31	0.07	0.15	1.26	TraesCS3A02G474000	gape score
S3B_815282615	3B	815282615	0.00098	3.01	0.11	0.14	1.12	TraesCS3B02G590100	CHA damage

Figure 1: Ideal parents for hybrid wheat. A is an ideal female plant with florets widely open, receptive stigmas and male sterile. B. is an ideal male plant with open florets and anthers extruding out of the floret and producing pollen.

Figure 2: Hybrid Crossing Block Design where the female is surrounded by a male on both sides.

Figure 3: Female floral traits measured in the phenotypic evaluation. The gape date is the Julian date when the whole plot reached maximum gaping, gape score is the opening of the florets at gape date in a 1-9 scale where 1 denoted the least gaping, and 9 denoted widely open florets, and CHA damage is the level of phytotoxicity of the plant after emasculated with CHA, measured in a 1-9 scale where 1 denoted no visible head damage and 9 denoted heads significant injuries measured in this study (Stoll., 2019).

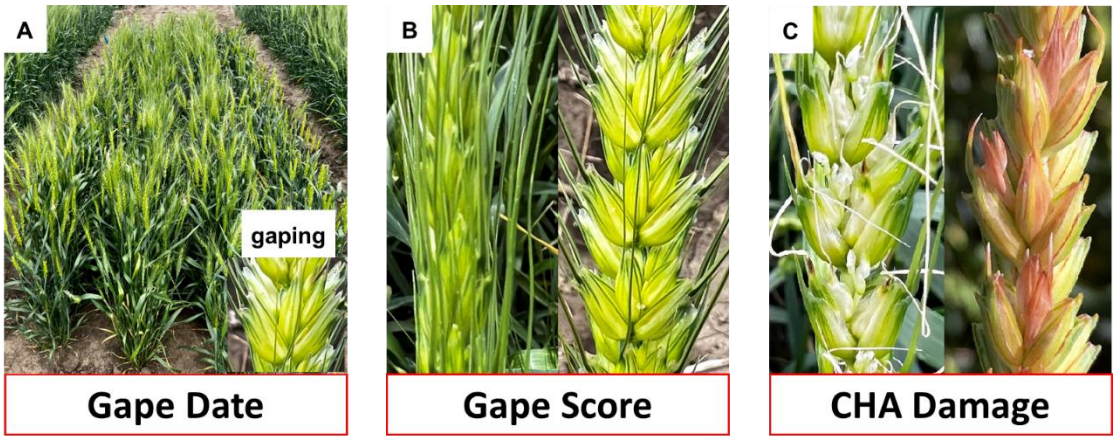

Figure 4: Distributions of the best linear unbiased prediction (BLUPs) of gape date showing wide variation of gape date (144-147)

