MAPPING OF TRAITS ADAPTIVE TO THE U.S. SOUTHERN AND CENTRAL GREAT PLAINS IN A 'TAM 204'/'IBA' POPULATION AND DEVELOPING SUPERIOR PRE-HARVEST SPROUTING TOLERANT HARD WHITE WINTER WHEAT VERSION OF 'TAM 114' USING CRISPR-CAS9 GENE EDITING

A Dissertation

by

JORGE LUIS VALENZUELA ANTELO

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Amir M.H. Ibrahim
Co-Chair of Committee,	Shuyu Liu
Committee Members,	Jackie C. Rudd
	Michael Thomson
Head of Department,	David Baltensperger

December 2020

Major Subject: Plant Breeding

ABSTRACT

The development of superior cultivars with outstanding agronomical performance, enhanced pest resistance, or special traits could be facilitated with the use of molecular breeding techniques. To achieve this, the Texas A&M wheat breeding and genetics program have developed the bi-parental population derived from ‘TAM 204’ and ‘Iba’ to identify genomic regions with favorable alleles for pest resistance and agronomical traits. The 221 recombinant inbred lines of this population were screened for Greenbug and Hessian fly under greenhouse conditions as well as grown in 11 environments across Texas. Also, a second project to develop a white grain isoline of ‘TAM 114’ was established.

In the first study, the accuracy of the linkage groups was validated by mapping highly heritable traits. Quantitative Trait Loci (QTL) associated with Hessian fly were detected on chromosome 1A and 3B. A major QTL for Greenbug resistance was identified in chromosome 7D. For agronomical traits, a total of 86 QTL were identified, for which 38 were consistent. For grain yield, ten consistent QTL were identified on chromosomes 2A, 2B, 2D, 3D, 5A and 6A, with the QTL located at 29 Mb on 2D having a high additive effect. For plant height, 11 QTL were identified in chromosomes 2A, 2B, 4A, 4B, 4D, 5A, 6A and 6D, whereas for days to heading, four consistent QTL were found in chromosomes 2B, 2D and 5B. With regards to the test weight and thousand kernel weight traits, eight and three consistent QTL were detected, respectively.

For the second project, two highly specific sgRNA for multiplex editing of the Tamyb-10 homeologous genes were designed and their cleaved efficiency was validated in vitro. The Immature embryo transformation was performed using biolistics. A total of 63 plants were regenerated out of 266 immature embryos, demonstrating that TAM 114 has a modest

regeneration capability. Genotyping of the targeted regions using Sanger sequencing was performed, but no results were obtained. Additionally, an alternative tissue transformation protocol using mature embryos was tested. This second protocol significantly reduced the time and labor associated with tissue culture; however, the transformation efficiency of both protocols has not been tested yet.

DEDICATION

I dedicate my dissertation to my mother, father and brother

To my future wife Irene Elizabeth

To Justo Antelo Leon and Maria de Jesus Parada

To all my family and friends

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to all the people who have helped me during the pursuit of my doctoral degree. It has been a wonderful and rich experience. I want to thank Dr. Amir M.H Ibrahim and Dr. Shuyu Liu for their guidance and support throught my journey. Also, I want to thank Dr. Jackie Rudd and Dr. Michael Thomson for their help as committee members.

I owe my gratitude to the Wheat Breeding and Genetics program at Texas A&M. This could not be possible without all of their help while conducting the experiments and processing and collecting data,. I extend my appreciation to Dr. Geraldin Opena, Dr. Bryan Simoneaux, and Dr. Chengenn chu, and all the staff who helped make this possible.

My sincere gratitude goes to the staff of the Crop Genome Editing Lab for their help and guidance while I was working on my gene-editing project. I want to thank Dr. Backki Kim, Dr. Nikolaos Tsakirpaloglou, Dr. Nithya Subramanian, Sudip Biswas and Oneida Ibarra.

I want to thank all my colleagues for their friendship and support. I thank Mehmet Dogan, Zhen Wang and Mustafa Cerit for their support in field phenotyping. I would like to thank Ilse Barrios for her friendship and unconditional support. I want to thank Margaret Krause for her help in the data analysis. I want to thank the wheat breeding and genetics former students, Betul Sade, Dr. Mahendra Bhandari, Dr. Anil Adhikari and Dr. Smit Dhakal.

I want to thank Dr. Charlie Johnson and Dr. Shichen Wang for their support in genotyping and processing the genetic markers data.

I want to thank Dr. Xiangyang Xu and Dr. Ming Shun Chen, as well as their staff for the evaluation of Greenbug and Hessian.

I owe my gratitude to the administrative staff at the Department of Soil and Crop Sciences and Texas A&M AgriLife Research for their support and facilitation of my doctoral degree program.

I owe my gratitude to Bayer Crop Science, formerly Monsanto for their financial support. I thank the U.S. Department of Agriculture National Institute of Food and Agriculture and the Wheat Coordinated Agricultural project for the training in QTL mapping.

I want to thank Dr. Ravi Singh, Dr. Julio Huerta, Dr. Francisco Pinera and Dr. Leonardo Crespo from the Global Wheat breeding at International Maize and Wheat Improvement Center, and Dr. Hector Villasenor Mir and Dr. Ignacion Benitez for all the training and encouragement to pursue a Doctoral Degree at Texas A&M.

Thanks to my mother, father, and brother, whom I owe everything to. Finally, I would like to thank my fiancé, Irene Elizabeth, for her patience, support and love.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by the dissertation committee consisting of Dr. Amir M.H. Ibrahim, Dr. Shuyu Liu, Dr. Jackie C. Rudd and Dr. Michael Thomson.

The DNA extraction and sample preparation for genotyping of the recombinant inbred lines and parental lines were performed at the Texas A&M AgriLife Research station at Amarillo, Tx, and the AgriGenomics Lab at College Station, TX.

Genotyping was performed at the Genomics and Bioinformatics Center at Texas A&M University.

The data presented in chapter 2 was obtained in collaboration with the USDA/ARS in Manhattan KS and Stillwater, OK.

All the experiments presented in chapter 3 were conducted at the Texas A&M AgriLife research station. The phenotypic data presented in this chapter was collected by Texas A&M AgriLife staff and students.

The gene-editing project was conducted in the Crop Genome Editing Lab at Texas A&M, and all the work was done by the student in collaboration with the lab staff.

Funding Sources

Graduate study was supported by a fellowship from Bayer Crop Science, former Monsanto, and Texas A&M AgriLife research.

NOMENCLATURE

AbyE	Additive-by-environment interaction
ANOVA	Analysis of variance
BLAST	Basic Local Alignment Search Tool
BLUE	Best Linear Unbiased Estimators
CRISPR	Clustered Regularly Interspaced Palindromic Repeat
CTAB	Cetyltrimethylammonium Bromide
DNA	Deoxyribonucleic acid
DTH	Days to heading
GB	GreenBug
GBS	Genotype-by-Sequencing
GEI	Genotype-by-Environment Interaction
GFP	Green Fluoresecent protein
GMO	Genetically modified organism
GPM	Grains per meter
GYLD	Grain yield
HF	Hessian Fly
HRWW	Hard Red Winter Wheat
KASP	Kompetitive Allele-Specific
LOD	Logarithm of the odds
MAS	Marker-Assisted Selection
ME	Mega-Environment

NB-LRR	Nucleotide-binding leucine-rich repeat
PAM	Protospacer adjacent motif
PCR	Polymerase Chain Reaction
PH	Plant height
QTL	Quantitative trait loci
REML	Residual Maximum Likelihood
RIL	Recombinant Inbred Lines
RNA	Ribonucleic acid
RNP	Ribonucleoproteins
SAM	Shoot apical meristem
SGRNA	Single guided RNA
SNP	Single Nucleotide Polymorphism
TKW	Thousand Kernel Weight
TW	Test Weight
YFP	Yellow fluorescent protein

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

CHAPTER I

GENERAL INTRODUCTION

Wheat (*Triticum aestivum*) is a staple food crop for most of the world population. It is consumed by 35 % of the population and provides 20 % of the protein and daily dietary energy of the human diet (Velu et al., 2017). Rising demand for wheat is expected due to the increasing human population and consumption affluence (Ray et al., 2013). Wheat is grown worldwide on about 214 million ha with a total production of 734 million tons(FAO, 2020).

The increase in global wheat yields in the period of 1961 to 2007 was about 40 kg per year, whereas in the period from 2005/2007 to 2050 it is projected to have an increase of 24 kg per year (Alexandratos and Bruinsma, 2012). This represents a gap between the increase in population and global wheat production. Therefore, it is imperative to draw upon different technologies that would facilitate the development and selection of superior cultivars that can cope with future challenges and meet the demands. The identification of quantitative trait loci (QTL) along with the use of marker-assisted selection (MAS) facilitates the selection of genotypes with desirable adaptive traits. Another novel approach is gene-editing technologies that could be used to create a targeted genetic variation of existing cultivars.

The United States produced 51 million tons of wheat in 16 million ha in 2018 (FAO, 2020). The most popular wheat class grown in the country is Hard Red Winter Wheat (HRWW), and it is mostly grown in the Southern and Central Plains and the Pacific Northwest. Among the major constraints in HRWW production in the country stand out the impact of biotic stresses namely diseases and pest in grain production. The Texas A&M AgriLife wheat breeding program has been very successful in developing broadly adapted HRWW, mostly for the state of Texas and Southern Great Plains. Their improved genetic germplasm possesses outstanding characteristics

that can cope with the most prevalent biotic and abiotic stresses, as well as excellent high yield potential, drought tolerance, superior end-use quality attributes.

CHAPTER II

GENOMICS AND MOLECULAR MARKERS

2.1 Introduction

Recent advances in bioinformatics and genomics are permitting an easy adoption of molecular markers in more breeding programs given their increased availability, decreased cost and good genome coverage. Single Nucleotide Polymorphism (SNP) Arrays and Genotype-by-Sequencing (GBS) provide good insights while investigating genetic diversity in populations that could be used in breeding applications (Darrier et al., 2019). Molecular markers obtained from these platforms are commonly used in molecular breeding strategies such as Genomic Wide Selection (GWS) and Marker-Assisted Selection (MAS). In the first strategy above, cross-validation schemes can be obtained from four different scenarios comprising of combinations of tested lines, untested lines, tested environments and untested environments(Crossa et al., 2017). the preference of the prediction model differs depending on the trait of interest. For highly quantitative traits such as grain yield, Genomic Best Linear Unbiased Predictor (G-BLUP) and Ridge-Regression BLUP (RR-BLUP) are the most commonly used models, and for qualitative traits, Bayesian models are utilized. This is due to several assumptions concerning the models, of which the first two presume that a trait is controlled by an infinitesimal number of additive loci, which works very well for many quantitative traits. However, for traits that have large effect locus or loci, the assumption of Bayesian models' unique variances leads to a better prediction (Poland and Rutkoski, 2016).

On the other hand, MAS starts by identifying molecular markers that are tightly linked and co-segregate with genes of interest via QTL mapping. This genetic mapping approach started with bi-parental populations, in which parents with contrasting phenotypes are crossed and their

progeny, in the form of segregating populations after the F_2 generation, become Recombinant Inbred Lines (RIL) and are screened for phenotyping differences (Pascual et al., 2016). This information is used afterward in statistical analysis to determine the genetic marker-trait association based on critical recombination events. Another common population type is the association mapping, a.k.a. Linkage Disequilibrium Mapping, a diverse panel used to exploit historical and evolutionary recombination events in the population under study (Zhu et al., 2008). In the past few years, more complex mapping populations involving multiple parents crossing such as Multi-parent Advance Generation Intercross and Nested Association Mapping have also recently been adopted to improve mapping resolution(Yu et al., 2008; Anderson et al., 2018).The MAS approach is very practical for qualitative traits given its monogenic nature. Usually, it tends to have low prediction accuracy when used for quantitative traits since the markers used typically do not capture most of the phenotypic variation (Poland and Rutkoski, 2016).

The Texas A&M wheat genetics and breeding program has developed multiple mapping populations to identify causal markers linked to favorable genes in their germplasm to use for diagnosis when evaluating superior breeding lines using MAS. Among the QTL and genes that have been identified, a few genes that stand out include those that confer resistance to Wheat Streak Mosaic Virus 2 (*WSM2*), Greenbug (*GB3, GB7*) and Hessian fly (*H32*) (Weng et al., 2005; Dhakal et al., 2017; Tan et al., 2017).

In order to build on this success and take advantage of the most recent genotyping technologies, the Texas wheat program developed a bi-parental population from the cross between ‘TAM 204’ and ‘Iba’ (T4I) to map adaptive traits for the wheat-growing regions in the Southern and Central Great Plains.

2.2 Objectives

1. Genotype the lines that comprise the T4I population, using Double digest restriction-site associated DNA (ddRADseq).
2. Identify and select non-redundant polymorphic markers that are evenly distributed across the population and do not deviate from the Mendelian segregation ratio.
3. Create reliable linkage maps for QTL mapping with the obtained markers.

2.3 Materials and Methods

2.3.1 DNA extraction and genotyping

Genomic DNA from the RILs and parental lines was extracted using the Cetyltrimethyl Ammonium Bromide (CTAB) modified protocol (Doyle & Doyle, 1990; Liu et al., 2013a). DNA quantity and quality was analyzed using NanoDrop Ds-11 (DenovixTM) and visually compared to lambda DNA control on an agarose gel. Genotyping was done in the Genomics and Bioinformatics Center at Texas A&M University (https://www.txgen.tamu.edu/) using the ddRADseq genotyping protocol in Illumina Novaseq 6000 sequencing system. The SNP´s calls were aligned to the 'Chinese Spring' reference genome from the International Wheat Genome Sequencing Consortium (IWGSC) (Mayer et al., 2014) using Bowtie2 (Langmead and Salzberg, 2012). Subsequent steps of markers sorting, local realignment and mapping quality parameters were similar to Yang et al. (2020) using SAMtools (Li et al., 2009) and GATK V3.8 (McKenna et al., 2010). Imputation was done using Beagle V4.1 (Browning and Browning, 2016) with the parameters of "windowc = c5,000 overlapc = c500 burns-itsc = c10 impute-itsc = c10", similar to Yang et al. (2020). Non-polymorphic SNPs were eliminated. Genotypes of AA, CC, GG, and TT were converted to A or B depending on their allelic similarity with regards to the parental line, A

for the female parent (TAM 204) and B the for male (Iba). SNPs with more than 20% of missing values (≥ 44 lines) and 4 ≥ selected ≤ 0.25 A/B ratio were discarded. Posteriorly, the markers were binned to a chromosome using the BIN function in QTL IciMapping software version 4.1 (hereafter referred to as QTL Icimapping) (Meng et al., 2015). SNPs with more than 10% missing values across the population were removed. After this step, SNPs statistically segregate from the 1:1 segregation ratio and similarity more than 95% were discarded using JoinMap 4.0 (Van Ooijen, 2006). In the remaining SNPs set, false double crossovers were manually checked and removed according to the SNPs' alignment order based on their physical base pair location in the IWGSC RefSeq v1.0 (Appels et al., 2018).

2.3.2 Linkage groups construction and validation

Linkage groups were created using the obtained markers using JoinMap 4.0 and their reliability was tested by mapping QTL associated with highly heritable traits that segregate in the population using the BIP function in QTL IciMapping.

For linkage map construction, grouping trees were created using independence LOD as a grouping parameter with LOD values from 2-40, mapping algorithm by using maximum likelihood and Kosambi mapping function (Kosambi, 2016). For the QTL analysis, the Inclusive Composite Interval Mapping of Additive (ACIM-ADD) function was selected as the mapping method to detect additive QTL. The mapping parameters for this step were: set a step at 1.0 cM with a probability of a stepwise regression of 0.05. The LOD threshold for QTL detection was determined by 1000 permutation test analysis using type 1 error set at P <0.001. Data was collected from Bushland 2019 and College Station 2020, given both environments had the whole set of lines. Data was recorded zero and 100 for the absence and presence of awns, respectively. The presence of

pigmentation in the glume was recorded as 100 and absence as 0. For both traits, lines with mixed phenotypes were recorded as 50 %.

2.4 Results and Discussion

A total of 140503 markers were obtained from the sequencer and further imputed to reduce the number of missing markers, based on neighboring markers and haplotypes across lines. Despite adopting this procedure, 41831 markers had missing values for more than 20% of the lines and 3035 markers deviated from the A/B ratio, resulting in a total of 95636 markers. This new set of markers was binned to remove redundant markers that would not increase the resolution of the mapping results but would affect the computational performance and increase the analysis time. The BIN function from QTL ICIMapping permitted us to significantly reduce the number of markers, ranging from 9.5 to 44.51%, but overall, 32.85% of markers were removed. Given 31.4 K was still a significant number of SNPs to be used for QTL mapping, stringent criteria were applied to retain them with at least 10 % missing values, yielding into 12332 markers. Lastly, they were entered into Joinmap for a second round to remove segregation distortion markers and those with a similarity of more than 95%. This resulted in a final set of 10,186 markers. False double cross-overs created by genotyping errors were removed.

Subsequently, 60 linkage groups, with a varied number of groups per chromosome, brought the genetic map's total length to 3988.1 cM with an overall distance between two markers of 0.039 cM. Concerning the physical position, the genetic map's total length was 11409.76 Mb, with an overall distance between two markers of 0.89 Mb. The marker distribution for genome A, B and D was 30.7, 48.4 and 20.8%, respectively. In terms of the homoeologous chromosome distribution, chromosomes 5 and 1 have 20.4 and 20%, respectively, whereas 3 had the lowest number of markers (4.6%) (Supplemental Table 1, 2 and 3).

The linkage groups' accuracy was confirmed by mapping a QTL associated with the phenotypic expression for awns in chromosome 5A. In this chromosome, Kato et al., (1998) identified the *B1* gene, which plays a role in spike morphology by inhibiting the awn development. A study conducted by DeWitt et al. (2019) identified a candidate transcriptor repressor in the TraesCS5A02G542800. They also mentioned that the *B1* locus is a 25Mb region that goes from 681-706Mb in the Chinese Spring reference genome, which is the range where the QTL Qawns.tamu.5A.696 was detected (Table 2.1). For glume color, the QTL Qgc.tamu.1B.9 was found in chromosome 1B (Table 2.1). A previous study also detected a QTL chaff color in this genomic region when testing their linkage groups' accuracy (Hussain et al., 2017). Khlestkina et al., (2006) mapped a major gene known as *Rg1* in this region between the xgwm1078 and Xgwm0550. This QTL may explain a relatively low percentage of the Phenotypic Variation Explained (PVE) which could be due to the used binary phenotyping scale. A more representative scale that includes different phenotypes or a more precise phenotyping tool would improve PVE, given it was possible to see white, yellow, red and black glume colors. This would also help confirm this finding and potentially find other QTL associated with the homeologous Rg genes in chromosome 1A and 1D.

Table 2.1 Quantitative trait loci (QTL) for presence of awns and glume color detected in the TAM 204/Iba population in Bushland 2019 and College Station 2020.

QTL Name	Chrom	Position (Mb)	Traits	Threshold	Overall LOD	Total PVE ranges	Additive effects	Parental favorable alleles	25LGs	Peak position cM	Left SNPs	Left SNPs alleles	Left SNPs TAM 204	Left SNPs IBA	Right SNPs	Right SNPs alleles	Right SNPs TAM 204	Right SNPs Iba
Qawns.tamu.696	5A	695.72	Presences of awns	3.51	36.71-46.26	53.35-60.94	-34.16, -37.42	IBA	13	144.86	chr5A_695729235	A/G	A	G	chr5A_700348846	G/T	G	T
Qglc.tamu.1B.9	1B	8.9	Glume color	3.51	12.31	22.63	22.74	TAM 204	3	0	chr1B_9390786 chr1B_8968035	A/C	C	A	chr1B_8181483	T/C	C	T

Chrom = chromosome; LOD= logarithm of the odds; PVE= Phenotypic Variation Explained; LG= Linkage Groups; cM = centiMorgan; SNP = Single Nucleotide Polymorphism; chr_bp= physical position in base pairs.

2.5 Conclusion

A high-density and high-quality genetic map was successfully created. Mapping two highly heritable traits confirmed the reliability and accuracy of the genetic map, given their location in the Chinese Spring reference genome has been previously described. This gives the certainty that other mapped traits for this population, whose genetic position in the reference genome is currently unknown, could be factual.

CHAPTER III

QTL MAPPING FOR HESSIAN FLY AND GREENBUG RESISTANCE

3.1 Introduction

Hessian fly (HF) and Greenbug (GB) are two of the most destructive pests of wheat yield and end-use quality (Lu et al., 2010; Li et al., 2015; Tan et al., 2017). HF can damage the wheat plant at any growth stage. During the seedling stage, the larvae feed between leaf sheaths, limiting growing and leading to the death of the infested plant if new uninfected tillers do not develop; whilst in the reproductive stage, the larvae feed on the stem causing lodging and reduced grain filling (Byers and Gallun, 1972; Garcés-Carrera *et al*., 2014). GB causes highly visible damage ranging from chlorotic spots, severe necrosis to dead leaves and plant mortality when a heavy infestation occurs (Burton, 1986). Even though GB feeds from the aerial part of the plant, it causes a reduction in root development and biomass that may lead to a reduction in functionality and the full potential of the plant (Burton, 1986).