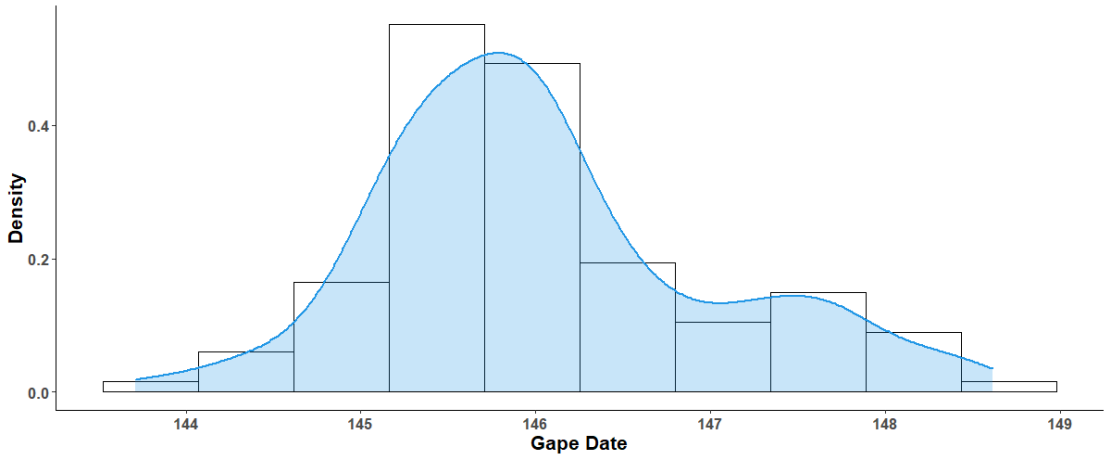

Figure 5: Distributions of the best linear unbiased prediction (BLUPs) of gape score. It is considered as a superior gape score the values >6

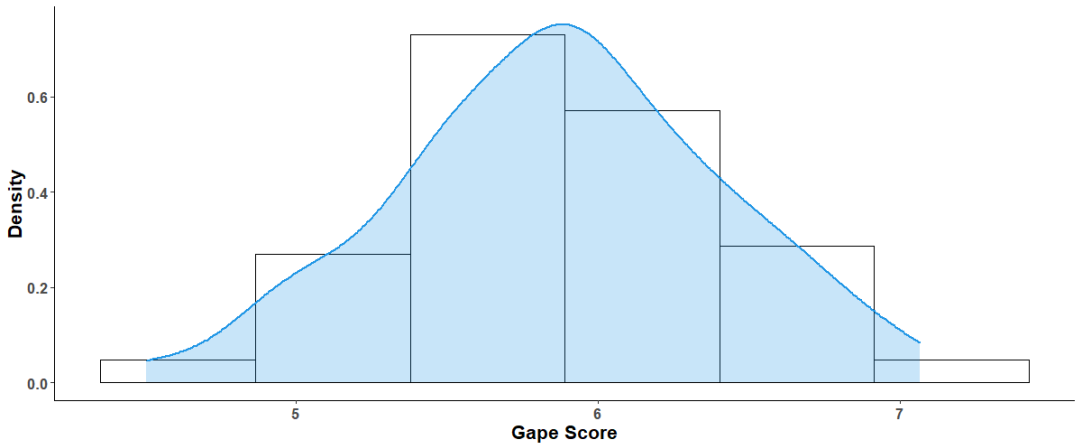

Figure 6: Distributions of the best linear unbiased prediction (BLUPs) of CHA damage. A line is considered as tolerant to CHA damage when values are <5

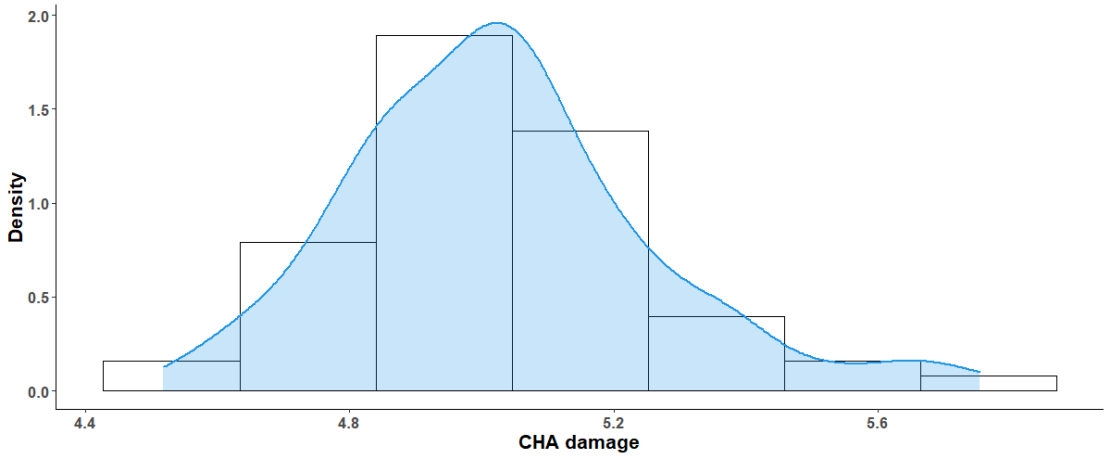

Figure 7: The most important three first principal components based on percentage of variation explained. The plot is showing the clustering of the female lines based on their similarity. Red dots represent Nebraska lines and Black dots represent Texas lines.

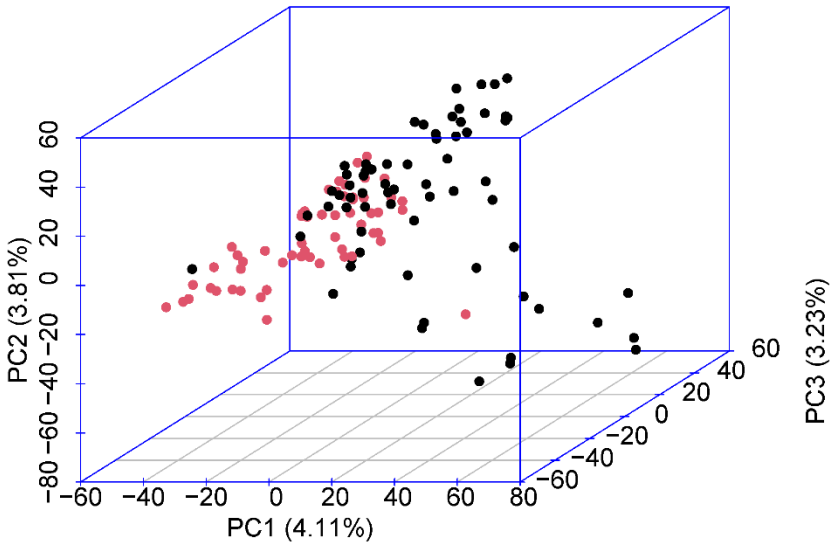

Figure 8: Eigen values plot of the first ten principal components

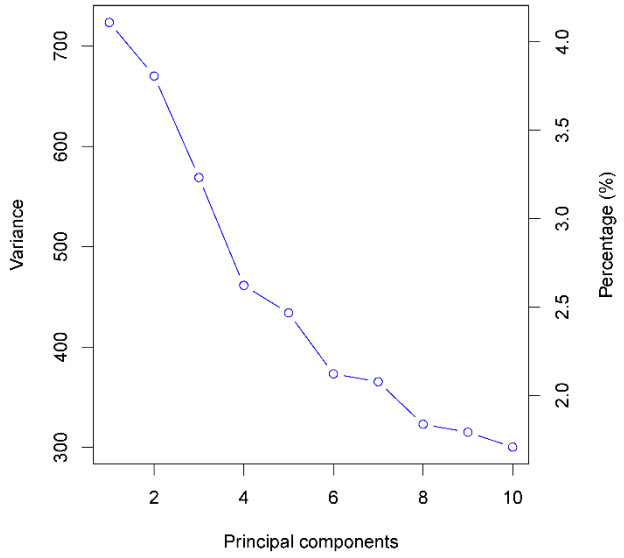

Figure 9: Manhattan plot of 20 MTA found in GWAS for gape date using a threshold of significance of P<0.001 (−log10(P) ≥ 3.0)