In order to avoid HF and GB infestations, it is important to have an Integrated Pest Management (IPM) program which includes cultural, chemical and biological control, as well as host-plant resistance (genetic control)(Hao et al., 2013; Garcés-Carrera et al., 2014). However, chemical treatments can control or reduce the infestation but lead to extra financial costs if several applications are needed. Agronomical control may also have some limitations, such as late planting only being an alternative in the northern United States (Chen et al., 2009; Garcés-Carrera et al., 2014). Therefore, the development of resistant genotypes is the most effective and economic component of the IPM approach.

Resistance genes tend to breakdown when they are deployed in a large area and over a long period of time due to the constant co-evolution of the HF and GB alongside the wheat, favoring the proliferation of virulent biotypes that are uncommon to race-specific genes (Hao et al., 2013; Garcés-Carrera et al., 2014; Tan et al., 2017). Hence the importance of the constant monitoring of population in order to safeguard the effectiveness of the plant resistance and know which biotypes are prevalent in each region (Chen et al., 2009; Garcés-Carrera et al., 2014). The southeastern region of the United States has historically been affected by HF, nevertheless, in recent years, heavy infections have become more intense and recurrent (Garcés-Carrera et al., 2014). The most prevalent biotype in the Great Plains is called GP (Li et al., 2015).

In pest evaluation, a gene is defined as highly resistant, moderately resistant and susceptible if ≥80%, 80-50% and <50% of the plants in a population, respectively, were resistant to HF in a virulence essay (Garcés-Carrera *et al*., 2014). 35 resistance genes have been identified and seven of them are located in the short arm of the chromosome 1A and confer resistance against biotype GP (*H5*, *H9*, *H10*, *H11*, *H16*, *H17*, and *Hdic)* (Williams et al., 2003; Li et al., 2015). Garcés-Carrera *et al*., 2014, in a study where five populations of HF were tested in a set of different cultivars carrying one or more resistant genes, found that *H12*, *H13*, *H17*, *H18*, *H21*, *H22*, *H25*, *H26* or *Hdic* genes showed highly resistance reaction to all three Texan HF populations. In contrast, *H3* or *H11* showed high resistance to two populations. The cultivar containing *H9*, *H16*, *H19*, or *H23* exhibited moderate resistance to at least one of the population. *H6*, *H7*/*H8*, *H10*, *H24*, and *H31* genes and gene combinations were susceptible to any one of these three populations (Garcés-Carrera et al., 2014).

In the case of GB, over 20 different biotypes have been classified through A to K but other isolates have been identified in New York, Kansas, Texas, and Florida (Lu et al., 2010; Tan et

al., 2017). Biotypes E and I have more of a host range and are the prevalent biotypes in Kansas, Nebraska, Oklahoma, and Texas states (Burd y Porter, 2006; Lu et al., 2010). The resistance genes *Gb2* and *Gb6* are present in the 1AL.1RS translocation (Lu et al., 2010), while Gb1 originated from durum wheat. *Gb4* was derived from *Aegilops Tauschii* and is located in the 7DL chromosome along with other resistant genes such as *Gba*, *Gbb*, *Gbc*, *Gbd*, G*bx1*, and *Gb3*. *Gb5* was transferred from *Ae. Speltoides L.* to the chromosome 7AL (Tan et al., 2017). *Gb3* has been identified in TAM 110 and TAM 112 cultivars, and *Gb7* was identified in a synthetic wheat line that possesses a high resistance to the prevailing GB biotypes (Lu et al., 2010; Tan et al., 2017).

3.2 Objectives

1. Evaluate the T4I mapping population for HF and GB resistance under controlled conditions.
2. Use the obtained phenotypic data to conduct QTL analysis for both traits.
3. Identify high confidence genes within the QTL flanking regions.

3.3 Materials and Methods

The wheat breeding program at Texas A&M has developed a bi-parental population from the cross between two elite varieties, TAM 204 and Iba. The maternal parent is TAM 204, a high yielding and drought-tolerant variety recommended for Texas plains that possess resistance to WSMV, Green Bug and Hessian Fly (Rudd et al., 2019). The paternal parent is Iba, a high yielding variety recommended for Kansas and Texas Panhandle with an outstanding SBMV resistance (Edwards, 2013). The population consisting of 221 F7 RILs, as well as parental lines, was screened for the response to infestation for GB in Stillwater OK, whereas the HF evaluation was performed in Manhattan KS. Both evaluations were done at the USDA/ARS labs at the above two locations.

Both of the evaluations were performed as described by Tan et al., 2017. Briefly, for GB, ten plants per line were grown and infested at the two-leaf stage and the response was scored 14 days after infestation. Plants were rated as resistant or susceptible and the percentage of resistant plants per line was recorded. Whereas for HF, mated females were used to infest the 20 plants per line at one-leaf stage and the rating was performed three weeks after infestation. Susceptible plants showed stunted phenotype and harbored live larvae, whereas resistant ones grew naturally with light green color and dead larvae. In both evaluations, the lines were scored based on the percentage of resistant lines. Histogram visualization of each evaluation was performed in R studio (RStudio Team, 2015) using ggplot2 in order to see the phenotypic distribution of the lines. Lines were considered resistant when 0-20% of the plants were susceptible and susceptible 80-100 % of lines were suspectible. Lines with a percentage of resistant plant within that range were considered heterogeneous. A Chi-square test was performed check the segregation ratio model of one gene (1R:1S), only considering susceptible and resistant lines. QTL mapping was performed using the BIP function in QTL IciMapping software version 4.1 (Meng et al., 2015) for all the traits. The parameters used were: deletion of missing phenotypes, Inclusive Composite Interval Mapping of Additive and Dominant QTL (ICIM-ADD) method with the 1000 permutation test to set the threshold, type I error of 0.05, step of 1 cM and PIN of 0.001.

3.4 Results and Discussion

In the HF resistance evaluation, it was possible to see a wide moderate phenotypic variation amongst the lines (Supplemental figure 1), although most of them tended to have either 100 or zero % of plants resistant to the GP biotype. Concerning the parental lines, TAM 204 exhibited a resistant phenotype, whereas Iba showed susceptibility symptoms. For the RIL population, 103 lines (47.76 %) showed a resistant phenotype and 53 (26.26 %) were susceptible. The remaining

57 lines (26.26 %) showed partial resistance. When evaluating the resistance inheritance in the population using a Chi-squared test considering either lines as resistant (>80 %) or susceptible (<20%), the phenotypic variation deviated ($\chi^2$1:1 = 19.55, df = 2, P-value = < .00001) from the expected 1:1 segregation ratio for a major gene. This explains the substantial number of lines that have plants with a partial resistant phenotype.

Concerning the QTL analysis, two QTL inherited in the population from TAM 204 were identified in chromosomes 1A and 3B (Table 3.1). According to the sequence alignment in the Chinese Spring reference genome, the QTL (*Qhf.tamu.1A.8*) in 1A was detected within the 8.29-8.59 Mb region. This QTL with a LOD score of 9.83, explained 12.69% of the PVE. The second QTL (*Qhf.tamu.3B.1*) was detected in 3B in a region of 0.37-0.67 Mb and with a LOD score of 11.64. This QTL explained a slightly higher phenotypic variation of 15.30 %.

In the IWGSC, six different annotated genes with high confidence protein-coding were found in the region where the QTL *Qhf.tamu.1A.8* is located. The TraesCS1A02G01560, TraesCS1A02G015700, and TraesCS1A02G015800 have multiple domains and features; their transcripts belong to the NADH oxidoreductase-related family according to the PANTHER (protein analysis through evolutionary relationships) gene classification system. Nevertheless, no phenotype is associated with them. The other three TraesCS1A02G015900, TraesCS1A02G016000 and TraesCS1A02G016100, also have multiple domains and features, and their transcripts belong to the serine/threonine-protein kinase subfamily. Amongst all these candidate genes, the one with more information is the TraesCS1A02G016000, and this gene belongs to the rust resistance kinase *Lr-10* related family. The rust resistance gene *Lr-10* has been mapped in the 1AS and it encodes a protein that has a coiled-coil domain (CC), a central nucleotide-binding site (NB), and a C-terminal leucine-rich repeat (LRR)(Feuillet et al., 2003).

Table 3.1 Quantitative trait loci (QTL) for Hessian Fly and Greenbug resistance detected in the TAM 204/Iba population.

QTL Name	Chrom	Position (Mb)	Threshold LOD	LOD	PVE	Additive effects	Parental favorable alleles	25LGs	Peak position cM	Left SNPs	Left SNPs alleles	Left SNPs TAM 204	Left SNPs IBA	Left SNPs chr_bp	Right SNPs	Right SNPs alleles	Right SNPs TAM 204	Right SNPs Iba	Right SNPs chr_bp
Qhf.tamu.1A.8	1A	8.29	3.51	9.8	12.69	15.55	TAM 204	1	7.25	chr1A_8295595	A/G	A	G	8295595	chr1A_8595496	T/C	T	C	8595496
Qhf.tamu.3B.1	3B	0.37	3.51	11.64	15.30	17.09	TAM 204	12	0	chr3B_374873	T/C	T	C	374873	chr3B_675212	T/G	T	G	675212
Qgb.tamu.7D.596	7D	596.68	3.51	108.3	88.24	44.88	TAM 204	60	128.87	chr7D_596683352	G/A	G	A	596683352	chr7D_596853157	C/T	C	T	596853157

Chrom = chromosome; LOD= logarithm of the odds; PVE= Phenotypic Variation Explained; LG= Linkage Groups; cM = centiMorgan; SNP = Single Nucleotide Polymorphism; chr_bp= physical position in base pairs.

The NB-LRR class of proteins are normally encoded when the avirulent genes are recognized by their counterpart R-genes in the plant in a typical interaction between the plant and pathogen (Juliana et al., 2018) Multiple HF resistance genes (*H9*, *H10*, *H11*, *Hdic*) have been found clustered in the 1AS(Liu et al., 2014; Li et al., 2015); and the gene *H9* is known to be linked to the *Lr-10* and the Powdery Mildew (*Blumeria graminis* f. sp. *tritici* (*Bgt*)) resistance gene *Pm3*. This finding indicates that the QTL *Qhf.tamu.1A.8* is neighboring the correct genomic region that will reveal the causal gene that confers HF resistance.

Contrasting to chromosome 1A where multiple resistance genes have been found, only the gene *H35* has been postulated to reside in chromosome 3B (Zhao et al., 2020). Zhao et al. (2020) postulated the *H35* gene in the 0.2-4.8 Mb region, where we found *Qhf.tamu.3B.1*. In their study, Zhao et al. (2020) confirmed the presence of the *H35* gene in the breeding line SD06165 (Wesley/SD97049) by saturating the region with Simple-Sequence Repeats (SSR) and Kompetitive Allele-Specific (KASP) (generated from GBS) markers on 3BS. Nevertheless, even though they identified multiple high confidence genes, none were associated with a typical disease resistance gene. Similarly, with the other HF resistance genes, *H35* is not associated with a gene in the IWGSC reference genome. In our results, within the *Qhf.tamu.3B.1* flanking markers interval, multiple high confidence genes were found, and, therefore, another strategy is needed in order to continue with fine-mapping *H35*.

In the Greenbug evaluation, the lines also showed variation in the symptomatic reaction to Biotype E (Supplemental Figure 2). Similar to the HF screening, TAM 204 exhibited a resistant phenotype, whereas Iba showed susceptibility symptoms. For the RIL population, 103 lines (41.55 %) showed resistant phenotype, and 93 (42.46 %) were susceptible. The remaining 35 lines (15.9%) were heterogeneous and had ranges of 20-80 % resistant plants. The Chi-square

test results of the resistance inheritance in the population fits ($\chi^2$1:1 = 0.125, df = 2, P-value = 0.72301) the expected 1:1 segregation ratio which is expected for a major gene.

In this analysis, a major QTL (*Qgb.tamu.7D.596*) was found in chromosome 7D (596.68-596.85 Mb) and explained up to 88.24 % of phenotypic variation (Table 3.1). The *Gb3* gene is known to be present in Texas A&M wheat breeding germplasm and it was identified in TAM 112, one of the parental lines of TAM 204. Azhaguvel et al. (2012) fine mapped this gene in *Aegilops tauschii* in the long arm of chromosome 7D, and narrowed it down to an interval between the two markers (HI067J6-R and HI009B3-R) of 1.1 cM. The forward marker primer for HI067J6-R overlaps the physical position 596.30 Mb of the reference genome, whereas the forward primer HI009B3-R overlaps the position 598.35 Mb. Azhaguvel et al., (2012) also hypothesized that the *Gb3* could belong to the NB-LRR R-genes, given that the cloned aphid resistance genes are from that family. No annotated genes are located within the 185kb QTL region in the Chinese Spring reference genome.

For both HF and GB mapping, the second replication of screening will permit us to consolidate our findings. Nevertheless, previous studies have found similar QTL to the ones identified in our study. The subsequent step is to convert the SNP markers linked to the QTL to KASP-based markers which will eventually be used for diagnostic assays using MAS in the Texas A&M breeding program.

3.5 Conclusion

This study successfully mapped two QTL for HF and one with GB linked to resistance genes. Two of them have been characterized as *H35* and *Gb3* and one located in 1A could be associated with *H9*. Further studies, namely KASP-based validation and fine mapping, will be needed to narrow down the distance between flanking markers. These strategies will permit the association

of the found QTL with candidate genes in the reference genome, especially for *H35*, given multiple high confidence genes were found in the QTL interval.

CHAPTER IV

QTL MAPPING FOR AGRONOMICAL TRAITS

4.1 Introduction

Grain yield is the most important trait to breed for in an agronomical crop, and, therefore, the majorly of the breeding efforts have been invested in it. Nevertheless, it is very complex given it is modulated by multiple genetic and environmental factors, as well as the genotype-by-environment interaction (GEI). For instance, environmental conditions, namely drought and heat, during the late season would favor early genotypes since they would avoid the stress and their growing cycle would be completed as normal; whereas a later heading genotype would be affected due to abortion of developing florets or lack of water during grain filling stage (Shokat et al., 2020). Therefore, it is important to understand the GEI by doing envirotyping (Xu, 2016) and delimited Mega-Environments (ME) by clustering environments with similar geographical and/or climatological conditions (Braun et al., 1996).

Breeding for grain yield per se has certain limitations given its nature and a large number of genes and gene networks that regulate the expression of several pathways that affect grain development and production (Kuzay et al., 2019). It has been hypothesized that increments in grain yield can be achieved by increasing yield components (Ferrante et al., 2017), and certain avenues as breeding approaches can be used. Grain yield can be partitioned into two major components, namely grain weight and the number of grains per m^{-2} (GPM); however, these two are generally negatively correlated (Slafer et al., 2014). GPM can further be divided into the number of spikes per m^{-2} and the number of grains per spike and their dynamics and impact on the final grain yield change throughout the growing season. GPM and its subcomponents are determined early in the growing season; whereas grain weight is determined after anthesis

(Slafer et al., 2014). It is commonly known that grain weight manifests less phenotypic plasticity under changing environments compared to GPM, being a consequence of a paradox of an evolutionary process (Sadras and Slafer, 2011). Grain size is prioritized as a part of an adaptive process that aims to conserve the offspring as its survival depends on the reserves that are stored in the embryo (Sadras, 2007). This trade-off has been minimized with the current agricultural practices permitting to exploit both components to increase grain yield. A diverse number of mapping populations have been developed to scrutinize markers linked with QTL related to these yield components (Sukumaran et al., 2014; Assanga et al., 2017; Kuzay et al., 2019; Voss-Fels et al., 2019; Kuang et al., 2020). Grain weight is a suitable trait for indirect selection for grain yield as its phenotypic evaluation via Thousand Kernel Weight (TKW) could be easily performed compared to GPM. This strategy has shown to be effective at CIMMYT, leading to increases in grain yield and grain size in Northwest Mexico (Lopes et al., 2012). Multiple genomic regions have been associated with TKW and the genes *TaSus2-2B*, *TaGS-D1*, *TaCKX-D*, *TaGASR-A1*, *TaCwi-4A*, *TaCwi-5D*, *TaMoc-7A*, and *TaTGW6-A1* have been identified (Rasheed et al., 2016; Lozada et al., 2018; Jamil et al., 2019; Li et al., 2019). QTL for the number of spikes, spikelet per spike, and grain number has been identified in multiple populations; similarly, genes for these traits, namely *TaMoc1-7A*, *TaCwi-4A* and *WAPO-A1*, respectively, have been characterized (Jamil et al., 2019; Kuzay et al., 2019; Li et al., 2019; Lozada et al., 2018; Rasheed et al., 2016). The introgression of dwarfing genes in elite cultivars was a major shift in wheat breeding back in the 1960s, in which the change in the plant archetype led to higher grain yields as a consequence of increments in harvest index and straw, as a product of a higher tiller production and reduced lodging under fertilized conditions (Borlaug, 1968). Multiple dwarfing/reduced height (Rht) genes, majorly associated with gibberellic acid insensitivity, have been characterized (Mo et al.,

2018). The dwarfing genes *Rht1* (*Rht-B1b*) and *Rht2* (*Rht2-D1b*) are the most widely deployed worldwide due to their advantageous characteristic under high yielding conditions; however, they are associated with poor emergence and short coleoptiles (Rebetzke et al., 2007). As an alternative, other genes such as Rht8, which does not affect the coleoptile characteristics, are used under deep planting conditions (Rebetzke et al., 2007; Mo et al., 2018).

The phenology in winter wheat is mostly conditioned by the vernalization genes, the homeologous group 5 chromosome *Vrn-A1*, *Vrn-B1* and *Vrn-D1* as well as *Vrn-D3* in 7D (Kato et al., 1998; Liu et al., 2014). Some other genomic regions with a similar role in plant development have been identified although they have not been characterized (Sukumaran et al., 2014; Ogbonnaya et al., 2017; Jamil et al., 2019). These genes have a significant impact on days to heading and maturity as well as a pleiotropic effect on important agronomic traits such as TKW and grain yield.

4.2 Materials and Methods

4.2.1 Plant Material and Field Experiment Design

The RILs, along with the parental lines, TAM 204 and Iba, as well as other Texas A&M varieties, TAM112, TAM114 and TAM205 were grown in the environments listed in Table 3.1. Not all the varieties were grown in all environments. Also, due to the limited number of seeds, not all the lines were planted in all environments. The experiments were laid out in an alpha-lattice design (0, 1) with an incomplete block size. Most of the experiments had two replications, excluding McGregor 2020 and Prosper 2020, which only had one replication. The plot size varied depending on the irrigation condition. For full irrigation, the size was 10x5 ft. (3.0 x 1.5 m), whereas it was 15x5 ft. (4.5 x1.5 m) in the dryland.

Table 4.1. Summary of the trials established in the state of Texas for testing the T4I population.

Year	Location name	Location code	Mega-environment	Water Level	Rep	No. plots	No. Lines
18	College Station	18CS	1	dryland	2	196	96
18	McGregor	18MCG	2	dryland	2	154	75
18	Dumas	18DMS	4	Irrigated	2	270	133
19	College Station	19CS	1	dryland	2	448	211
19	McGregor	19MCG	2	dryland	2	448	214
19	Bushland	19BI	4	Irrigated	2	448	221
19	Emeny Land	19EMN	2	dryland	2	448	205
19	Chillicothe	19CH	2	dryland	2	448	218
20	College Station	20CS	1	dryland	2	448	221
20	McGregor	20MCG	2	dryland	1	224	221
20	Prosper	20PRO	2	dryland	1	224	221
20	Emeny Land	20EMN	2	dryland	2	448	221

(ME1=South Texas, ME2=Rolling Plains, ME3=High Plains dryland, ME4=HP irrigated).

4.2.2 Phenotypic Data Collection

Agronomic data collected depended on the location due to logistical reasons (Table 4.2). Grain yield (GYLD) was collected using a Wintersteiger combine harvester. Days to heading (DTH) was recorded as the number of days from January 1st when more than 50 % of the plants were a at the 10.1 Feekes' scale. Plant Height (PH) was taken at physiological maturity (11 Feekes' scale) from the base of the plant to the tip of the spike. Thousand Kernel Weight (TKW) was measured by counting 300 grains and extrapolating it to 1000 for both replications of MCG18. In the case of MCG19, samples from one replication were analyzed using the Single Kernel Characterization System (SKCS) 4100 (Perten Instruments North America Inc.), which

calculated a single grain weight based on the average of 300 kernels per sample. Test weight (TW) was measured using the Seedburo equipment (www.seedburo.com, Des Plaines, IL, USA).

Table 4.2. Year, location, location code and traits evaluated in the detected in the TAM 204/Iba population.

Year	Location name	Location code	GYLD	DTH	PH	TW	TKW
2018	College Station	18CS	✕	✓	✕	✕	✕
2018	McGregor	18MCG	✓	✓	✓	✓	✓
2018	Dumas	18DMS	✓	✓	✓	✕	✕
2019	College Station	19CS	✓	✓	✓	✓	✕
2019	McGregor	19MCG	✓	✓	✓	✓	✓
2019	Bushland	19BI	✓	✓	✓	✕	✕
2019	Emeny Land	19EMN	✓	✕	✓	✓	✕
2019	Chillicothe	19CH	✓	✕	✕	✕	✕
2020	College Station	20CS	✓	✓	✓	✕	✕
2020	McGregor	20MCG	✓	✓	✓	✓	✓
2020	Prosper	20PRO	✕	✓	✕	✕	✕
2020	Emeny Land	20EMN	✓	✕	✕	✓	✕

GYLD = Grain yield; DTH = Days to heading; PH = Plant height; TW = Test weight; TKW = Thousand kernel weight, ✓= evaluated, ✕ = no evaluated.

4.2.3 Statistical Analysis

The statistical analyses were performed in the META-R program which is linked to an interface that uses R (Alvarado et al., 2020). The REML produced was used in the lme4 package to calculate variance components and broad-sense heritability for the individual environments (Eq. 4.1 and 4.2) and combined analysis (Eq. 4.3 and 4.4) considering all factors as random effects.

$$\text{(Equation 4.1) } Y_{ijk} = \mu + R_i + B_{j(Ri)} + G_k + \varepsilon_{ijk}$$

Where y_{ijk} is the trait of interest; μ the mean effect; R_k is the effect of the i^{th} replicate; $B_{j(Rk)}$ is the effect of the j^{th} block within the i^{th} replicate; G_i is the genetic effect of the k^{th} genotype; and ε_{ijk} is the residual (Alvarado et al., 2020).

$$\text{(Equation 4.2) } H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_\varepsilon^2}{nReps}}$$

Where H^2 is broad-sense heritability, σ_g^2 and σ_ε^2 are genotype and error variances, respectively, and nReps is the number of replications.

$$\text{(Equation 4.3) } Y_{ijk} = \mu + E_i + R_{j_{(Ei)}} + B_{j(EiRj)} + G_l + GxE_{il} + \varepsilon_{ijkl}$$

Where E_i are the effects of the i^{th} environment and GxE_{il} are the genotype-by-environment interaction. The other factors are similar to Equation 4.1.

$$\text{(Equation 4.4) } H^2 = \frac{\sigma_g^2}{\sigma_g^2 + + \frac{\sigma_{ge}^2}{nEnvs} + \frac{\sigma_\varepsilon^2}{nEnvs\ x\ nReps}}$$

Where σ_{ge}^2 is the genotype by environment variance, $nEnvs$ is the number of environments. The other factors are similar to Equation 4.2.

Phenotypic correlations between traits were estimated using Pearson correlations. Best Linear Unbiased Estimators (BLUEs) were estimated across all the locations, assuming genotypes as fixed effects. Biplots were generated using the two main principal components of the phenotypic variance in order to classify Mega-Environments (ME) for each trait (Yan and Tinker, 2006). Environments clustered in the same quadrant with an acute angle between their vectors were

assigned to the same ME. QTL analyses were performed using the BLUEs of all the traits collected by the environment, trait across environments, and trait across ME. The GYLD and PH data from 20MCG were not included in the statistical analysis given they were unreplicated but the obtained BLUE values were used for the QTL analysis. QTL mapping was performed using the BIP function in QTL IciMapping software version 4.1 (Meng et al., 2015) for all the traits. The parameters used were: deletion of missing phenotypes, Inclusive Composite Interval Mapping of Additive and Dominant QTL (ICIM-ADD) method with the 1000 permutation test to set the threshold, type I error of 0.05, step of 1 cM and PIN of 0.001. The epistatis effects were calculated with the function inclusive composite interval mapping of epistatic QTL (ICIM-EPI).The threshold in the ICIM-EPI was te same as the one determined in the 1000 permutation test of the ICIM-ADD method, or a LOD = 10 if too many interaction existed.

Identified QTL were designated as *Qtrait.tamu.chrom.Mb*, where *trait* is an abbreviation of the trait, *tamu* is the acronym for Texas A&M University, *chrom* is the chromosome where the QTL was found, and Mb is the physical position of the left marker linked to the QTL according to the alignment with the IWGSC RefSeq v1.0 reference genome (Appels et al., 2018).

4.3 Results and Discussion

4.3.1 Analysis of variance and heritability

The combined analysis of variance showed highly significant differences ($P < 0.001$) among genotypes, environments and genotype-by-environment interaction for all traits (Table 4.3). The large magnitude of the environment's variation is not desirable in genetic studies given it can affect the results obtained when performing the QTL analysis. Also, the significant genotype-by-environment interaction had a differential response to the environmental changes for all traits. It is important to point out that the genetic variation was higher than the variation explained by the

genotype by environment component. The coefficient of determination (R^2) was high for all the traits and ranged from 76.62 to 99.27 for PH and DTH, respectively. Similarly, high heritability estimates were obtained, ranging from 0.69 to 0.93. Other studies have found similar heritability estimates for these traits (Assanga et al., 2017; Gao et al., 2015; Yang et al., 2020). GYLD is expected to have lower heritability estimates compared to TKW given the complexity of the trait; however, previous studies have found similar estimates (Jamil et al., 2019; Kuang et al., 2020). The low coefficient of variation for all traits suggests a good model fit and data reliability. The average GYLD pooled across environments was 316.28 g m^{-2}. The population average DTH and PH were 102.66 days and 84.52 cm, respectively. TKW had an average of 28.68 grams and TW 733.63 Kg m^{-3}. Regarding comparing both population parents, Iba had higher BLUEs values for all traits, except for DTH, where the variation difference between them was one day.

Table 4.3. The combined variance component and heritability estimates, assuming all sources of variation as random, and mean performance for all traits across environments.

		Variance Components									BLUEs	
Traits[a]	Units	Gen[b]	Env[c]	Gen x Env[d]	Residual	R^{2e}	Heritability	Trait Mean	LSD[g]	CV[g]	TAM 204	Iba
GYLD	g m^{-2}	1303.85***	21448.8***	2380.08***	885.71	96.68	0.82	316.28	52.52	9.38	305.2	356.9
DTH	days	44.55***	397.97***	16.46***	3.56	99.27	0.85	102.66	4.95	1.83	101	100
PH	cm	22.53***	22.88***	4.08***	15.78	76.62	0.93	84.52	4.04	4.7	83.3	86
TKW	g	7.50***	28.49***	5.69***	1.64	96.28	0.69	28.68	7.08	4.42	24.7	32.3
TW	Kg m^{-3}	411.33***	1531.36***	298.58***	50.81	97.82	0.88	733.63	25.95	0.96	707.3	757.1

[a]Abbreviation of traits: GYLD grain yield form combine harvest of plot, DTH Days to heading, PH Plant height, TKW Thousand kernel weight, TW Test Weight,[b]Genotype Variance, [c]Environmental variance, [d]Genotype by Environment Variance, [e]R^2 Coefficient of determination (%), [f]Least significant difference ($\alpha = 0.05$), [g]coefficient of variation. **, *** significant at 0.01, and 0.001 probability levels, respectively, and ns not significant.

The GYLD performance of the population showed high phenotypic plasticity across all the tested locations. Figure 4.1 shows that the environments that belong to the same ME, shown in Table 4.1, had superior yields to environments from ME1 in the previous classification. The higher

yielder environment was 19BI, whereas the lower one was 20CS, with a grand mean of 569 and 119.9 g m^{-2}, respectively. It is essential to point out that the higher BLUE value is from the line 133 in 19BI with an outstanding yield of 693.52 g m-2, 50 g m^{-2} more compared to Iba, the higher yielder parent (Supplemental Table 4).

The reason underlying the differences in the lines' performance across environments, leaving aside environmental factors such as temperature or soil type, was the irrigation management. However, this is irrelevant, given the modest performance observed in 18MCG even though the experiment was managed under rainfed conditions. This factor is also seen when looking at the differences in performance from year to another since the rain pattern is not consistent year to year. Moreover, 20CS was the only environment in which some lines failed to vernalize, given in February and March 2020, the temperatures were considerably higher than the historical average. The lines that did not vernalize were taken out of the analysis in order to remove that bias.

Figure 4.2 shows the biplot constructed with the first two principal components, which explain 89.32 % of the phenotypic variation. This figure also shows that the environmental mean and phenotypic variation was different across environments. The biplot classified clustered environments with similar geographical location and irrigation conditions. The MEs for GYLD (according to their position and named clockwise) are ME1 (19CH, 19MCG, 20CS, 19CS, and 18MCG); ME2 (18DMS and 20EMN) and ME3 (19EMN and 19BI).

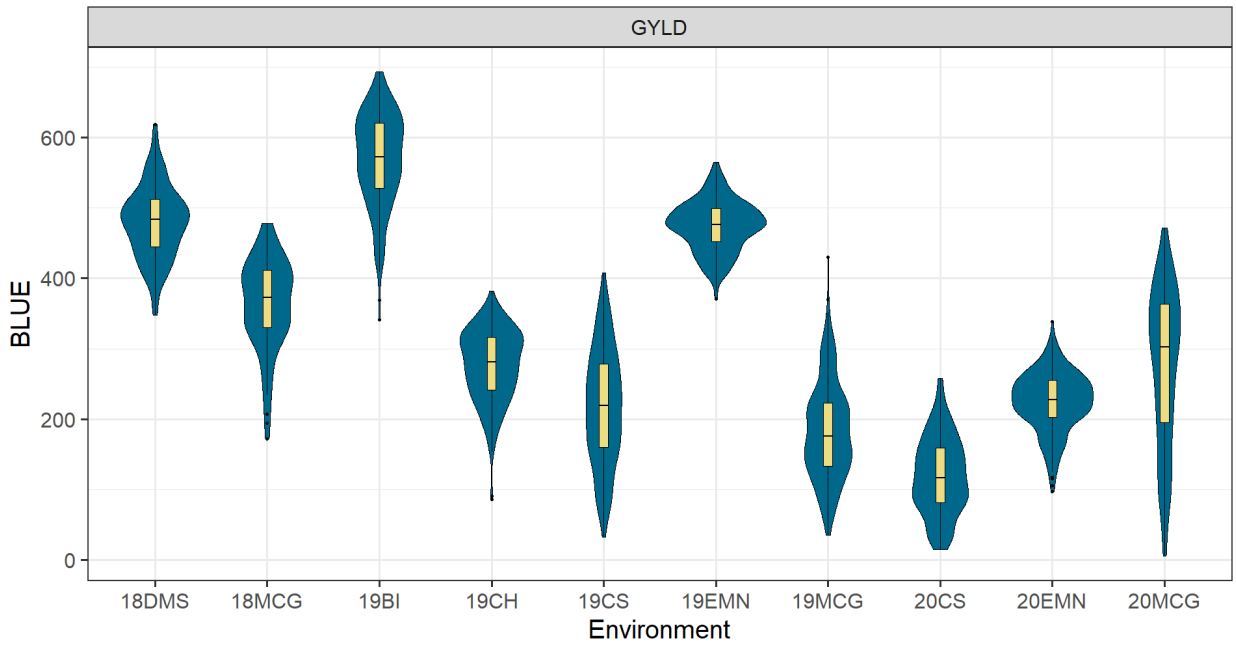


Figure 4.1. Phenotypic distribution of the Best Linear Unbiased Estimators (BLUE) values for grain yield (GYLD) in all evaluated environments. 18DMS, Dumas 2018; 18MCG, McGregor 2018; 19BI, Bushland Irrigated 2019; 19CH, Chillicothe 2019; 19CS, College Station 2019; 19EMN, Emeny Land 2019; 19MCG, McGregor 2019; 20CS, College Station 2020; 20EMN, Emeny Land 2020; 20.MCG, McGregor 2020.

The phenotypic variation for DTH depended on the environment. Figure 4.3 shows two marked patterns for this trait. In the first one, Central Texas environments showed different days to heading ranging from 79 to 104 for 20CS and 20MCG, respectively. Furthermore, the days to heading trend reflects the vernalization issues at 20CS. It is well known that heading is highly modulated by the environment and increments in the temperature at the booting stage accelerate the plant development. As a consequence, some lines started heading earlier compared to other environments and even compared to other years at this location.

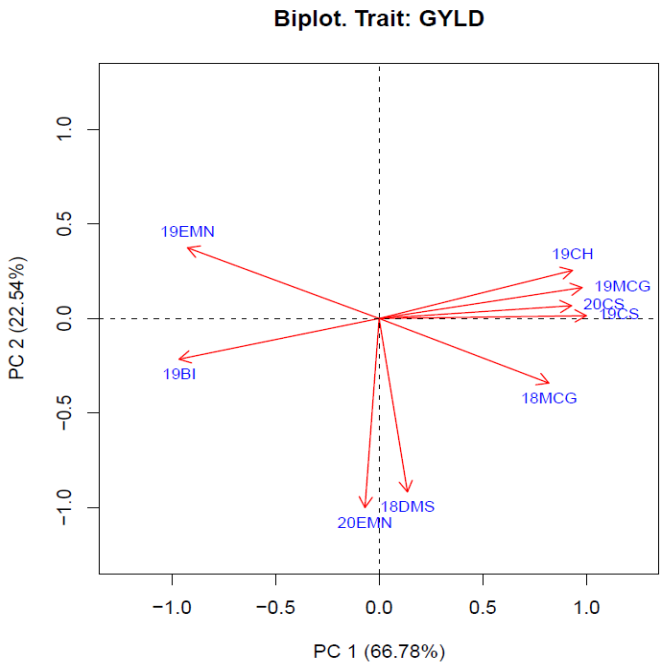


Figure 4.2. Biplot for mega-environments classification using the two principal components according to the grain yield (GYLD) performance of the T4I population lines across the tested environments. 18DMS, Dumas 2018; 18MCG, McGregor 2018; 19BI, Bushland Irrigated 2019; 19CH, Chillicothe 2019; 19CS, College Station 2019; 19EMN, Emeny Land 2019; 19MCG, McGregor 2019; 20CS, College Station 2020; 20EMN, Emeny Land 2020; 20.MCG, McGregor 2020

On the other hand, lines inthe North Texas environments (18DMS and 19BI) were later (>125 days after planting) and with a smaller range when compared to the other environments. These two environments are also considerably colder, and it snows during the growing season; a factor that extends the growing cycle due to slowing down plant development. In the biplot to group ME, based on the performance of the lines across these environments (Figure 4.4), there is an apparent year effect, where 18CS and 18MCG are grouped into one ME (ME1), 18DMS and 19BI in ME2. Relatively close to them is 19MCG, although it was grouped in ME3 along with 20MCG, 19CS and 20CS, given the acute angles between them and close geographical proximity.

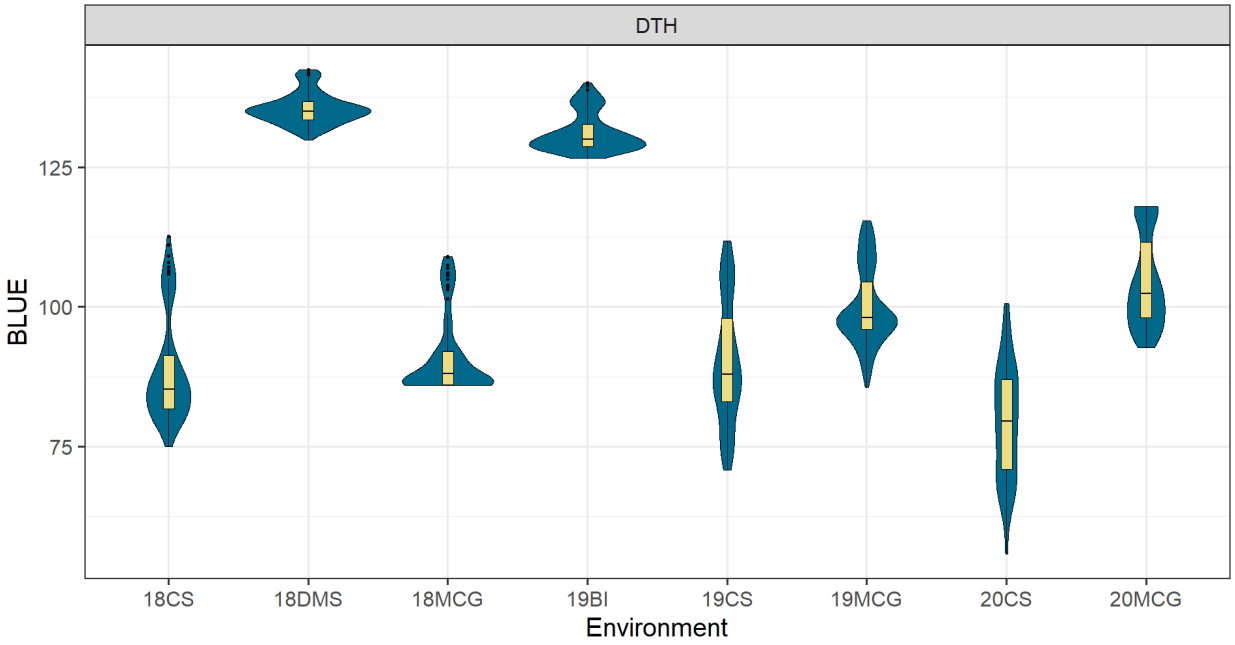


Figure 4.3. Phenotypic distribution of the Best Linear Unbiased Estimators (BLUE) values for days to heading (DTH) in all evaluated environments. 18CS, College Station 2018; 18DMS, Dumas 2018; 18MCG, McGregor 2018; 19BI, Bushland Irrigated 2019; 19CS, College Station 2019; 19MCG, McGregor 2019; 20CS, College Station 2020; 20.MCG, McGregor 2020.

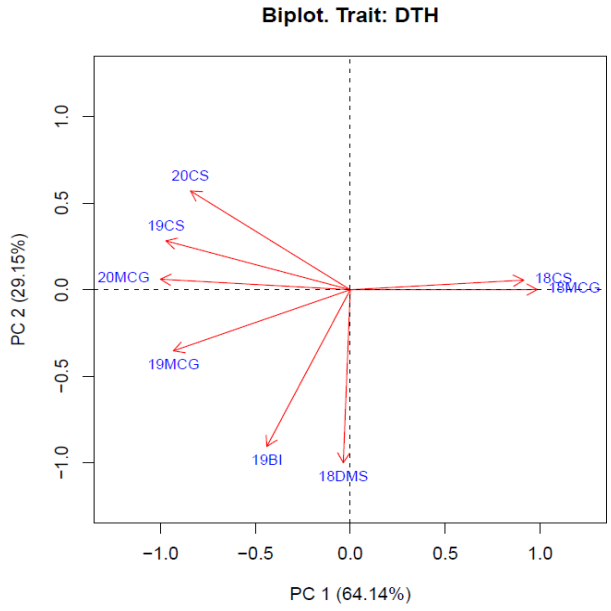


Figure 4.4. Biplot for mega-environments classification using the two principal components according to the days to heading (DHT) performance of the T4I population lines across the tested environments. 18CS, College Station 2018; 18DMS, Dumas 2018; 18MCG, McGregor 2018; 19BI, Bushland Irrigated 2019; 19CS, College Station 2019; 19MCG, McGregor 2019; 20CS, College Station 2020; 20.MCG, McGregor 2020.

Among all the traits evaluated, PH's phenotypic distribution was the most consistent across environments (Figure 4.5), although the BLUE mean values varied from 80.26 to 93.57 cm, for 20CS and 19BI, respectively. This reflects the effect of weather conditions and irrigation level on the plant phenotype, given 19BI and 20CS also had higher and lower values with 111.2 and 60 cm, respectively. The biplot in Figure 4.6b shows two ME, the ME2 comprising 19EMN and 19BI and the ME3 which includes 19MCG, 19CS and 20CS. For the ME1, despite 18MCG and 18DMS not being located in the same quadrant of the biplot, they were considered as one ME given they were clustered together in the Ward dendrogram (Figure 4.6a).

Phenotypic variation was observed for both TKW and TW (Figures 4.7a and 4.7b), and similarly, a different genotype-by-year effect was manifested in MCG and EMN locations, where the BLUE values obtained in the first year of evaluation are higher than the second year. Nevertheless, their phenotypic distribution was similar. 19MCG was the environment in which the broader range of phenotypic variation was observed for TW, with a difference of 198 kg m^{-3} between the higher and the lower BLUE values. In TW's case in 20EMN, observations with low values (>3sd from the mean) negatively skewed the box plot, but these observations were kept because they belonged to the same lines.