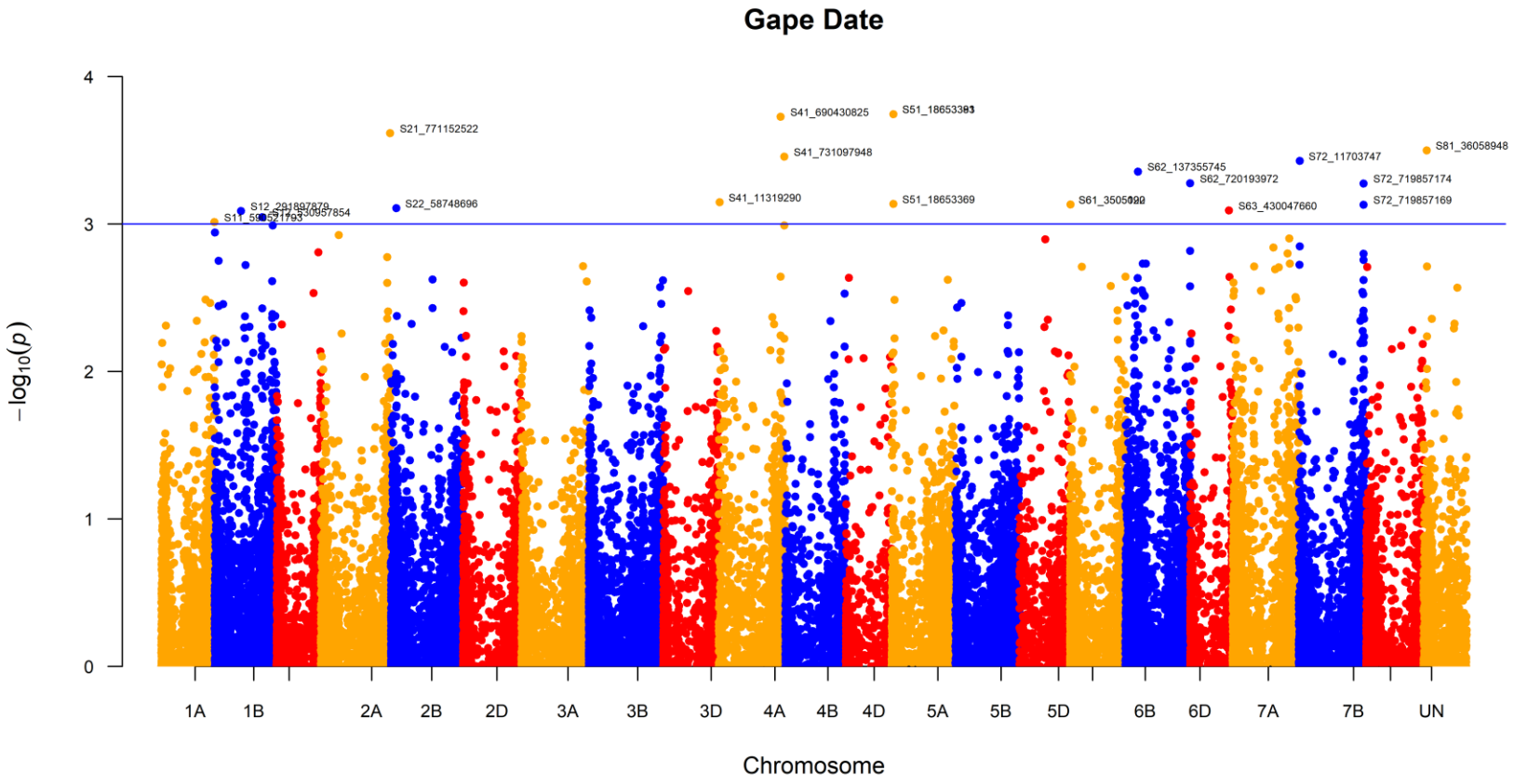

Figure 10: Manhattan plot of 33 MTA found in GWAS for gape score using a threshold of significance of P<0.001 (−log10(P) ≥ 3.0)

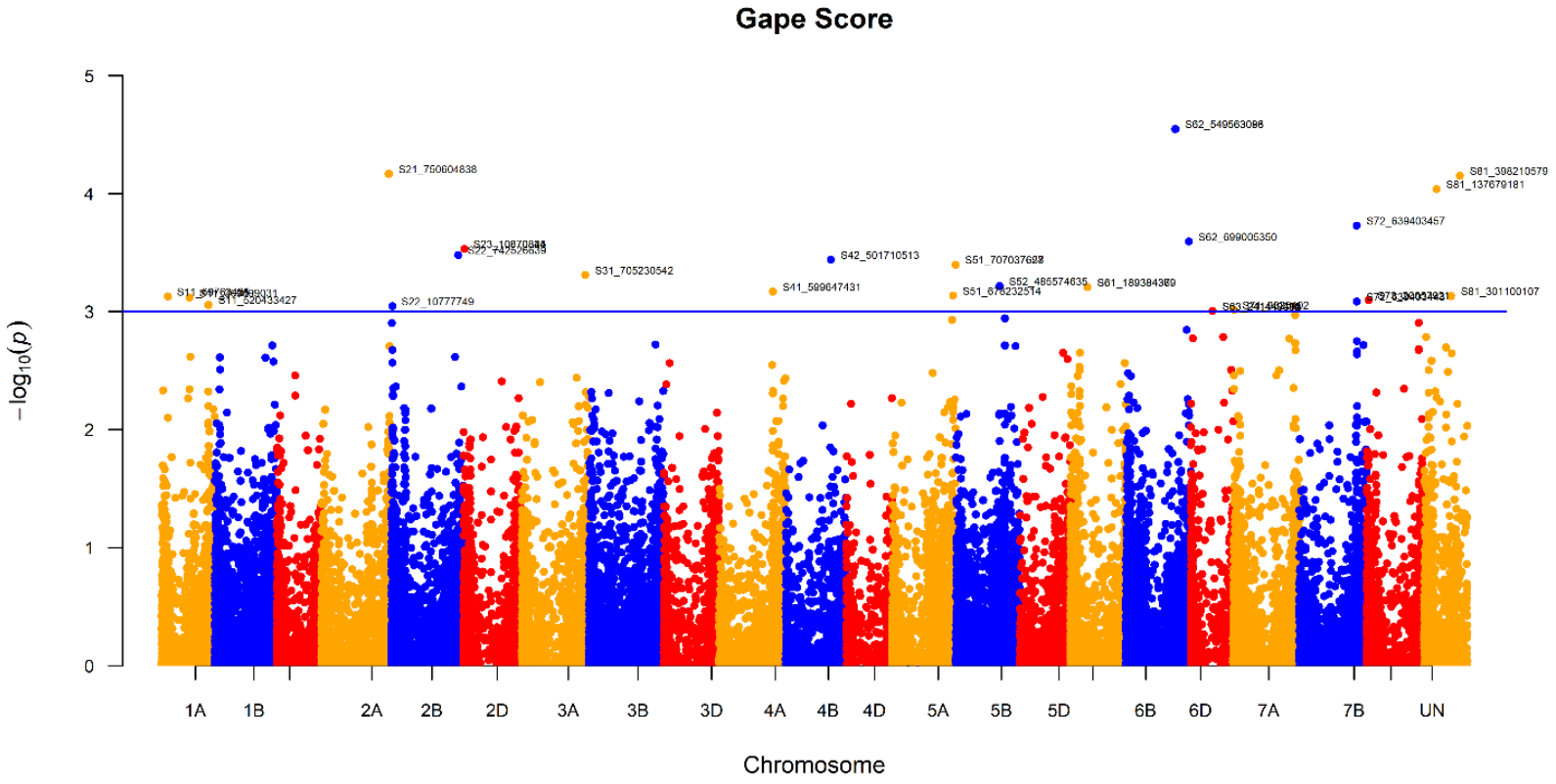

Figure 11: Manhattan plot of 20 MTA found in GWAS for CHA Damage using a threshold of significance of P<0.001 (−log10(P) ≥ 3.0)