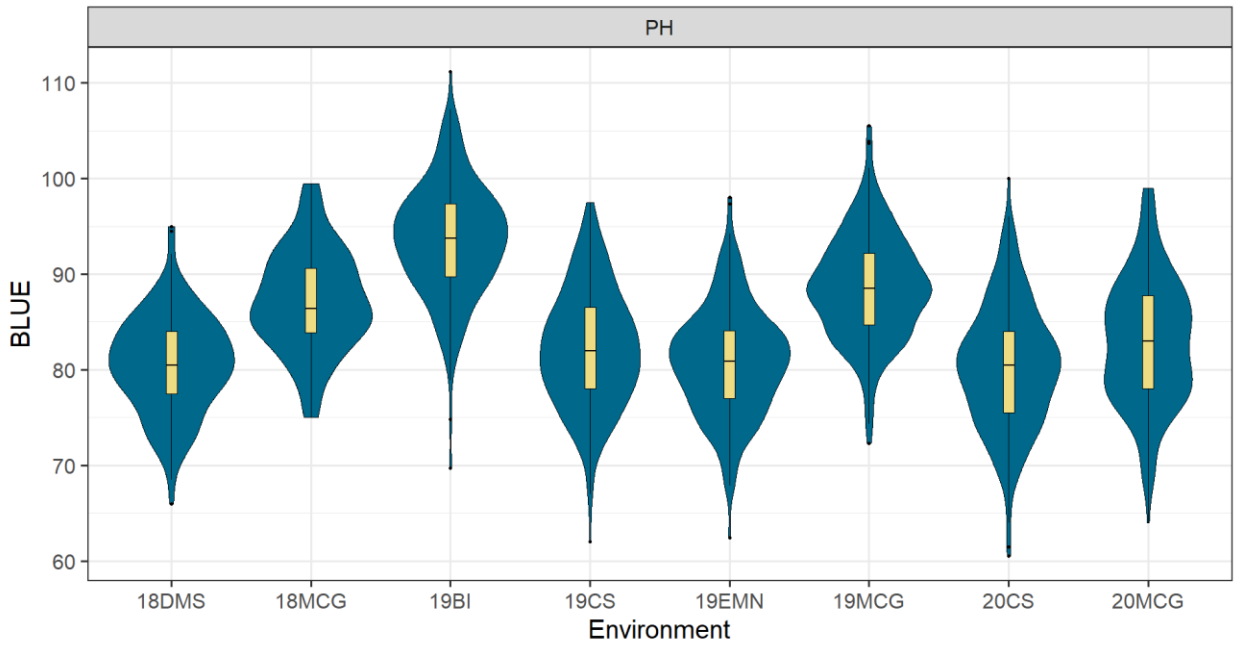


Figure 4.5. Phenotypic distribution of the Best Linear Unbiased Estimators (BLUE) values for plant height (PH) in all evaluated environments. 18DMS, Dumas 2018; 18MCG, McGregor 2018; 19BI, Bushland Irrigated 2019; 19CS, College Station 2019; 19EMN, Emeny Land 2019; 19MCG, McGregor 2019; 20CS, College Station 2020; 20.MCG, McGregor 2020.

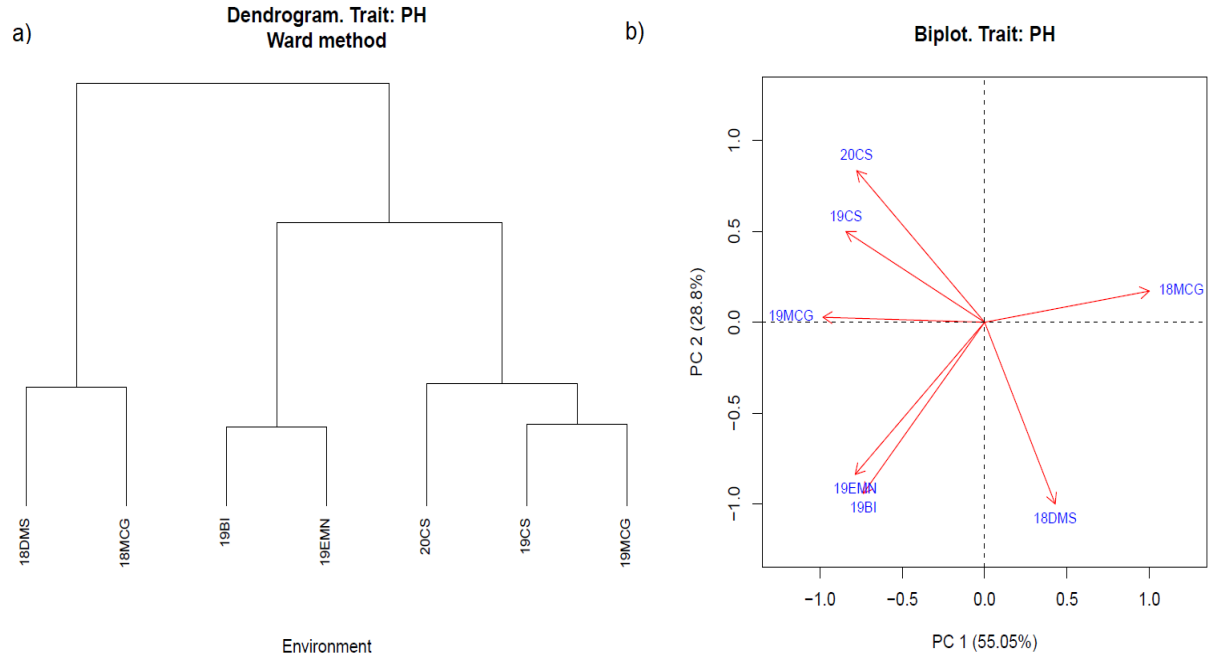


Figure 4.6. Dendrogram (a) and Biplot (b) for mega-environments classification using the two principal components according to plant height (PH) performance of the T4I population lines. 18DMS, Dumas 2018; 18MCG, McGregor 2018; 19BI, Bushland Irrigated 2019; 19CS, College Station 2019; 19EMN, Emeny Land 2019; 19MCG, McGregor 2019; 20CS, College Station 2020; 20.MCG, McGregor 2020.

For TW, two ME were obtained (Figures 4.8a and 4.8b). The ME1 comprises of 18MCG, 19CS and 19MCG, whereas ME2 includes 19EMN and 20EMN. Although the biplot showed 18MCG and 19MCG were not positively correlated, both environments were grouped as ME1 to utilize all the environments. However, there is some relationship between them according to the dendrogram (Figure 4.8a).

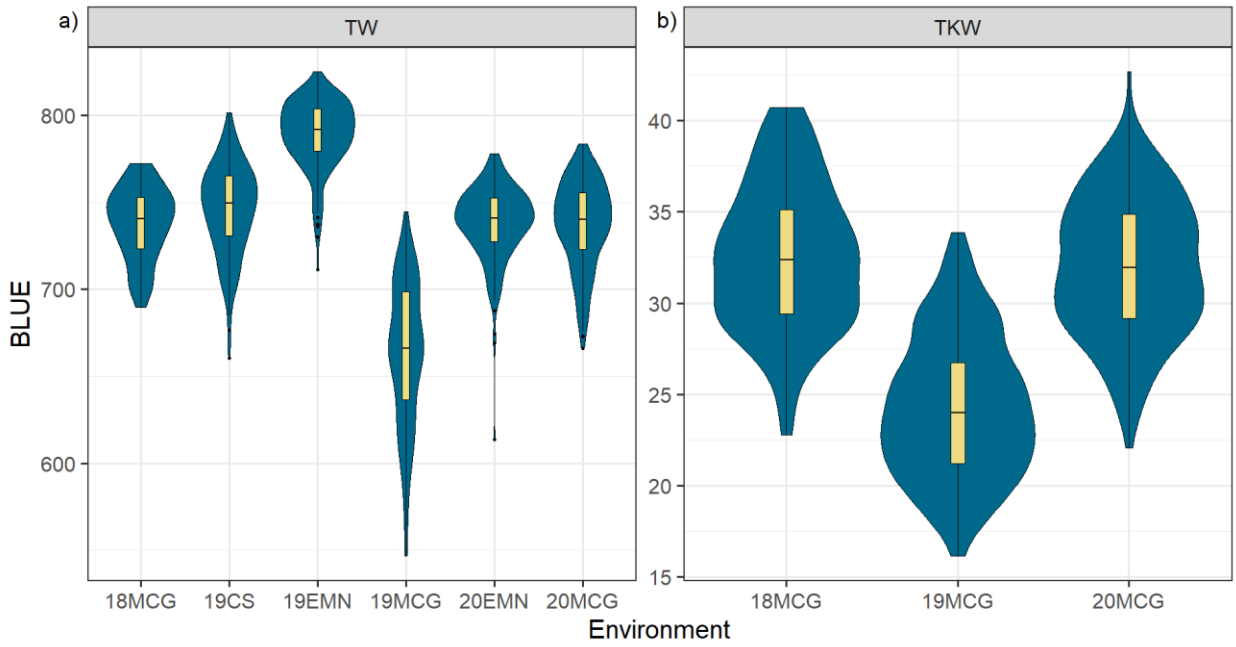


Figure 4.7. Phenotypic distribution of the BLUE values for test weight (TW) (a) thousand kernel weight (TKW) (b) in all evaluated environments. 18MCG, McGregor 2018; 19CS, College Station 2019; 19EMN, Emeny Land 2019; 19MCG, McGregor 2019; 20EMN, Emeny Land 2020; 20.MCG, McGregor 2020.

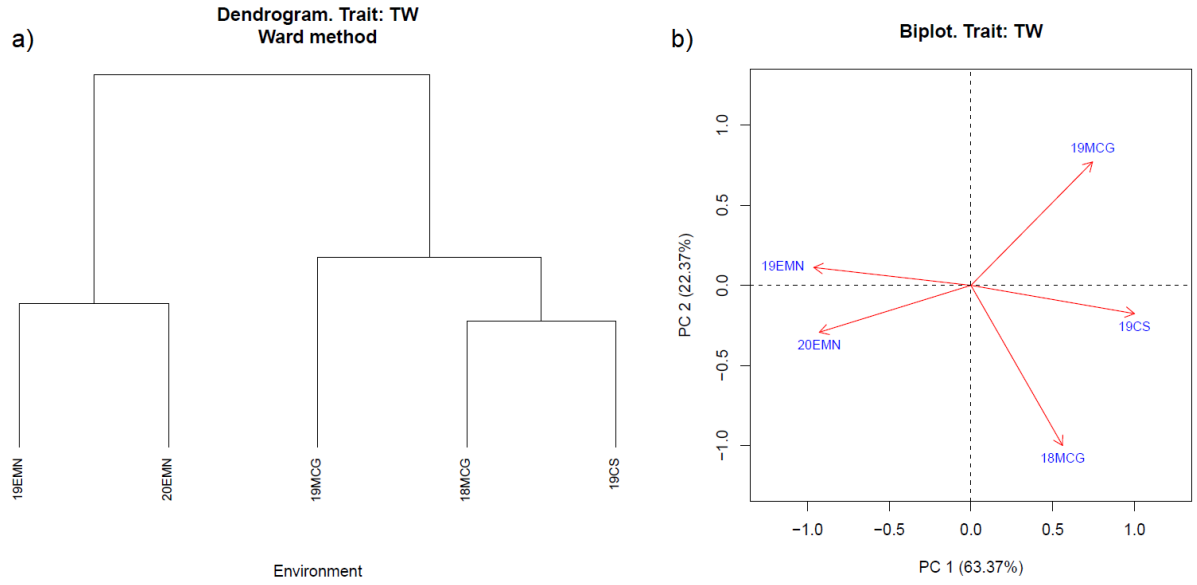


Figure 4.8. Dendrogram (a) and Biplot (b) for mega-environments classification using the two principal components according to test weight (TW) performance of the T4I population lines. 18MCG, McGregor 2018; 19CS, College Station 2019; 19EMN, Emeny Land 2019; 19MCG, McGregor 2019; 20EMN, Emeny Land 2020.

4.3.2 Phenotypic correlation between traits

The magnitude and direction of the correlation between grain yield, yield components and agronomic traits were varied (Table 4.5). Mixed results were obtained for the correlations between PH and GYLD; positive correlations were observed in 19CS, 19MCG and 20CS, whereas the two traits were negatively correlated in both environments located in the Texas Panhandle. PH was not significantly correlated with GYLD in 18MCG and 19EMN. TKW and TW were positively correlated with GYLD across different tested environments, except for TW in 19EMN. Regarding the correlation among traits in the combined analysis (Table 4.6), increments in TKW and TW were positively correlated with GYLD. PH was positively correlated with TKW but not with TW. Lastly, both yield components were positively correlated.

Ogbonnaya et al., 2017, found that GYLD was negatively correlated with DTH under heat-stress conditions, whereas PH and TKW were positively correlated with GYLD. Although no stress conditions were simulated, we found similar correlations on Central Texas environments that are considerably drier and warmer, compared experiments under full irrigation conditions evaluated in the Texas Panhandle. Assanga et al., 2017, found similar correlations between GYLD, PH and TKW. These significant correlations within environments and combined analyses provide a good insight into the potential pleiotropic effects or linkage of genomic regions regulating these traits. This high correlation between TKW and TW with yield indicates that both traits could be used for indirect selection.

Table 4.4. Phenotypic correlations between grain yield, agronomic data and yield components by individual environment.

Enviroments / Traits[a]	18DMS	18MCG	19BI	19CS	19EMN	19MCG	20CS	20EMN
				GYLD				
DTH	-0.51**	-0.75***	0.03ns	-0.50***		-0.50***	-0.32***	
PH	-0.33***	0.09ns	-0.44***	0.21**	-0.04ns	0.31***	0.39***	
TKW		0.46***				0.70***		
TW		0.57***		0.57***	0.08ns	0.66***		0.45***

[a]Abbreviation of traits: GYLD grain yield form combine harvest of plot, DTH Days to heading, PH Plant height, TKW Thousand kernel weight, TW Test Weight. **, *** significant at 0.01, and 0.001 probability levels, respectively, and ns not significant.

Table 4.5. Phenotypic correlation among traits in the combined analysis.

Traits[a]	GYLD	DTH	PH	TKW
DTH	-0.576***			
PH	0.038 ns	0.163*		
TKW	0.581***	-0.521***	0.278***	
TW	0.598***	-0.541***	0.132ns	0.61***

[a]Abbreviation of traits: GYLD grain yield form combine harvest of plot, DTH Days to heading, PH Plant height, TKW Thousand kernel weight, TW Test Weight. **, *** significant at 0.01, and 0.001 probability levels, respectively, and ns not significant.

4.3.3 Quantitative trait loci for grain yield, yield components and agronomic data

A set of 86 unique genomic regions associated with GYLD, DTH, PH, TW and TKW were identified across 11 environments over three years when analyzing by individual and mega-environment. Among them, a set of 35 QTL was consistent, i.e., they were identified in at least two of the three analyses (individual environment, across environments, or mega-environments). A set of six pleiotropic QTL was found to be associated with at least two traits. Also, all the pleiotropic QTL were consistent for at least one trait. Also, all the pleiotropic QTL were consistent for at least one trait.

For GYLD, ten consistent QTL were located in chromosomes 2A, 2B, 3D, 5A and 6A (Table 4.6). Four major QTL, inherited from Iba, were detected at 29.92, 29.3 and 57.70 Mb on 2D and 680.6 Mb in 5A that increased GYLD between 18 – 40.67 g m^{-2} on single environment basis (18MCG, 19CH, 19CS, 19BI and 20MCG), and from 2.6 – 5.43 g m^{-2} for across environments and across mega-environments. Two major QTL with favorable alleles from TAM 204 were detected at 59.47 and 93.30 Mb on chromosome 2B. These QTL increased GYLD from 15.9-24.3 g m^{-2} in the single environment analysis (19CH, 19MCG and 20CS) and 3.3-9.4 g m^{-2} in the across environments and mega-environments analyses. Two minor QTL at 40.06 Mb on 2D and 61.45 Mb on 3D with the favorable allele from Iba were also identified; these QTL increases GYLD from 12 -20 g m^{-2} for single environment analysis (18DMS and 19MCG) and 4.13-5.75 g m^{-2} for across environments and mega-environments analyses.

Two more minor QTL with favorable alleles from TAM 204 at 560.08 Mb in 2A and 646.07 Mb in 6A, increased GYLD from 14.9-18.3 g m^{-2} for single environment analysis (19BI and 19EMN) and 1.16-4.86 g m^{-2} for across Environments analyses. All the major and minor QTL had larger LOD and R^2 for additive-by-environment interactions (AbyE) effect than only the additive

effect. For the major QTL with larger additive effects in the across environment analysis, the AbyE of *Qgyld.tamu.2D.28* and *Qgyld.tamu.5A.681* increased GYLD by 11.75-20.19 g m^{-2} at 19BI, respectively, whereas the *Qgyld.tamu.2B.93* increased GYLD by 20.37g m^{-2} at19MCG. Lastly, increments in grain yield due to the AbyE effect associated with the *Qgyld.tamu.2D.28* where detected at 19BI and 20MCG, with 11.775 and 20.53 g m^{-2}, respectively, although the allele from TAM 204 was the one that increased grain yield in 19BI, whereas the Iba allele in MCG20. For the consistent QTL detected across mega-environments, both *Qgyld.tamu.2B.59* and *Qgyld.tamu.2D.28* interact positively in mega-environments M1 and M3 which comprises environments from Central Texas, 19BI and 19EMN, respectively. On the other hand, *Qgyld.tamu.2D.29* interacted positively with ME2.

For PH, 13 consistent QTL were detected in chromosome 2A, 2B, 4A, 4B, 4D, 5A, 6A and 6D (Table 4.7). Three major QTL at 135.8, 240.2 and 250.7 Mb on chromosome 6D with favorable alleles from TAM 204 were identified in the population. These QTL increased PH by 2.1-2.45 cm in individual analysis, by 0.24-0.61 across environments, and by 0.79 cm across mega-environments. Besides, three minor QTL with favorable alleles from TAM 204 at 467.47Mb on 4B, 402.47 on 4D and 594.6 on 5A were identified. These minor QTL increased PH by 1.2-1.35 cm in the individual environments, by 0.42-0.72 cm across environments, and by 0.59 across mega-environments. The QTL inherited from Iba had minor effects and were identified at 65.112, 703.9 and 715.2 Mb on 2A, 57.16 Mb on 2B, 714.08 Mb on 4A, and 53.14 and 197.05 Mb on 6A. These QTL additive effects varied from 1.27-1.68 cm for the individual environments, 0.30-0.49 cm across environments, and from 0.68-0.88 cm across mega-environments. Amongst the major QTL, *Qph.tamu.6D.136*, *Qph.tamu.6D.240* and

Qph.tamu.6D.251 had larger LOD and R^2 for AbyE than additive effects. All the minor QTL had larger LOD and R^2 additive than AbyE effects, except for *Qph.tamu.2A.704, Qph.tamu.2B.57* and *Qph.tamu.6A.197* that had larger additive than AbyE effects.

The larger additive-by-environment effects of the major genes across environments were detected in 19BI for *Qph.tamu.6D.136*, 19EMN for *Qph.tamu.6D.240* 19CS and 19MCG for *Qph.tamu.6D.251*. This last-mentioned QTL interacted positively with the environments from Central Texas, where the population was grown under rain-fed conditions, whereas the other two interacted favorably with environments from the Texas Panhandle. The minor QTL *Qph.tamu.2A.704* also showed a high AbyE in 19CS, whereas *Qph.tamu.2B.57* showed a high AbyE in 19BI. No other minor QTL showed a higher AbyE effect >1cm in any location.

Concerning the AbyE across mega-environments, the major *Qph.tamu.6D.251* had a significant effect in ME3, which comprises 19CS, 19MCG and 20CS, increasing PH up to 1.49 cm. No other QTL had an AbyE effect >1cm across mega-environments.

In the case of DTH, four consistent QTL were found in at least three environments, across environments and mega-environments (Table 4.8). Three were considered major effect QTL given their high LOD values >18 and they explained a phenotypic variance exceeding 20% in at least one environment. These major QTL were detected at 25.88 and 29.34 Mb in chromosome 2D with favorable alleles from TAM 204, and at 575.05 on 5A and with favorable alleles from Iba.

Table 4.6. Consistent quantitative trait loci (QTL) for grain yield detected in the TAM 204/Iba population in individual environment, across environments and mega-environments.

QTL Name	Chrom[a]	Position (bp)	Env[b]	Threshold	Overall LOD	LOD(A)	LOD (AbyE)	Total PVE ranges[c]	PVE (A) (%)	PVE (AbyE) (%)	Additive effects[d]	Parental favorable alleles	Pleiotropic [c]
Qgyld.tamu.2A.560	2A	560.08	19BI and AcrossEnv	4.5-8	5-10.54	0.256	10.2837	5.59-7.31	0.146	5.44	1.16-18	TAM 204	
Qgyld.tamu.2B.59	2B	59.47	20CS, 19CH, AcrossEnv and AcrossME	3.46-8	6.21-20.89	7.1626	13.7285	9.067-24.32	4.11-12.32	5.03-11.61	6.17-22.1	TAM 204	y
Qgyld.tamu.2B.93	2B	93.3	19MCG and AcrossEnv	3.4-8	7.79-9.73	2.08	7.65	6.87-12.33	1.2	5.6626	-3.34-(-24.28)	Iba	
Qgyld.tamu.2D.28	2D	27.92	20MCG, 19CH, AcrossEnv and AcrossME	3.44-8	4.35-18	0.701-8.22	5.84-9.58	7.71-14.32	0.98-4.93	9.39-10.60	-2.62-(-28.1)	Iba	y
Qgyld.tamu.2D.29	2D	29.34	18MCG, AcrossEnv and AcrossME	3.48-8	5.63-11.19	0.80-3.57	4.83-7.61	4.24-38.33	1.24-2.01	2.22-6.95	-2.95-(-40.67)	Iba	y
Qgyld.tamu.2D.41	2D	40.64	18DMS, 19MCG and AcrossEnv	3.4-8	3.65-12.42	5.92	6.49	7.29-8.98	3.51	3.77	-5.75-(-20)	Iba	y
Qgyld.tamu.2D.58	2D	57.7	19CS and AcrossEnv	3.43-8	8.45-11.13	5.55	5.58	9-16.58	3.17	5.9	-5.43-(-28)	Iba	
Qgyld.tamu.3D.61	3D	61.45	19EMN and AcrossEnv	3.42-8	8.29-12.05	3.13	8.92	3.1-6.1	1.84	1.27	-4.13-(-12)	Iba	
Qgyld.tamu.5A.681	5A	680.62	19BI and AcrossEnv	3.97-8	9.81-14.39	2.32	12.07	8.99-13.68	1.32	7.66	-3.5-(-25)	Iba	
Qgyld.tamu.6A.464	6A	464.07	19BI, 19EMN and AcrossEnv	3.37-8	4.59-15.48	4.25	11.22	5.51-9.32	2.52	2.98	4.86-14.9	TAM 204	

Chrom[a] = chromosome; bp = base pairs; Env = Environments; Env[b] = AcrossEnv, across environments; AcrossME, across mega-environments; 18DMS, Dumas 2018; 18MCG, McGregor 2018; 19BI, Bushland Irrigated 2019; 19CH, Chillicothe 2019; 19CS, College Station 2019; 19EMN, Emeny Land 2019; 19MCG, McGregor 2019; 20CS, College Station 2020; 20EMN, Emeny Land 2020; 20.MCG, McGregor 2020; LOD= logarithm of the odds; PVE= Phenotypic Variation Explained; A= Additive effect; AbyE= Additive by environment effect; y[c] = yes.