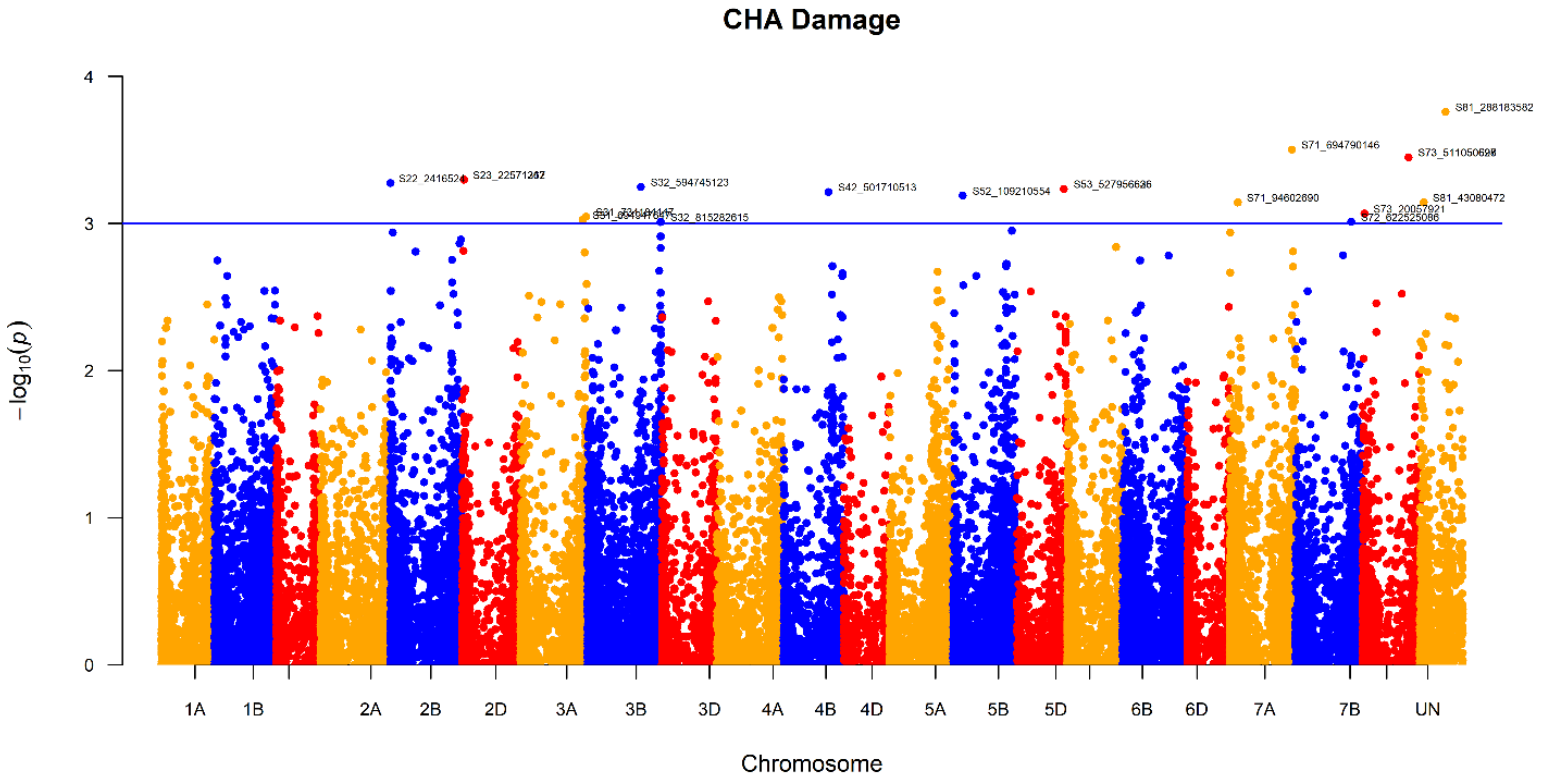

Figure 12: Q-Q plots for gape date, gape score, and CHA damage