In the analysis by the environment, the *Qdth.tamu.2D.26* additive effect ranged from 1.59-5.9 days, for *Qdth.tamu.2D.29* from 1.12-6.2 days, and for *Qdth.tamu.5B.575* from 1.13-4.9 days. According to the AbyE effect, the *Qdth.tamu.2D.26* was a major QTL with a smaller effect across environments with a range of 0.088-0.87 days. In contrast, the *Qdth.tamu.2D.29* QTL had a higher variation across environments with a range of 0.96-2.75 days. However, the *Qdth.tamu.5B.575* had a higher AbyE effect across environments with an increment of 3.16 days for 20CS. This can be confirmed when looking across mega-environments given it had a large AbyE in ME3, where 20CS was grouped, although it had a higher effect on ME2 (18DMs and 19BI). The remaining DTH QTL with a minor effect was identified at 50.53 Mb in chromosome 2B. This QTL had an additive effect on that ranged from 2.45-3.67 days in the individual environment. Interestingly, the AbyE effect of this minor QTL was higher than the *Qdth.tamu.2D.26* across environments and mega-environments, denoting a high QTL-by-environment interaction.

A set of eight consistent QTL was detected for TW at 520.53 Mb on 1A, 57.16 and 57.18 Mb on 2B, 29.34 and 40.64 Mb on 2D, 25.73 Mb on 4A, 695.72.9 Mb on 5A and 663.127 Mb on 5B (Table 4.9). From this set, five were considered major QTL with LOD scores higher than eight and PVE > 10 %. Both major QTL in chromosome 2B, *Qtw.tamu.2B.57.1* and *Qtw.tamu.2B.57.2*, had favorable alleles from TAM 204 associated with an increment of 7.11-12.80 kg m^{-3}. Whereas the other ones in 2D and 5A, *Qtw.tamu.2D.29*, *Qtw.tamu2D.41* and *Qtw.tamu.5A.696*, had favorable alleles from IBA and were associated with an increment of 8.53-12.17 kg m^{-3}.

Table 4.7. Consistent quantitative trait loci (QTL) for plant height detected in the TAM 204/Iba population in individual environment, across environments and mega-environments.

QTL Name	Chrom[a]	Position (bp)	Env[b]	Threshold	Overall LOD	LOD(A)	LOD (AbyE)	Total PVE ranges[c]	PVE (A) (%)	PVE (AbyE) (%)	Additive effects[d]	Parental favorable alleles	Pleiotropic[c]
Qph.tamu.2A.65	2A	65.11	19EMN and AcrossEnv	3.38-7.08	4.02-7.20	4.8611	2.3486	6.57-4.02	2.29	1.27	-0.45-(-1.27)	Iba	
Qph.tamu.2A.704	2A	703.94	19CS and AcrossEnv	3.48-7.08	5.62-7.57	2.6819	4.8882	3.6-7.86	1.26	2.64	-0.36-(-1.55)	Iba	
Qph.tamu.2A.715	2A	715.21	19BI, AcrossEnv and AcrossME	3.37-7.08	4.93-12.74	6.02-7.20	1.54-5.53	5.56-7.16	3.42-5.76	1.39-2.14	-0.59-(-1.41)	Iba	
Qph.tamu.2B.57	2B	57.16	19BI and AcrossEnv	3.37-7.08	4.26-8.12	1.87	6.33	3.46-5.41	0.89	2.56	-0.30-(-1.31)	Iba	y
Qph.tamu.4A.714	4A	714.08	AcrossEnv and AcrossME	4.65-7.08	7.29-7.9	4.89-5.66	2.23-2.4	3.61-7.59	2.3-4.49	1.2-3.1	-0.49-(-0.68)	Iba	
Qph.tamu.4B.467	4B	467.47	19BI and AcrossEnv	3.37-7.08	4.55-7.09	3.6	3.48	3.61-5.71	1.71	1.6	0.42-1.35	TAM 204	
Qph.tamu.4D.402	4D	402.4	AcrossEnv and AcrossME	4.65-7.08	4.66-7.98	3.59-5.06	1.07-2.37	4.17-5.68	2.63-3.39	1.53-2.28	0.52-0.59	TAM 204	
Qph.tamu.5A.595	5A	594.68	19BI and AcrossEnv	3.37-7.08	3.64-12.97	10.09	2.5806	4.49-6.98	4.7716	2.2116	0.72-1.120	TAM 204	
Qph.tamu.6A.53	6A	53.14	19BI, AcrossEnv and AcrossME	3.37-7.08	6.83-10.28	5.55-7.92	1.76-4.73	4.54-8.80	2.59-7.57	0.99-1.95	-0.51-(-1.68)	Iba	
Qph.tamu.6A.197	6A	197.05	19CS and AcrossEnv	3.43-7.08	3.94-7.57	3.51	4.05	4.10-5.39	1.67	2.43	-0.41-(-1.29)	Iba	
Qph.tamu.6D.136	6D	135.83	19BI and AcrossEnv	3.37-7.08	12.34-12.49	1.22	11.12	5.88-16.92	0.58	5.3	0.24-2.34	TAM204	
Qph.tamu.6D.240	6D	240.2	19EMN and AcrossEnv	3.38-7.08	11.80-11.92	2.16	9.76	6.18-19.97	1.01	5.17	0.32-2.28	TAM204	
Qph.tamu.6D.251	6D	250.71	19MCG, 19CS, AcrossEnv and AcrossME	3.4-7.08	10.2-23.50	6.26-7.70	5.69-15.80	15.06-21.23	3.69-6.15	9.14-10.80	0.61-2.45	TAM204	y

Chrom[a] = chromosome; bp = base pairs; Env = Environments; Env [b] =AcrossEnv, across environments; AcrossME, across mega-environments; 19BI, Bushland Irrigated 2019; 19CS, College Station 2019; 19EMN, Emeny Land 2019; LOD= logarithm of the odds; PVE= Phenotypic Variation Explained; A= Additive effect; AbyE= Additive by environment effect; y[c] = yes.

Although the minor *Qtw.tamu.1A.521* and *Qtw.tamu.5B.663*, did not explain a high PVE in the environment where were identified, their additive effects are moderate-high with 9.38 and 8.19 kg m^{-3}, respectively. All the stable QTL had larger LOD and R^2 for the additive effect than the AbyE effect across environments, except for Qtw.tamu.5A.696 that had higher AbyE LOD, and *Qtw.tamu.1A.521* and Qtw.tamu.2D.41 with higher AbyE R^2. Across-megaenvironments most of the QTL had larger LOD and R^2 for the additive effect than the AbyE effect, except for *Qtw.tamu.2D.29* and *Qtw.tamu.2D.4* than had higher LOD and R^2 for the AbyE effect.

No consistent QTL was detected when comparing individual environments for TKW (Table 4.9). Four QTL were detected in 19MCG, whereas two were detected in 18MCG, and one in 20MCG. Six out of this set were identified across environments. The QTL *Qtkw.tamu.2B.59*, *Qtkw.tamu.2B.83*, *Qtkw.tamu.2D.29* and *Qtkw.tamu.6A.478* were considered major QTL given they explained more than 15% of the PVE in the individual environment analysis. The first and second major QTL increased TKW by 1.6 and 1.4 g, respectively, and had favorable alleles from TAM 204, whereas the other two increased it by 2.54 and 2.01, respectively, with favorable alleles from Iba. With regards to across environment analysis, *Qtkw.tamu.2D.29* had larger additive R^2 than AbyE, and the *Qtkw.tamu.2B.59* and *Qtkw.tamu.2B.83* had larger AbyE R^2 than the additive effect. Modest AbyE effects were observed in 19MCG, where the *Qtkw.tamu.2B.83* increased TKW 0.64 g. However, a major AbyE effect was observed at 20MCG, where the *Qtkw.tamu.2B.59* increased TKW by 0.99 g. The *Qtkw.tamu.6A.478* QTL was not detected in the across environment analysis.

Table 4.8. Consistent quantitative trait loci (QTL) for days to heading detected in the TAM 204/Iba population in individual environment, across environments and mega-environments.

QTL Name	Chrom[a]	Position (bp)	Env[b]	Threshold	Overall LOD	LOD(A)	LOD (AbyE)	Total PVE ranges[c]	PVE (A) (%)	PVE (AbyE) (%)	Additive effects	Parental favorable alleles	Pleiotropic[c]
Qdth.tamu.2B.51	2B	50.53	20MCG, 20CS, 19MCG, 19CS, AcrossEnv and AcrossME	3.4-7.26	7.97-40.72	5.85-30.90	3.59-9.81	8.07-14.78	4.34-7.23	7.54-9.01	-0.93-(-3.67)	Iba	
Qdth.tamu.2D.26	2D	20.88	20MCG, 18MCG, 18CS, AcrossEnv and AcrossME	3.45-7.26	4.36-18.17	4.28-9.76	1.76-3.55	3.32-40.96	2.15-3.11	1.16-2.07	0.78-5.90	TAM 204	
Qdth.tamu.2D.29	2D	29.34	20MCG, 20CS, 19MCG, 19CS, 19BI, AcrossEnv and AcrossME	3.37-7.26	12.07-113.75	37.16-99.61	7.78-14.13	21.2-48.48	26.64-32.73	11.56-15.75	1.26-6.27	TAM 204	y
Qdth.tamu.5B.575	5B	575.03	20MCG, 20CS, 19MCG, 19CS, AcrossEnv and AcrossME	3.4-7.26	4.38-45.82	42.25	2.071-3.56	4.32-25.41	10-10.66	9.96-11.74	-1.46-(-4.93)	Iba	

Chrom[a] = chromosome; bp = base pairs; Env = Environments; Env[b] = AcrossEnv, across environments; AcrossME, across mega-environments; 18CS, College Station 2018; 18MCG, McGregor 2018; 19CS, College Station 2019; 19MCG, McGregor 2019; 20CS, College Station 2020; 20.MCG, McGregor 2020; LOD= logarithm of the odds; PVE= Phenotypic Variation Explained; A= Additive effect; AbyE= Additive by environment effect; y[c] = yes.

All the consistent QTL discovered for GYLD, yield components and agronomic traits were found in 9 chromosomes, where the majority were in chromosome 2B and 2D, with eight and nine QTL respectively. These findings are confirmed by the multiple studies in which several QTL have been found affecting GYLD and yield components in all of these chromosomes (Assanga et al., 2017; Bennett et al., 2012; Jamil et al., 2019; Ogbonnaya et al., 2017). These genomic regions have a high potential to be exploited in wheat breeding programs.

4.3.4 Pleiotropic QTL

A set of seven QTL regions were considered to have pleiotropic effects, affecting more than one trait. All the six were consistent for at least one trait. They were located at 57.16 Mb on 2B, affecting PH and DTH, with favorable alleles increasing the traits from Iba, and TW with favorable alleles from TAM 204. The second and third QTL on 2B at 57.18 and 59.47 Mb affect GYLD and TW with favorable alleles from TAM 204. The fourth QTL on 2D at 27.92 Mb, affect GYLD and TW with favorable alleles from Iba, and DTH with favorable alleles from TAM 204. The fifth pleiotropic QTL detected on 2D at 29.34 Mb is the most interesting one as it affected all the traits and was consistent for three of them. For this QTL, Iba provided the favorable alleles for GYLD, TW and TKW, whereas TAM 204 was the source of favorable alleles for PH and DTH . The pleiotropic effect of this QTL may be due to the photoperiod sensitivity gene *Ppd-D1*, which is at 34 Mb on 2D The QTL at 40.64 on chromosome 2D was pleiotropic for GYLD, TW and TKW with all the favorable alleles from Iba. The last QTL with pleiotropic effect for the agronomic traits was at 250.71 Mb on chromosome 6D. This QTL affected GYLD and PH with favorable alleles from Iba and TAM 204, respectively.

Table 4.9. Consistent quantitative trait loci (QTL) for test weight and thousand kernel weight detected in the TAM 204/Iba population in individual environment, across environments and mega-environments.

QTL Name	Chrom[a]	Position	Traits[b]	Environments[c]	Threshold	Overall LOD	LOD(A)	LOD (AbyE)	Total PVE ranges[c]	PVE (A) (%)	PVE (AbyE) (%)	Additive effects	Parental favorable alleles
Qtw.tamu.1A.521	1A	520.53	TW	19MCG, AcrossEnv and AcrossME	3.4-5.68	4.5-7.83	5.15-6.13	0.02-1.69	6.17-8.13	3.96-4.59	1.58-4.16	2.94-9.38	TAM 204
Qtw.tamu.2B.57.1†	2B	57.16	TW	20MCG, 19MCG, and AcrossEnv	3.4-5.68	5.44-16.39	15.37	1.01	9.67-17.79	10.26	7.52	4.73-12.8	TAM 204
Qtw.tamu.2B.57.2†	2B	57.18	TW	19EMN and AcrossME	3.42-4.01	4.41-15.80	15.73	0.06	6.51-19.26	15.3	3.95	5.03-8.85	TAM 204
Qtw.tamu.2D.29†	2D	29.34	TW	20EMN, AcrossEnv and AcrossME	3.46-5.68	7.79-12.17	3.96-6.41	4-5.74	4.25-15.41	3.4-4.18	0.55-2.6	-3.02-(-8.97)	Iba
Qtw.tamu.2D.41†	2D	40.64	TW	19MCG, AcrossEnv and AcrossME	3.4-5.68	7.79-8.88	6.44-6.51	1.92-2.44	11.18-13.60	4.3-5.91	6.13-9.3	-3.08-(-12.17)	Iba
Qtw.tamu.4A.26	4A	25.73	TW	19EMN, AcrossEnv and AcrossME	3.42-5.68	4.77-7.68	4.33-4.7	1.21-3.34	4.15-6.96	2.86-4.14	0.0034-1.60	-2.49-(-5.19)	Iba
Qtw.tamu.5A.696	5A	695.72	TW	19EMN, AcrossEnv and AcrossME	3.42-5.68	7.3-13.93	4.9-5.98	2.4-7.94	4.35-18.59	3.92-4.25	0.1-3.39	-2.93-(-8.53)	Iba
Qtw.tamu.5B.663	5B	663.12	TW	19MCG and AcrossEnv a	3.4-5.68	3.72-6.55	3.3	0.24	5.14-7.33	4.2	3.136	-3.02-(-8.19)	Iba
Qtkw.tamu.2B.59†	2B	59.47	TKW	20MCG and Across Env	3.45-4.64	10.64-10.83	6.11	4.72	17.31-19.99	6.64	10.66	0.55-1.6	TAM 204
Qtkw.tamu.2B.83†	2B	82.79	TKW	19MCG and Across Env	4.4-4.64	11.07-11.18	5.08	6.09	13.98-16.09	5.52	8.64	0.50-1.4	TAM 204
Qtkw.tamu.2D.29†	2D	29.34	TKW	18MCG and Across Env	3.48-4.64	7.19-11.21	4.82	2.37	6.46-29.68	5.1	1.32	-0.48-(-2.54)	Iba
*Qtkw.tamu.6A.478**	6A	477.76	TKW	18MCG	3.48	7.86	NA	NA	18.55			-2.01	Iba

Chrom[a] = chromosome; Traits[b] = TW, Test weight; TKW, Thousand kernel weight; Environments[c] = AcrossEnv, across environments; AcrossME, across mega-environments; 18MCG, McGregor 2018; 19EMN, Emeny Land 2019; 19MCG, McGregor 2019; 20EMN, Emeny Land 2020; 20MCG, McGregor 2020; LOD= logarithm of the odds; PVE= Phenotypic Variation Explained; A= Additive effect; AbyE= Additive by environment effect; † = pleiotropic; * = no consistent QTL.

A consistent QTL for TW was also found to be pleiotropic for awns presence, both traits with favorable alleles from Iba. This association between the presence of awns and TW has been previously described (DeWitt et al., 2019). DeWitt et al. (2019) hypothesized that the relationship between them is primarily due to the role that the gene *B1* plays in florets development rather than decrement of photosynthetic capacity due to the absence of awns. Also, in most cases, the pleiotropic effects were not found within the same environment despite all traits being collected in almost every environment, excluding TKW that was collected in 18MCG, 19MCG, and 20MCG. However, these three environments were the only ones in which a pleiotropic was detected. In 18MCG and 20MCG, the QTL detected on 2D at 29.34 Mb had a pleiotropic effect on GYLD and TKW, and PH and DTH, respectively (Figure 4.9). Whereas the QTL at 40.64 on chromosome 2D affected GYLD and TW in 19MCG.

It is essential to distinguish whether these genomic regions effect on multiple traits is due to a real pleiotropic effect of a major gene or linkage between multiple genes. In pleiotropy, the same biological process regulates the traits; whereas linkage results from selection or an evolutionary process that caused stacking of favorable gene (Saltz et al., 2017). Further fine mapping will elucidate the genetic control of these traits.

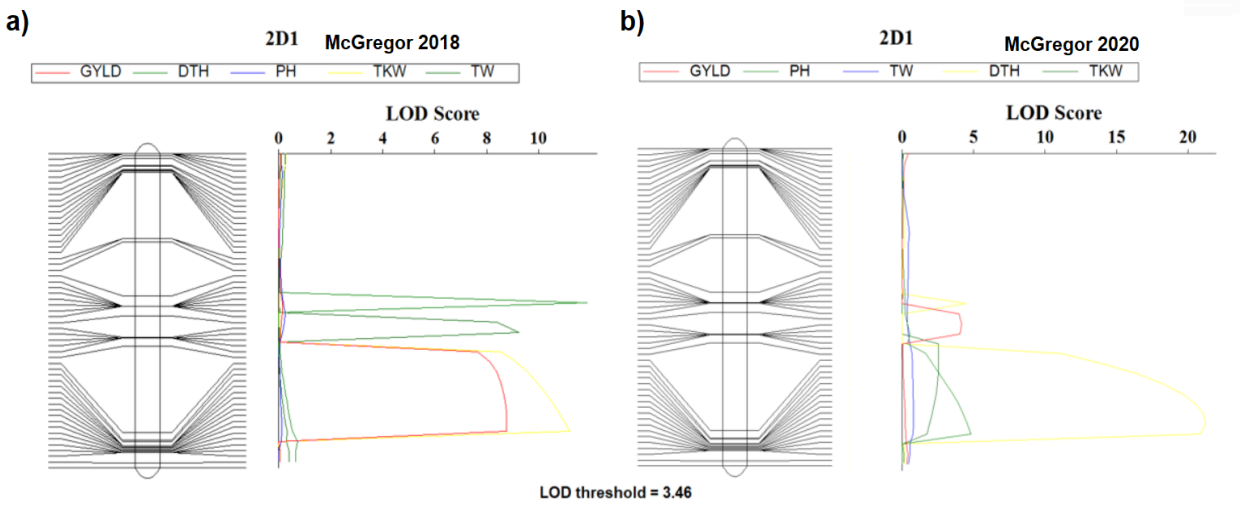


Figure 4.9. Pleiotropic effect of the QTL detected between 29.34-35.53 Mb on chromosome 2D affecting grain yield (GLYD), days to heading (DTH), plant height (PH) and thousand kernel weight (TKW). TW= test weight.

4.3.5 Epistasis, epistasis-by-environment, and additive-by-environment interactions

For most of the traits, only epistatic interactions involving consistent QTL or overall LOD scores >10.0 were summarized for epistasis and AbyE interactions. Among all the 138 epistatic interactions identified for GYLD with LOD scores >10, 16 of them had overall LOD scores >15.0 (Figure 4.10). From these 16 epistatic interactions, ten had additive-by-additive LOD scores >10.0, and two had AbyE interactions with LOD > 10.0. Two of the major consistent QTL were involved in epistatic interaction with other genomic regions. The QTL *Qgyld.tamu.2D.29* interacted with another genomic region located at 556.02 on 1B. Positive AbyE effects were detected from this epistatic interaction at 19CS, 19BI and 20MCG that increased GYLD from 7.0–19.9 g m^{-2} with favorable allele from TAM 204, whereas in 18MCG and 19MCG increments in yield ranged from 3-8.7 g m^{-2} and were associated with Iba alleles. The other consistent QTL, Qgyld.tamu.2B.93, interacted with a genomic region located at 603.24 Mb on 4A. This second epistatic interaction, increased GYLD at 18MCG, 19BI and 19MCG from 0.9-20.3 g m^{-2} with favorable alleles from TAM 204, and in 19CS and 20MCG it

increased AbyE iby 6.0-11.0 g m^{-2} and was associated with Iba alleles. The *Qgyld.tamu.2D.29* was also involved in an epistatic interaction with a genomic region at 556.02 on 1B, where exceptional epistasis-by-environment interaction was detected in 20MCG and was associated with an increment of 12.72 g m^{-2} with favorable alleles from TAM 204. The *Qgyld.tamu.5A.681* was the only major QTL found interacting with another genomic region at 622.28 Mb on 2A with an overall LOD of 6.36 across mega-environments. In this epistasis, the AbyE effect of the major QTL increased up to 2.41 and 2.44 g m^{-2} in ME1 (18MCG, 19CH, 19MCG, 19 CS and 20CS) and ME2 (18DMS and 20 EMN), respectively. Nevertheless, the other genomic region involved in this epistasis increased GTLD by a magnitude of 4.33 and 5.2 g m^{-2} for ME1 and ME2, respectively. This epistatic interaction affected the ME by magnitude no larger than 0.8 g m^{-2}. Similarly, the epistasis-by-environment interaction increased grain yield in ME1 and ME2 4.42 by 4.23 g m^{-2}.

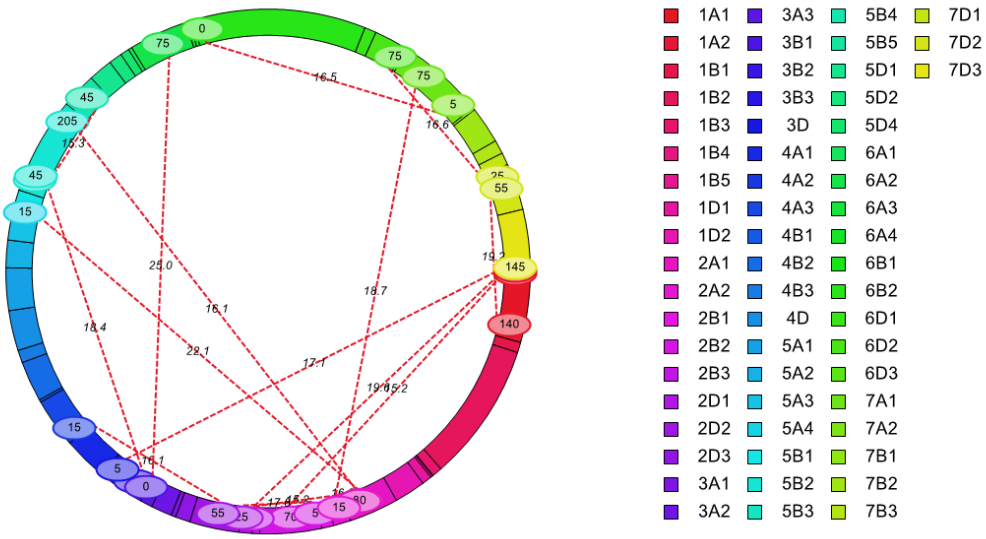


Figure 4.10. Whole-genome significant epistatic interaction at LOD >15 for grain yield across environments. Each color represents a different linkage group described in the right part of the figure. Numbers inside the circles represent the peak cM of the SNPs involved in the epistatic interaction. The numbers in the dashed lines represent the LOD score of that epistasis.

For DTH, a total of 305 epistatic interaction with an overall LOD score >10 were detected, from which 72 had overall LOD score >15 (Figure 4.11), and nine > 20. Besides, from this set of nine epistatic interactions, three had epistasis LOD >15 and two of them had AbyE interactions LOD > 10.0. Three out of the four consistent QTL were found to be interacting with other genomic regions affecting this trait. The *Qdth.tamu.2D.29* was found in an epistatic interaction with six genomic regions on chromosome 2B, 3A, 4A, 4B and 6A. Among these, the epistasis with a genomic region at 57.18 on 2B had an overall, epistasis and epistasis-by-environment LOD of 40.22, 27.48, and 12.77, respectively. This QTL increased DTH in all the six epistatic interactions, but the major effect was in 18DMS and 19CS with 2.29-2.64 and 2.67-2.9 days, respectively. Interestingly, for 18DMS the favorable alleles were from Iba, whereas for 19CS they were from TAM 204. Similarly, the epistatic interaction between *Qdth.tamu.2B.51* and a QTL at 49.69 days in 6A also impacted DTH, increasing it by 0.98-2.22 days across locations. The *Qdth*.tamu.2D.26 was involved in 3 epistasis, although the epistatic effect of this QTL did not increase the DTH by more than one day. The *Qdth.tamu.2D.29* QTL and the QTL at 57.18 on 2B were involved in the only epistatic interaction with a LOD score >10 Across mega-environments, all the epistatic interaction involving consistent QTL had higher LOD and R^2 epistasis than epistasis-by-environment. The AbyE effect of the major QTL in this epistatic interaction increased 1.17 days for ME1 (18CS and 18MCG), 1.21 days for ME2 (18DMS and 19BI), and 2.38 days for ME3 (19CS, 19MCG, 20CS and 20MCG). Also, increments of 0.44, 0.45 and 0.89 days for ME1, ME2 and ME3 were obtained as a result of the epistasis-by-environment interaction.

Table 4.10. Epistatic interaction involving a consistent quantitative trait loci (QTL) for grain yield detected in the TAM 204/Iba population in across environments and mega-environments analyses.

Analysis	LeftMarker1	LeftMarker2	LOD	LOD AA	LOD AbyE	PVE	PVE AA	PVE AAbyE		19CS	19MCG	18MCG	19BI	19CS	20MCG			
AcrossEnV	chr1B_556028797	chr2D_29344035	13.6	8.6	4.9	1.5	0.9	0.63	6.43	-3.88	6.49	-8.76	6.61	-4.46	12.72			
AcrossEnV	chr2B_57183671	chr2D_29344035	17.5	12.8	4.7	2.4	1.3	1.07	7.75	16.08	-2.97	-8.68	6.98	-4.69	19.89			
AcrossEnV	chr2B_93308427	chr4A_603246965	13.8	9.20	4.66	1.7	1.1	0.62	-6.96	-6.03	20.28	2.16	0.85	-6.71	-11.03			
									AbyA	A1byE_1	A1byE_2	A1byE_3	A2byE_1	A2byE_2	A2byE_3	AAbyE_1	AAbyE_2	AAbyE_3
AcrossME	chr2A_622280405	chr5A_680628854	6.3	5.2	1.1	2.5	2.1	0.37	7.42	-5.24	4.43	0.80	2.44	-2.41	-0.02	4.4	-4.2	-0.19

LOD= logarithm of the odds; PVE= Phenotypic Variation Explained; AA= Additive by additive effect; AAbyE= Epistasis-by-environment effect. AbyA =Epistasis additive effect. A1 in the Additive by environment epistatis refers to the LeftMarker1 and A2 to the LeftMarker2. AcrossEnv= Across Environments; AcrossME= Across mega-environments; 18DMS, Dumas 2018; 18MCG, McGregor 2018; 19BI, Bushland Irrigated 2019; 19CH, Chillicothe 2019; 19CS, College Station 2019; 19EMN, Emeny Land 2019; 19MCG, McGregor 2019; 20CS, College Station 2020; 20EMN, Emeny Land 2020; 20MCG, McGregor 2020. In the AcrossME, E_01 refers to ME1 (18MCG, 19CH, 19MCG, 19 CS and 20CS, E_02 to ME2 (18DMS and 20 EMN), and E_03to ME3 (19BI and 19EMN).

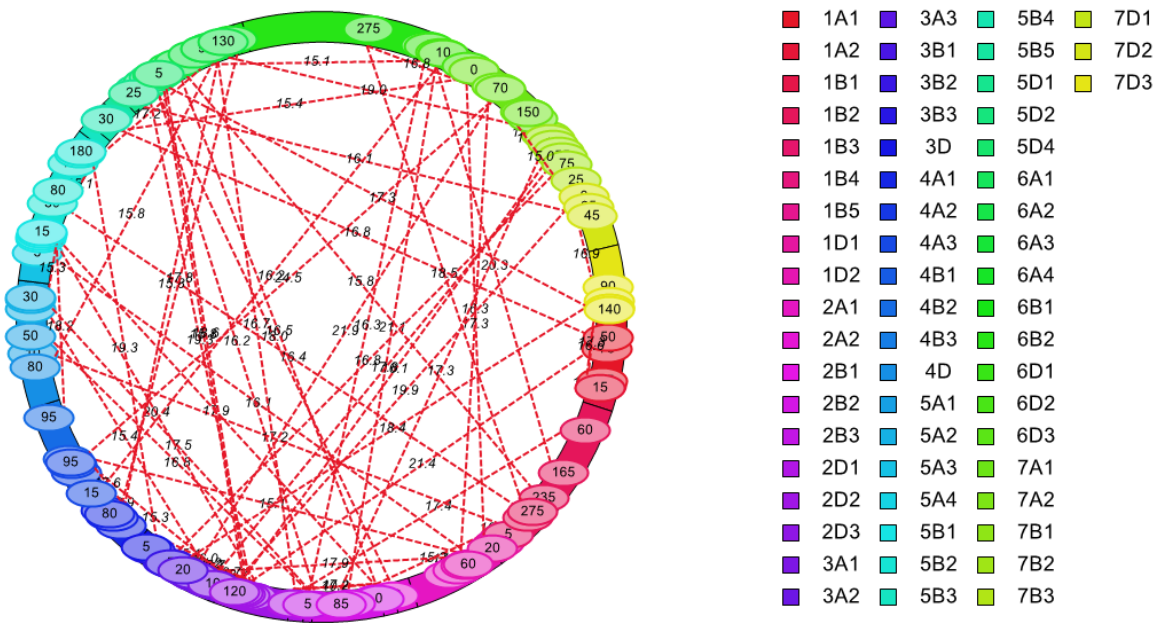


Figure 4.11. Whole-genome significant epistatic interaction at LOD >15 for days to heading across environments. Each color represents a different linkage group described in the right part of the figure. Numbers inside the circles represent the peak cM of the SNPs involved in the epistatic interaction. The numbers in the dashed lines represent the LOD score of that epistasis.

For PH, a set of 278 epistatic interactions had a larger overall LOD score than the one obtained from the 1000 permutation test. Among them, 50 had overall LOD>10 (Figure 4.12), and out of this 27 set, only five had epistasis LOD>10. None of them had epistasis-by-environment LOD>10. The major QTL *Qph.tamu.2A.65* was involved in two epistatic interactions with a genomic region at 7.16 Mb on 5A and 7.48 Mb on 6D with an overall LOD score of 10.11 and 8.85, respectively. In both cases the LOD and R^2 for epistasis were larger than epistasis-by-environment interaction. The larger AbyE effect on PH of *Qph.tamu.2A.65* was detected in 19EMN with increments between 0.72-7.32 cm, with favorable alleles from Iba. The epistasis-by-environment effect was manifested in18DMS, 18MCG and 20CS with increments in PH of 0.36-0.54, 0.60-0.61 and 0.36-0.66 cm, respectively. Each region at an overall LOD of 8.58 and

10.11. No epistatic interaction with overall LOD> 10 was detected for PH across mega-environments.

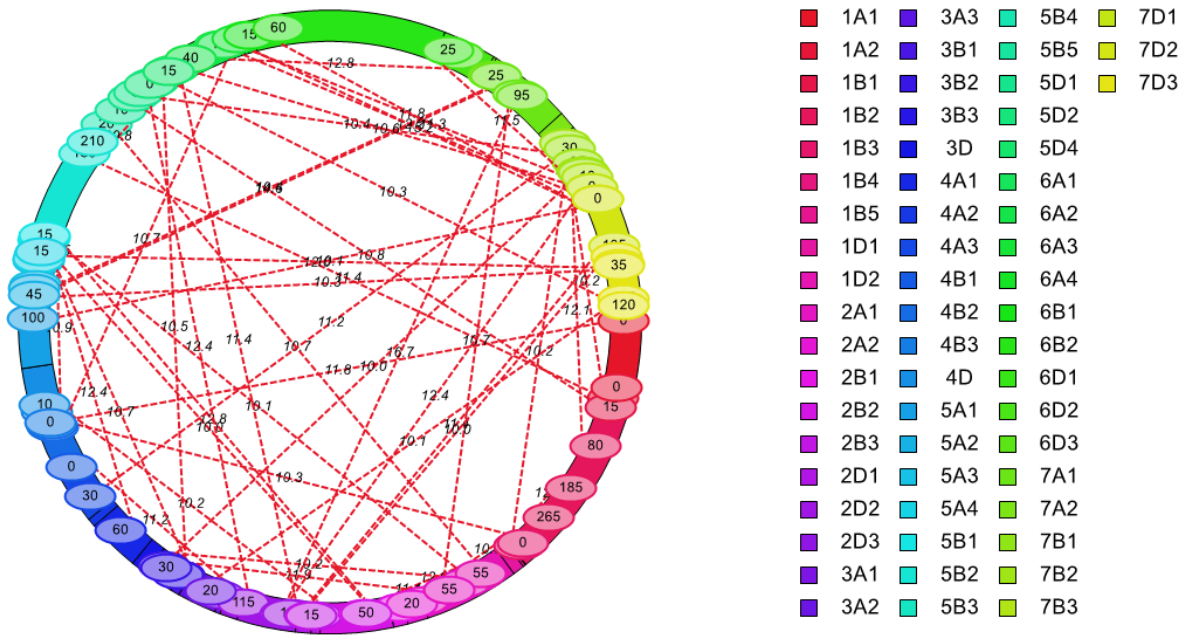


Figure 4.12. Whole-genome significant epistatic interaction at LOD >10 for plant height across environments. Each color represents a different linkage group described in the right part of the figure. Numbers inside the circles represent the peak cM of the SNPs involved in the epistatic interaction. The numbers in the dashed lines represent the LOD score of that epistasis.

For TW, only two epistatic interactions were found to have LOD>10 when evaluating across environments. None of them were associated with a consistent QTL. Nevertheless, when considering the threshold obtained from 1000 permutations test, among the 96 epistatic interactions found (Figure 4.13), four were involved the consistent *Qtw.tamu.2D.29* (three times), and *Qtw.tamu.2B.57.2* were obtained. The three epistases of *Qtw.tamu.2D.29* had larger LOD and R^2 for epistasis than for epistasis-by-environment; the opposite of the epistasis interaction of *Qtw.tamu.2B.57.2*. The larger the AbyE effects of the *Qtw.tamu.2D.29* were obtained at 18MCG and 20EMN with increases of 2.22-3.05 and 3.88-4.03 kg m^{-3}, respectively.

In the case of *Qtw.tamu.2D.29*, the larger AbyE effects were found at 19EMN and 19MCG with increases of 1.90 and 1.08 kg m^{-3}, respectively. Moderate increments in TW in the three epistatic interactions involving *Qtw.tamu.2D.29* were obtained in 18MCG and 19MCG in the range of 1.58-2.33 and 2.01-3.31 kg m^{-3}, respectively. Also, the epistasis-by-environment effect between *Qtw.tamu.2B.57.2* and a QTL at 44.28 Mb on 2A, increased TW in 18MCG, 19EMN, 19MCG and 19CS by 1.54-2.33 kg m^{-3}. The *Qtw.tamu.2D.29* was also involved in an epistatic interaction in the across the mega-environments analysis and was associated with increments of 1.90 kg m^{-3} in ME1 (18MCG, 19CS and 19MCG) and ME2 (19EMN and 20EMN). Similarly, the epistatic-by-environment effect increased by 2.12 kg m^{-3} in both mega-environments.

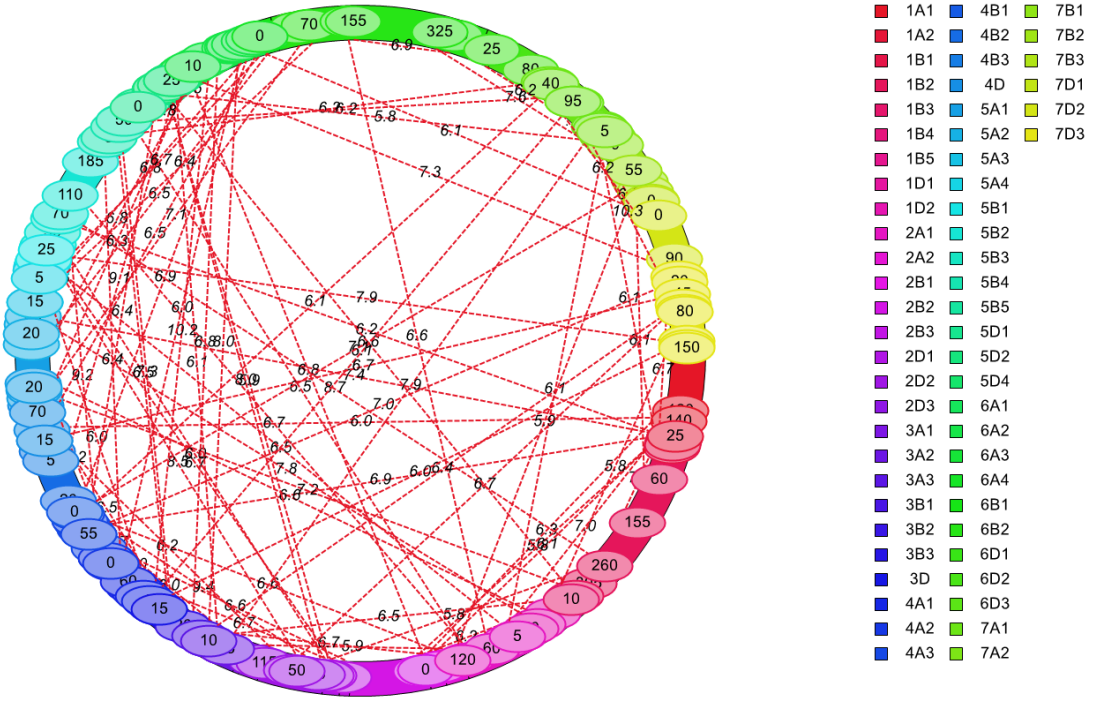


Figure 4.13. Whole-genome significant epistatic interaction at LOD >5.8 for test weight across environments. Each color represents a different linkage group described in the right part of the figure. Numbers inside the circles represent the peak cM of the SNPs involved in the epistatic interaction. The numbers in the dashed lines represent the LOD score of that epistasis.

In the case of TKW, 37 epistatic interactions were found (Figure 4.14). Amongst these epistatic interactions, one involved the major *Qtkw.tamu.2D.29* QTL at 57.18 Mb on 2B. For this genes interaction, the LOD and R^2 were larger for epistasis-by-environment than for the epistasis effect. The major gene AbyE effect increased TKW by 0.19, 0.36 and 0.17 g in 18MCG, 19MCG and 20MCG. Whereas the AbyE of the QTL on 2B increased TKW by 0.36, 0.12 and 0.24 g for the same environments. Regarding the epistasis-by-environment effect, the major increments in TKW were observed in 18MCG and 20MCG with 0.51 and 0.57 g, respectively. Any of these epistases, the additive, additive-by-environment, or epistasis-by-environment effects, increased TKW >1 gr.

The contribution of the epistasis to the total genetic variation complicates the selection of QTL for MAS, as it makes the diagnostic of superior cultivars more challenging. However, a reduced number of epistatic interactions involving the consistent genes were identified. In addition, some of the epistatic interactions involved two consistent QTL, facilitating their use in MAS.

Table 4.11. Epistatic interaction involving a consistent quantitative trait loci (QTL) for days to heading (DTH) and plant height (PH) detected in the TAM 204/Iba population across environments and mega-environments analyses.

Trait	Analysis	LeftMarker1	LeftMarker2	LOD	LOD(AA)	LOD(AAbyE)	PVE	PVE(AA)	PVE(AAbyE)	AddbyAdd	18DMS A1byE_5	19CS A1byE_7	18DMS A2byE_5	19CS A2byE_7
DTH	AcrossEnV	chr2B_57183671	chr2D_29344035	40.2	27.49	12.73	3.31	2.53	0.78	-1.27	0.2396	-0.1603	-2.2883	2.6925
DTH	AcrossEnV	chr2B_30141165	chr2D_29344035	17.2	10.47	6.73	1.15	0.75	0.4	-0.69	0.1094	0.0205	-2.6464	2.9047
DTH	AcrossEnV	chr2D_29344035	chr6A_220892045	16.7	10.81	5.93	1.48	1.11	0.37	0.83	-2.291	2.6912	0.359	-0.5688
DTH	AcrossEnV	chr2B_50536274	chr6A_49698383	16.4	7.56	8.93	1.25	0.71	0.54	0.68	1.4162	-2.2249	0.1218	-0.1208
DTH	AcrossEnV	chr2D_25882515	chr4A_699842612	13.6	7.8	5.89	1.01	0.75	0.26	-0.68	-0.7222	0.1867	0.1071	0.0538
DTH	AcrossEnV	chr2D_25882515	chr5B_481863102	12.1	7.62	4.51	1.12	0.73	0.39	0.68	-0.7277	0.1996	-0.176	0.0246
DTH	AcrossEnV	chr2D_29344035	chr4B_656219662	11.6	6.23	5.44	0.82	0.63	0.18	0.64	-2.3088	2.7271	0.4386	-0.8553
DTH	AcrossEnV	chr2D_29344035	chr3A_734508576	10.6	7.96	2.67	0.98	0.84	0.14	-0.73	-2.2675	2.6779	0.1173	-0.462
DTH	AcrossEnV	chr2D_29344035	chr4A_25105207	10.5	7.2	3.4	0.87	0.73	0.14	0.69	-2.2776	2.6701	0.4576	-0.2445
DTH	AcrossEnV	chr2D_25882515	chr5B_529855145	10.1	5.93	4.19	1.01	0.59	0.42	0.6	-0.7211	0.1872	0.0123	0.0217
											A2byE_1	A2byE_2	A2byE_3	
DTH	AcrossME	chr2B_57183671	chr2D_29344035	13	10.66	2.38	3.19	2.54	0.66	-1.24	-1.1743	-1.2124	2.3867	
											19EMN	18DMS	18MCG	19MCG
											A1byE_3	AAbyE_2	AAbyE_1	AAbyE_5
PH	AcrossEnV	chr2A_65112158	chr5B_7165731	10.11	6.41	3.7	1.19	0.8	0.39	-0.54	-0.7244	0.3626	0.614	-0.2529
PH	AcrossEnV	chr2A_65112158	chr6D_7483566	8.58	4.48	4.11	0.93	0.58	0.35	-0.46	-0.732	0.5453	0.6053	-0.2224

LOD= logarithm of the odds; PVE= Phenotypic Variation Explained; AA= Additive by additive effect; AAbyE= Epistasis-by-environment effect. AddbyAdd =Epistasis additive effect. A1 in the Additive by environment epistatis refers to the LeftMarker1 and A2 to the LeftMarker2. AcrossEnv= Across Environments; AcrossME= Across mega-environments; 18CS, College Station 2018; 18DMS, Dumas 2018; 19BI, Bushland Irrigated 2019; 19CS, College Station 2019; 19EMN, Emeny land 2019; 19MCG, McGregor 2019; 20CS, College Station 2020; 20MCG, McGregor 2020 . In the across mega-environments for DTH. E_1 refers to ME1 (18CS and 18MCG), E_02 to ME2 (18DMS and 19BI), and E_03to ME3 (19CS, 19MCG, 20MCG and 20CS).

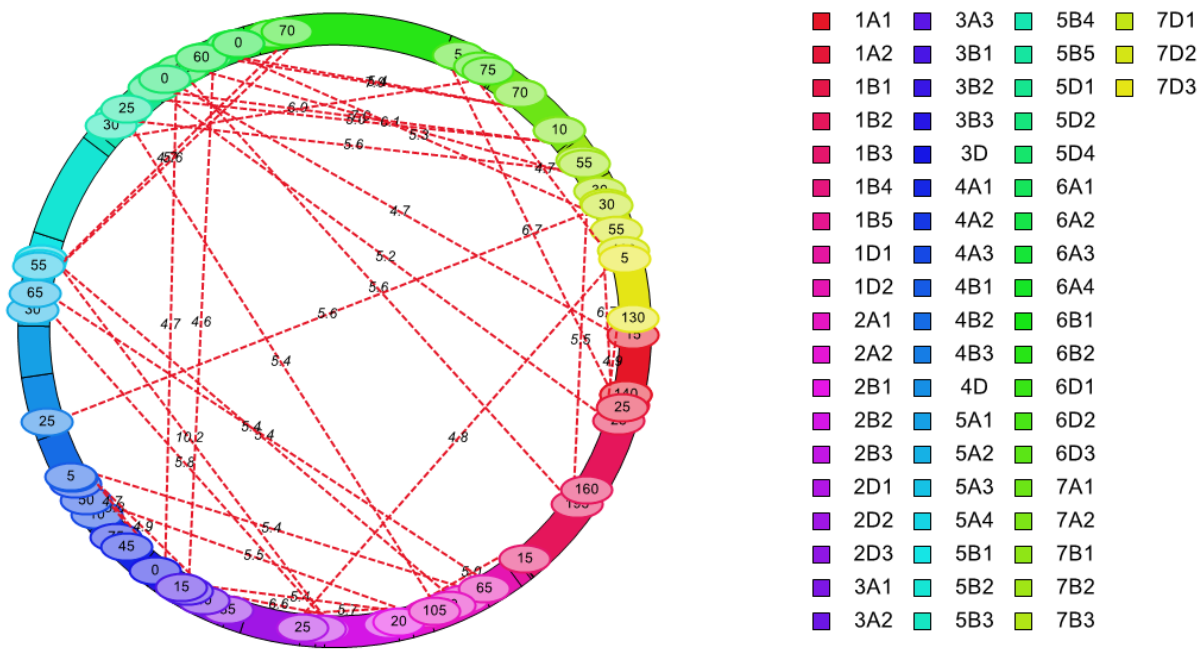


Figure 3.14. Whole-genome significant epistatic interaction at LOD >4.6 for thousand kernel weight across environments. Each color represents a different linkage group described in the right part of the figure. Numbers inside the circles represent the peak cM of the SNPs involved in the epistatic interaction. The numbers in the dashed lines represent the LOD score of that epistasis.

Table 4.12. Epistatic interaction involving consistent quantitative trait loci (QTL) for Test weight (TW) detected in the TAM 204/Iba population across environments and mega-environments analyses.

Analysis	LeftMarker1	LeftMarker2	LOD	LOD AA	LOD AAbyE	PVE	PVE AA	PVE AAbyE	AbyA	20EMN A1byE_2	18MCG A1byE_5	20EMN A2byE_2	18MCG A2byE_5	19MCG A2byE_3	19EMN A2byE_1	19MCG AAbyE_3	18MCG AAbyE_5
AcrossEnV	chr2D_29344035	chr5B_6533551	6.68	4.45	2.23	1.22	0.8	0.42	2.43	-4.28	3.06	0.07	-0.75	0.53	0.42	3.31	-2.33
AcrossEnV	chr2D_29344035	chr5D_344895472	5.87	4.31	1.56	1.12	0.79	0.33	-2.44	-3.88	2.23	0.8	0.82	-2.36	1.63	-2.98	1.87
AcrossEnV	chr2A_44281148	chr2B_57183671	6.23	2.1	4.12	0.84	0.41	0.43	1.73	-0.55	-0.29	0.27	-0.94	-1.09	1.9	0.63	-2.34
AcrossEnV	chr1B_541073782	chr2D_29344035	5.83	4.43	1.4	1.08	0.83	0.25	2.49	-0.26	-0.04	-4.3	3.02	0.5	-0.78	2.02	-1.58
										A1byE_1	A1byE_2	A2byE_1	A2byE_2	AAbyE_1	AAbyE_2		
AcrossME	chr2D_29344035	chr5D_344895472	5.55	4.27	1.28	2.61	2.07	0.54	-4.25	1.91	-1.91	-1.07	1.07	-2.12	2.12		

LOD= logarithm of the odds; PVE= Phenotypic Variation Explained; AA= Additive by additive effect; AAbyE= Epistasis-by-environment effect. AbyA =Epistasis additive effect. A1 in the Additive by environment epistasis refers to the LeftMarker1 and A2 to the LeftMarker2. AcrossEnv= Across Environments; AcrossME= Across mega-environments; 18MCG, McGregor 2018; 19CS, College Station 2019; 19EMN, Emeny Land 2019; 19MCG, McGregor 2019; 20EMN, Emeny Land 2020; 20MCG, McGregor 2020. In the across mega-environments E_1 refers to ME1 (18MCG, 19CS and 19MCG) and E_02 to ME2 (19EMN and 20EMN).

Table 4.13. Epistatic interaction involving consistent quantitative trait loci (QTL) for thousand kernel weight detected in the TAM 204/Iba population in across environments analysis.

LeftMarker1	LeftMarker2	LOD	LOD(AA)	LOD(AAbyE)	PVE	PVE(AA)	PVE(AAbyE)	AddbyAdd	18MCG A1byE_1	19MCG A1byE_2	20MCG A1byE_3	18MCG A2byE_1	19MCG A2byE_2	20MCG A2byE_3	18MCG AAbyE_1
chr2B_57183671	chr2D_29344035	5.67	1.59	4.08	1.79	0.45	1.34	0.26	0.37	-0.12	-0.25	-0.2	0.37	-0.17	-0.51

LOD= logarithm of the odds; PVE= Phenotypic Variation Explained; AA= Additive by additive effect; AAbyE= Epistasis-by-environment effect. AddbyAdd =Epistasis additive effect. A1 in the Additive by environment epistasis refers to the LeftMarker1 and A2 to the LeftMarker2. 18MCG, McGregor 2018; 19MCG, McGregor 2019; 20MCG, McGregor 2020.

4.4 Conclusion

The genetic variation for yield, yield components, and agronomical traits were evaluated at six locations of Central and High Plains of Texas for up to three years, giving a total of 11 environments. The phenotypic data collected for grain yield, test weight, thousand kernel weight, days to heading and plant height was used to identify genomic regions associated with favorable alleles for further use in marker-assisted selection. Increments in grain yield were found consistently correlated with test weight and thousand kernel weight; similarly, grain yield was negatively correlated with days to heading. Mixed results were obtained for correlation between grain yield and plant height, where increments in yield were positively correlated in some environments under rain-fed conditions and negatively correlated in irrigated ones. For the QTL analyses, each trait was analyzed by individual environment, across multiple environments and mega environments. These analyses permitted us to scrutinize additive effects of the QTL, additive-by-environment interactions, and gen- by-gene interactions as epistasis and epistasis-by-environment interactions. We found 86 QTL for the four traits, from which 38 were consistent as they were found in more than one environment or analysis. Among these consistent QTL, seven had a pleiotropic effect on more than one agronomical trait. These QTL were located at 57.16, 57.18 and 59.47 Mb on 2B, at 27.92, 29.34, and 40.64 Mb on 2D and 250.71 on 6D. A total of 25 epistatic interactions involved consistent QTL for all four traits. Amongst these, epistatic interactions between the consistent QTL located at 57.18 Mb on 2B and 29.3 Mb on 2D were observed for GYLD, DTH and TKW.

CHAPTER V DEVELOPING SUPERIOR PRE-HARVEST SPROUTING TOLERANT HARD WHITE WINTER WHEAT USING CRISPR-CAS9 GENE EDITING

5.1 Introduction

5.1.1 Grain color genes and the importance of Pre-Harvest Sprouting resistance

The grain color pigmentation of wheat (*Triticum aestivum* L.) is composed of catechin and proanthocyanidin pigments regulated by the flavonoid biosynthetic pathway (Himi et al., 2011). This pathway is controlled by three homologous Myb-type transcription factors (TF), known as Tamyb10, located in the chromosomes 3A, 3B, and 3D (Himi and Noda, 2005). Myb TFs are involved in the regulation of various functional genes under particular development and stress conditions, such as abscisic acid response, as well as interacting with other TF (Ambawat *et al*., 2013). With regards to the grain color, in order to obtain a white-grained variety, the three homologous must be recessive for the three alleles *R-A1a, R-B1a, and R-D1a,* since one or more dominant alleles (*R-A1b, R-B1b, or R-D1b*) would confer a red-grained wheat phenotype (Himi et al., 2011).

In Texas and other wheat production areas in the U.S. Great Plains, the Hard Red Winter Wheat (HRWW) class dominates in the number of acreages grown per the six classes. The Hard White Winter Wheat (HWWW) class adapted to Western Kansas and Eastern Colorado has been challenging to grow in Texas due to its predisposition to pre-harvest sprouting (PHS). Grain germination before harvest leads to poor end-use quality of the flour and renders wheat as feed grade. It starts by triggering the production of enzymes that breakdown the starch and protein to provide energy for the germinating embryo (Vetch *et al*., 2019). When starch is highly damaged it absorbs more water, affecting the dough mixing and end-use properties, impacting the final product. In HRWW, because of the high water-holding ability and porosity of the dough, the

dough gets stickier. Rheological properties also change, leading to a decrease in dough consistency and low bread volume, reducing the baking performance (Barrera *et al*., 2007). Seed color has been known to be associated with pre-harvest sprouting (caused by a lack of seed dormancy) in wheat due to the accumulated pigment in the grain which inhibits germination (Lin et al. 2016). Therefore, HWWW has a higher predisposition to PHS compared to HRWW. Nevertheless, the tight linkage between grain color (*Tamyb10-A, B,* and *D*) and PHS has been broken with various degrees of success (Kottearachchi et al., 2006; Shorinola et al., 2016). This association between red color and seed dormancy in weedy red rice (*Oryza sativa L*.) was suggested to be a pleiotropic effect (Gu et al., 2011). In wheat, the gene associated with seed dormancy is known as *taVp1* and has been identified in the long arm of the group of chromosomes 3A, 3B and 3D (Osa et al., 2003; Himi and Noda, 2004; Yang et al., 2007). *taVp1* is an orthologue of the maize transcription factor VIVIPAROUS-1 (*Vp1*), which plays a role in the induction and maintenance of dormancy (Osa et al., 2003). Other seed dormancy or pre-harvest sprouting resistance genes in wheat have been mapped to chromosomes 2, 3, 4 and 7 (Yang *et al*., 2007; Liu *et al*., 2013; Cabral *et al*., 2014; Zhang *et al*., 2014; Shorinola *et al*., 2016).

A definitive test would be to knock out the function of the three *Tamyb10* genes on chromosomes 3A, 3B and 3D to convert HRWW lines to HWWW lines. Subsequently, tests could be done to see if they still retain their superior tolerance to pre-harvest sprouting.

5.1.2 Gene editing using CRISPR/Cas9

In 2012, Clustered Regularly Interspaced Palindromic Repeat (CRISPR) technology exploded onto the scene with the discovery that CRISPR biology could be used for precise gene editing with the Cas9 protein and a single guide RNA (Doudna and Charpentier, 2014). The CRISPR/Cas9 system was discovered in bacteria and archaea that have developed an RNA mediated adaptive defense system to defend themselves from viruses and plasmid attack (Jinek et al., 2012). The CRISPR loci, also known as protospacers, are short fragments of foreign sequences that are integrated into the host chromosome at the proximal end of the CRISPR array and are responsible for recognizing when viral or plasmid infection occurs (Jinek et al., 2012). Jinek et al. (2012) identified that in order to create DNA breaks using the Cas9 protein, it was required to guide it with a single-guided RNA (sgRNA) with a length of 20 nucleotides (nt) along with a hairpin structure. The cleavage efficiency of the sgRNA is determined by the complementarity of the sequence with the DNA and an adjacent 3 nt region called protospacer adjacent motif (PAM) (Jinek et al., 2012). The exonuclease is a critical component required in the system. The most popular protein used is Cas9, although, different variants such as Cpf1 (Cas12a) have also been used. The differences between the usage and selection of the exonuclease lie in the availably of the specific PAM sequences in the target region, as well as the type of a double-strand break (DSBS). The Cas9 creates a DSBS at 3 base pairs upstream the PAM site 5'-NGG-'3, whereas Cpf1 recognizes 5'-TTV-'3 and creates staggered DNA DSBS genomes (Fagerlund et al., 2015). Also due to this last property, Cpf1 is preferably used in A/T rich genomes.

5.1.3 CRISPR/Cas9 as a breeding tool in plants

The CRISPR/Cas9 systems have become a promising technology to target genomic regions of interest and create precise modifications to introgress genes, gene validation or create novel allelic variation faster than conventional breeding (Chen et al., 2019). Gene or genome editing using CRISPR-Cas9 technology is quite different from genetic engineering that transfers genes from one organism to another and results in a genetically-modified organism (GMO). Genetically modified crops have transgenes, most often from other species, that insert randomly in the genome. In contrast, CRISPR-Cas9 editing can precisely modify the DNA sequence of an existing gene, either to knock out its function, thus deleting the gene entirely, or replace a specific allele at that locus with a similar allele from another genotype, although insertions are also possible (Chen et al., 2019). The resulting CRISPR-Cas9 edited lines are not considered transgenic *per se* when the Cas9 gene is transiently expressed, delivered into the cell as a protein, or is removed from the genome in subsequent generations (Zhang et al., 2016a). The edited lines then contain only the precise DNA modifications and are otherwise indistinguishable from genotypes that had obtained similar genome changes through natural mutations. This technology has been used in many plants, ranging from model to agronomical crops (Belhaj et al., 2018). With regards to agronomical crops, in rice mutations using CRISPR/Cas9 have been used to evaluate the effect of mitogen-activated protein kinase (MPK) genes, known to be essential for its development (Minkenberg et al., 2017). They found that mutations in the *MPK1* gene produced dwarf and sterile plants, whereas *MPK6* knockouts caused no seed development. Disease and bacterial resistance enhancement, grain weight improvement, photoperiod controlled male sterile line development, herbicide resistance enhancement, amongst other traits, are the other applications for genome editing in rice (Mishra et al., 2018). In maize, this technology has

been used in the industry to develop a reduced plant height type of corn by creating novel allelic variants of the brachytic gene (*br2*), using *Agrobacterium tumefaciens* (Bage et al., 2019). Several studies using gene editing have been conducted in wheat with the aim of gene validation or new allelic variation. Wang et al. (2014) induced mutations in the three homoeologs that encode the Mildew-Resistance Locus (*MLO*), obtaining transgenic plants with improved resistance to powdery mildew (*Blumeria graminis* f. sp. *Tritici* (*Bgt*)). Sánchez-León et al. (2017), developed non-transgenic wheat mutants with a significant reduction in gluten production by modifying the *α-gladin* gene family, encoded by the three homeologs located in the chromosomes 6A, 6B and 6D. CRISPR/Cas9 has also been used in wheat to facilitate the creation of sterile lines for hybrid wheat production. Singh et al. (2018) knocked-out the three homeologs of *TaMs45*, an ortholog of the male-sterile 45 (*Ms45*) gene. Their triple homozygous recessive mutant plants produce nonviable anthers with a shrunken shape that do not extrude as the wild-type counterpart. Another effort to enable the production hybrid wheat was made by Li et al. (2020), where they edited the three *TaNP1* homoeologs, another gene associated with male sterility which is the *OsNP1* and *ZmIPE1* orthologous in rice and maize, respectively. Contrary to the *TaMs45* study, their triple knockout of the *TaNP1* genes created mutants that did not even produce pollen, therefore displaying male sterility. These aforementioned studies are examples of how this technology can change plant breeding and facilitate the creation of lines that could be used for hybrid wheat production, attracting private companies and public breeding programs to invest in it. CRISPR/Cas9 technology is not limited to target only one gene or gene families, given that simultaneous editions of different genes have been accomplished in one round of transformation. Wang et al. (2018) developed multiple gene-editing methods in which multiple sgRNA were arrayed in tandem in a single construct. Their objective was to knock-out the

TaGW2 grain size gene, the homoeologs of *TaMLO* that provided resistance against powdery mildew and silenced the *TaLpx-1* to achieve resistance to *Fusarium graminearum*, obtaining multiple combinations of mutated genes. Multiplexed CRISPR/Cas9 complex has also been used to activate genomic regions that are selected due to methylation (Lowder et al., 2018). The *MPK* proteins previously mentioned study is another example of multiplex editing in agronomical crops (Minkenberg et al., 2017).

5.1.4 Delivery methods and tissue types for wheat transformation

The selection of the tissue type and delivery method relies mostly on the species of study, via either transgenic or non-transgenic approach, as well as the budget. For some crops such as lettuce (*Lactuca sativa L.*), *Arabidopsis thaliana* and rice, protoplast transformation and plant recovery can be routinely performed (Abdullah et al., 1986; Woo et al., 2015). However, in wheat and other monocots, protoplast transformation is merely one step in the gene-editing pipeline given the difficulties of regenerating plants out of them (Liang et al., 2017). Alternatively, immature embryo and callus transformation have been widely used for wheat since it is possible to recover a decent number of plants with a modest transformation efficiency (Hamada et al., 2018; Liang et al., 2018). Nevertheless, plant regeneration is still relatively low and highly genotype-dependent. Usually, modern elite cultivars are more difficult to regenerate. Some model cultivars, namely ‘Bobwhite’ and ‘Fielder’ are amenable to transformation and are known for their moderate regeneration efficiency (Zale et al., 2004; Hayta et al., 2019). In order to overcome this bottleneck, mature embryo transformation has been proposed as an alternative to eliminate the tissue culture steps as well as to speed up the development of plants to genotype and screen for potential mutants in a faster manner (Hamada et al., 2017).

Agrobacterium-mediated T-DNA transformation is one of the most popular methods to deliver the CRISPR/Cas9 construct. One downside associated with that is the increased chance of off-target changes and the inability to perform transgene-free editing (Demirer et al., 2019; Yi Zhang et al., 2016). Alternative delivery methods with transient expression of the CRISPR/Cas9 complex have been proposed to overcome this problem. Liang et al., 2017 put forward a method that uses preassembled CRISPR/Cas9 coated with ribonucleoproteins (RNP), which showed low on-target mutations at the expenses of a high reduction of off-target mutations. Liang et al. (2018) published the complete protocol for bread wheat transformation using RNPs. This protocol will be further described in other sections of this chapter since it was implemented in this experiment. Recently, alternative delivery methods that incorporate nanomaterial such as carbon nanotubes, have been postulated to increase transformation efficiency (Demirer et al., 2019). Compared to other methods such as *Agrobacterium*-mediated transformation, this method has a broader range of plant species and tissue type, moreover, it could be transgene-free editing. Furthermore, compared to the gene-gun, it does not inflict tissue damage since it does not require the use of high bombardment pressure and it is not restricted to the tissue/specimen size given the limited transformation chamber (Demirer et al., 2019). This methodology, using single-walled carbon nanotubes delivered in leaves, has been successfully tested on wheat, in which Green Fluorescence Protein (GFP) was used to validate the transformation efficiency (Demirer et al., 2019)

5.2 Research objectives

1. Design sgRNAs for multiplexed editing of the red seed color genes.
2. Validate the sgRNA *in vitro*.
3. Establish the tissue culture and transformation protocol.

4. Compare CRISPR-Cas9 protocols in wheat that will enable regeneration.

5.3 Materials and methods

5.3.1 Genetic material

'TAM 114' was used as the model variety for the Crop Genome Editing Lab to establish the tissue culture and transformation protocol in HRWW. It is one of the most widely grown varieties from the Texas A&M wheat breeding program. The resulting findings will provide a proof of concept and pave the way for resurrecting cultivars that lost resistance to pests or need slight modifications to add value to them.

5.3.2 Tamyb10 characterization and sgRNA design

Genomic DNA from the TAM 114 was extracted using the Cetyltrimethyl Ammonium Bromide (CTAB) modified protocol (Doyle & Doyle, 1990; Liu et al., 2013). The allelic variation for the three *Tamyb10* homeologs genes in TAM114 was obtained by amplifying the three exons using polymerase chain reaction (PCR) using the primers designed by Wang et al. (2016) (Table 5.1). The obtained amplicons were separated by 1% agarose electrophoresis gel, extracted and purified using the QIAquick PCR Purification Kit (QIAGEN®), and sent to the Laboratory for Genome Technology at Texas A&M University for sequencing using the Sanger sequencing method. To better understand the three *Tamyb10* homeologous, the transcript sequences from the reference genome were obtained using Gramene (http://www.gramene.org/). The three exons, as well as introns from the genome A, B and D, were taken from the transcripts TRIAE_CS42_3AL_TGACv1_195150_AA0645460, TraesCS3B02G515900, TraesCS3D02G468400, respectively. Raw sequencing data were trimmed in order to remove low-quality reads. Subsequently, the reads were aligned to its corresponding transcript from the reference genome using Benchling (https://www.benchling.com/). Following this, a comparison

of the exons across the three genomes was performed to find unique conserved regions for multiplexed editing.

The obtained conserved regions of the three exons were blasted in CRISPR direct (https://crispr.dbcls.jp/) and CRISPR-P 2.0 (http://crispr.hzau.edu.cn/CRISPR2/). CRISPR direct also uses the IWGSC 1.0 reference genome to determine how many potential target regions are present in the wheat genome and select highly specific sgRNA with minimum off-target regions. This finding was confirmed by BLASTing them to the Chinese Spring wheat reference genome (Appels et al., 2018). The proposed sgRNAs were acquired as CRISPRevolution sgRNA EZ kit from Synthego (https://www.synthego.com/).

After finding a highly specific sgRNA, the exon was re-sequenced using a new set of primers (Table 5.2) that targeted smaller regions (606-334 bp) and facilitated the insertion of a blunt-end PCR product into a plasmid vector. The re-sequencing was done to confirm the consensus. This time the PCR products from each genome were cloned into a plasmid inserted in *Escherichia coli* using Zero Blunt™ TOPO™ PCR Cloning Kit (user guide located on https://www.thermofisher.com). Five mutant colonies resistant to kanamycin (selective media) were selected and the region with the insertion was amplified using M13 primers (provided in the kit) in a PCR. The subsequent cleaning steps are as mentioned before. The purified DNA fragments were sent for Sanger sequencing and the results were used for a new consensus across genomes.

Table 5.1. Primers sequence, amplicons size and PCR conditions of the primers designed by Wang et al. (2016) used to characterize the Tamyb10 homeologous genes exons in TAM 114.

Primer name	Abbreviation	Genome	Primer sequence	Fragment size (bp)	Annealing temp. (°C)	Annealing temp. PCR worked (°C)
Tamyb10-A F1	A1	3A	ATGGCTGCTCCCAAAGCTCTCA	1948	63.6	55
Tamyb10-A R1			CGATGAGCTCCTCTTCGTCGTT		61.7	
Tamyb10-A F2	A2		AATCGCTGCGGTAAGAGCTG	1081	59.9	55
Tamyb10-A R2			CCTGAGCAAGAGGATGCTGC		60.8	
Tamyb10-A F3	A3		TCAAGAACTACTGGAACACC	540	54.1	57
Tamyb10-A R3			CGTATTTTACTGCACGTAAC		52.3	
Tamyb10-B F1	B1	3B	ATGGGGAGGAAACCATGCTG	445	59.2	55 D*
Tamby10-B R1			CCGGCAGCTCTTTCCGCAC		63.6	
Tamyb10-B F2	B2		AATCGGTGCGGAAAGAGCTG	1198	60.1	61 D*
Tamyb10-B R2			CCGTATCGGGCTGCTGCTC		62.2	
Tamyb10-B F3	B3		TGCCGGGGCGAACAGACAAT	515	63.9	55 D*
Tamyb10-B R3			TGTCACCCGGGCCATCAAAG		62.2	
Tamyb10-D F1	D1	3D	ATGGGGAGGAAGCCATGCTG	1420	61.4	57
Tamyb10-D R1			CGGTCACTGTTATCTGACGCTGGAT		64.4	
Tamyb10-D F1	D2		ATGGGGAGGAAGCCATGCTG	1629	61.4	62
Tamyb10-D R2			ACTGCTGCTCGTGCCCTCC		63.6	
Tamyb10-D F3	D3		GGGCGAACAGACAATGAGAT	630	57.3	55
Tamyb10-D R3			CTTTGTTTACAGCA CCAC		51	

Bp= base pairs, *D= dimethyl fulsoxide was needed to complete the reaction.

The cleavage efficiency of the design sgRNA was tested *in-vitro* digestion reaction (Jinek et al., 2012; Larson et al., 2013; Mali et al., 2013). Briefly, the sgRNA, Cas9 nuclease, nuclease-free water, and NEBbuffer 3.1 were mixed and incubated at 25°C for 10 min. DNA was added, the reaction was mixed and spun in a microfuge. The reaction was incubated at 37°C for two hours. Then, 1 µL of Proteinase K was added and the mixing step was repeated. Finally, the sample was

incubated at 56°C for 10 min and followed by a fragmented analysis. Reagents are described in Supplemental Table 5.

Table 5.2. Primers sequence, amplicons size and PCR conditions of the primers used to amplify the second exons of the *Tamyb10* homeologous genes in TAM 114.

Primer name	Genome	Primer sequence	Fragment size (bp)	Annealing temp. (°C)	Annealing temp. used for PCR (°C)
Tamyb10-A JV F2	3A	TGAGTGTAATGTGTTTTCTGAAACT	634	58.6	60
Tamyb10-A JV R2		CTCCTCCACGACCAAAGACC		61.6	
Tamyb10-B JV F2	3B	AGAGAGGGGCAGGGGATATG	617	62.4	58
Tamyb10-B JV R2		GAAAATGCGCATGGTACGGT		60.9	
Tamyb10-D JV F2	3D	GAGAGGCGCAGAGGGAATG	606	61.9	58
Tamyb10-D JV R2		GAAATGTGCACGATGCAAACG		60.6	

Bp= base pairs

5.3.3 Transformation protocol and delivery method

The transformation protocol and reagents preparation was performed following the protocol developed by Liang et al. (2018) and is briefly described in this section.

Before the transformation, TAM 114 plants were grown under greenhouse conditions. Spikes were tagged at the flowering stage and harvested 10 and 11 days after tagging. The harvested spikes were washed in 75 % ethanol for 1 min. The ethanol was discarded and they were then washed with 2.5% (wt/vol) sodium hypochlorite (NaClO) for 20 min. The spikes were then rinsed six times with ddH2O, the excess was wiped and the kernels were detached and washed. Immature embryos were obtained by dissecting the kernels under a microscope. The embryos

were then placed in a high osmotic medium (Supplemental Table 5) and incubated for 3-4 hours at room temperature. Simultaneously, mixtures of gold nanoparticles, Cas9 protein, sgRNA, ammonium acetate, and 2-propanol were prepared for shooting in the gene gun. The transformation was performed using the PDS1000/He particle bombardment system (Bio-Rad®) with 1100 psi Helium pressure.

The treated embryos (30 per petri dish) were placed into the recovery medium (Supplemental Table 6) at 23°C for 15 days under dark conditions one day after transformation. Immature embryos were screened for GFP expression in an Olympus® SZX10 microscope with fluorescent light. Developed calli were transferred to a regeneration medium (Supplemental Table 6) with a density of 15 per Petri dish and stored at 23°C for 30 days with a photoperiod of 16:8 hours light: dark. Plantlets were transferred to a Magenta GA-7 containing rooting medium (Supplemental Table 6) at a density of two/three plants per box at the same conditions for 30 days to harden. Passed this period, the plants were vernalized in a new rooting medium at 8°C with the same photoperiod for 10 weeks, more time than required to guarantee possible vernalization issues. The boxes were moved to a growth chamber at the Institute for Plant Genomics and Biotechnology growth chambers with a temperature of 10°C with increments of 2°C for two days to temper the plants. Two plantlets were then planted in a pot with garden soil and covered with a plastic bag to maintain humidity and protect for potentially drastic changes in humidity. Once the plants started growing, bags were removed and the temperature was maintained at 14°C for 12 hours with light and 10°C for 12 hours in the dark for a month. Then maximum and minimum temperatures were increased 4°C every week for two weeks and maintained at 24°C and 20°C until the plants were harvested. At the heading stage, a set of

spikes were covered with glassine bags to prevent cross-contamination due to physical contact or insects and subsequently harvested when they reached physiological maturity.

In the case of the transformation of mature embryos by targeting the exposed Shoot Apical Meristem (SAM), we followed the protocol described by Hamada et al. (2017). Briefly, the grains were washed and rinsed as previously mentioned. They were then placed in a Petri dish with wet filter paper for germination in an incubator at room temperature. When the grains started to sprout, the Shoot Apical Meristem (SAM) was excised to expose the apical meristematic tissue. The embryo was then excised from the grain and placed in an osmotic medium for incubation for three to four hours at room temperature. In this experiment, a plasmid containing the Yellow Fluorescent Protein (YFP) at a concentration of 825 ng/µL was used for transformation using the gene-gun. The mature embryos were transferred to recovery medium 24 hours after transformation and the subsequent steps were performed as in the previous experiment. Nevertheless, no plants were regenerated in this study.

5.3.4 DNA extraction and genotyping of treated plants

DNA extraction of the recovered plants was performed using the CTAB protocol using leaf tissue taken at the early developmental stage. For genotyping, the second exon of the B genome was amplified, PCR products were purified and sequenced as previously described. This genome was selected because mutated alleles would produce a whiter grain due to the dosage effect compared to A and D genomes. Mutated lines for the B genome would be genotyped for the other genomes later. In the case of plants recovered from the SAM transformation, PCR and YFP insertion amplification were performed using the KAPA3G plant PCR kit (https://sequencing.roche.com/en-us.html) which allows to PCR directly from the leaf tissue, skipping the step of obtaining genomic DNA.

5.4 Results and Discussion

5.4.1 Allelic variation of the Tamyb10 homoeologous genes in TAM 114 and sgRNA design

The primers developed by Wang et al. (2016) served to amplify the three exons in the three genomes, their amplification size and annealing temperatures varied (Table 5.1 and Figure 5.1). Figure 5.1 shows the specific binding of the primers and the lack of smear in most cases. The primers that amplify D2 did not work, although D1 and D2 had a primer in common that overlap; therefore, eight fragments were obtained. The primers also amplified part of the introns, which was another main reason why the second set of primers was designed.

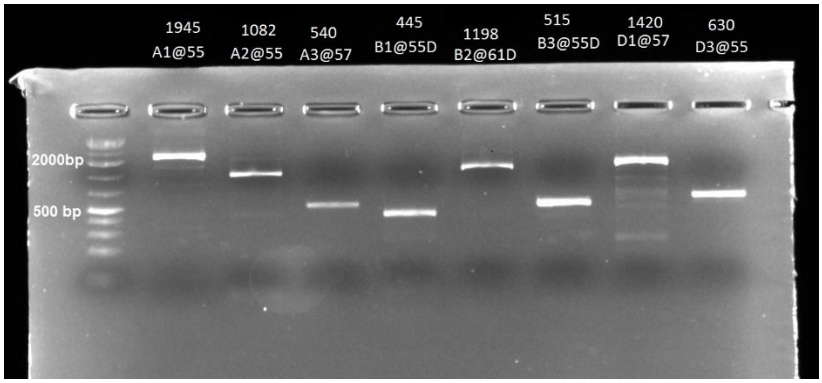


Figure 5.1. Agarose gel with the amplicons of the *Tamyb10* homeologous genes amplified by PCR using the primers designed by Wang et al. (2016). Bp= Fragment size in base pair; A1-D3 abbreviations described in table 1; @= annealing temperature; D= If dimethyl fulsoxide was needed to complete the reaction.

The three homeologous genes were partially constructed using the forward and reverse sequences from the three exons and were aligned to the transcript from the reference genome. Within these regions, minimal differences were found when comparing TAM 114 and the reference genome. The intragenic regions were trimmed, and a consensus of the exons across genomes was done to find a unique conserved region multiplex editing. There was no potential target region with the characteristics required for gene editing using CRISPR/Cas9 given multiple mismatches in the multiple alignments or the lack of adjacent PAM regions.

Nevertheless, there were potential genomic regions to be targeted in the second exon. The third exon was not examined for targeting regions because mutations at the end of the coding region of a gene are, in theory, less prone to change the phenotype as most of the protein was coded and its capacity may remain similar to the wild type.

The second exon was re-sequenced with a new set of highly specific primers that amplify a smaller region (606-634 bp) (Figure 5.2 and Table 5.2). This reduced fragment length facilitated the insertion of the blunt-end PCR product into the plasmid vector. The resequencing was performed in order to discard potential errors obtained from the PCR amplification from genomic DNA. After the second exon was inserted in a plasmid in *Escherichia coli*, DNA was extracted and sent for Sanger sequencing. Figures 5.3, 5.4 and 5.5, show the allelic variation among the five samples of DNA extracted from the plasmids for genome A, B and D. In these figures it is possible to see that there was minimal variation in the target region among the replicates, except for chromosome D in which two replications have cytosine. In contrast, the other three had adenine (Figure 5.5). This finding denotes the importance of conducting this step given that with the consensus obtained, we are more confident that the sgRNAs are well designed.

For the sgRNA design, the common region obtained from the multiple alignments was loaded in CRISPR direct and CRISPR-P 2.0. These databases of multiple sgRNA were not highly specific, given their similarity with a high number of targets across the wheat genome. After blasting the sgRNA generated in both databases, two common sgRNA (hereinafter referred to as sgRNA1 and sgRNA-2) (Table 5.3) were selected due to their lower number of off-targets. The sgRNA-1 off-targets are three homoeologous genes in chromosomes 5A, 5B and 5D with a similarity of 19

nucleotides. In the case of the sgRNA-2, it matches 17 nucleotides with two genomic regions in chromosomes 7B and 2B.

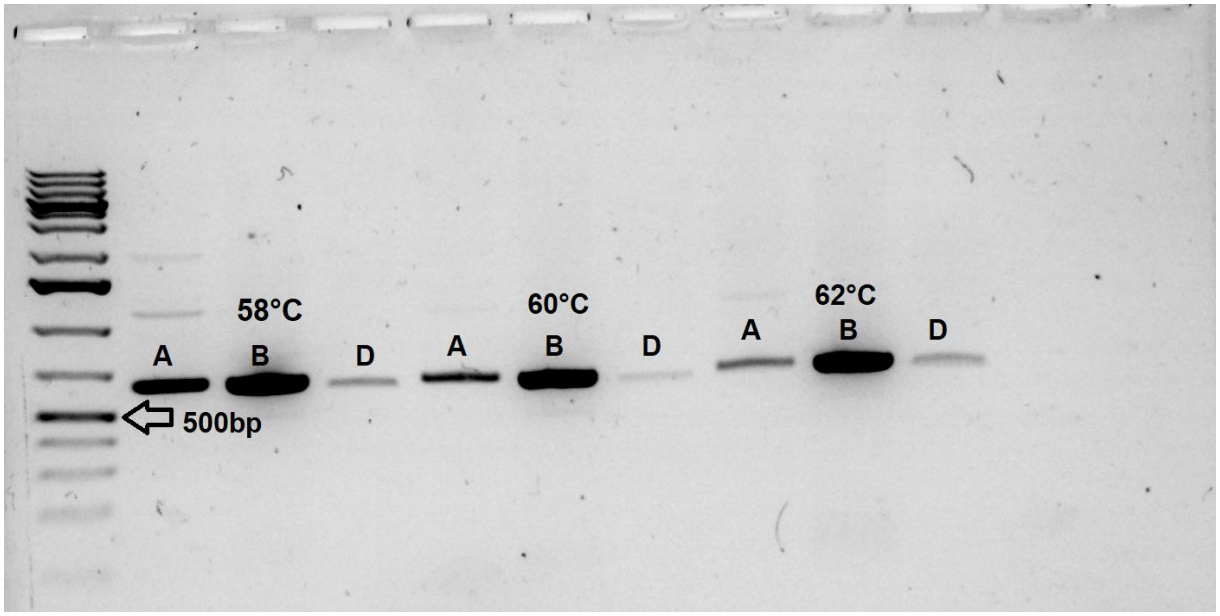


Figure 5.2. Agarose gel of the annealing temperature test for the second exon of the *Tamyb10* homeologous genes amplified by PCR using the second set of primers. A, B, and D = the second exon amplified from genome A, B and D, respectively.

An important step that ensures the cleavage efficiency of the sgRNA is the in-vitro digestion. This assay is described in S1. In our cleavage efficiency test, the sgRNA-2 shows some degree of cleavage activity, given two faint bands are seen in the gel, whereas some smear between the 100-200bp can be seen for the sgRNA-1 (Figure 5.7). It is essential to point out that the cleavage efficiency in this assay could be affected due to the shelf life expiration of both sgRNA. Nevertheless, the subsequent steps that will be described were performed immediately after receiving the sgRNA, and, therefore, their cleavage efficiency was not compromised.

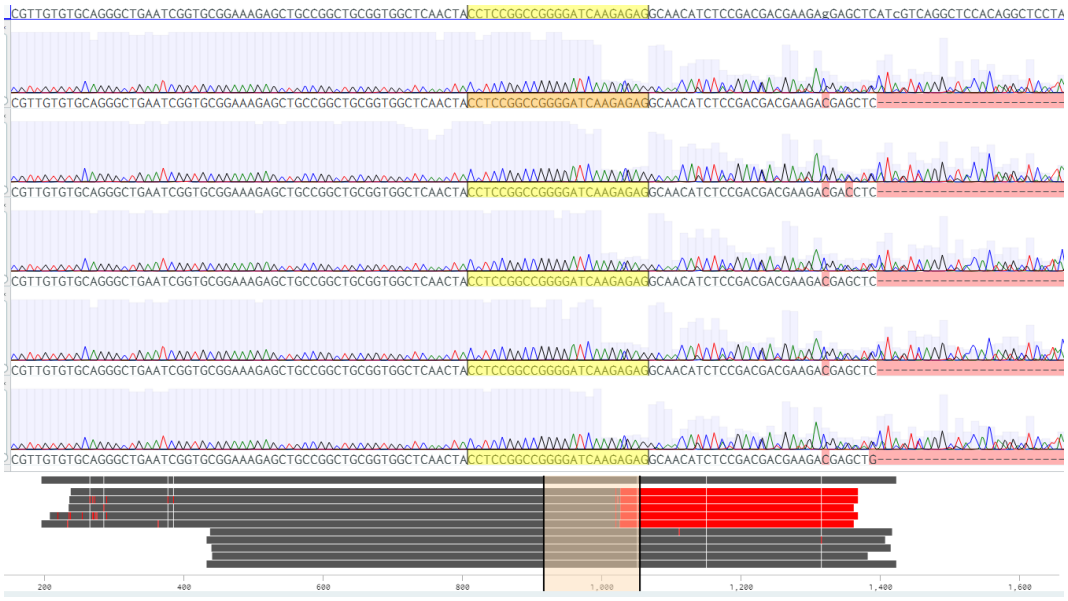


Figure 5.3. Sequences Consensus of five TAM 114 *Tamyb10*-A obtained from multiple alignments of DNA inserted into a *E. Coli*.

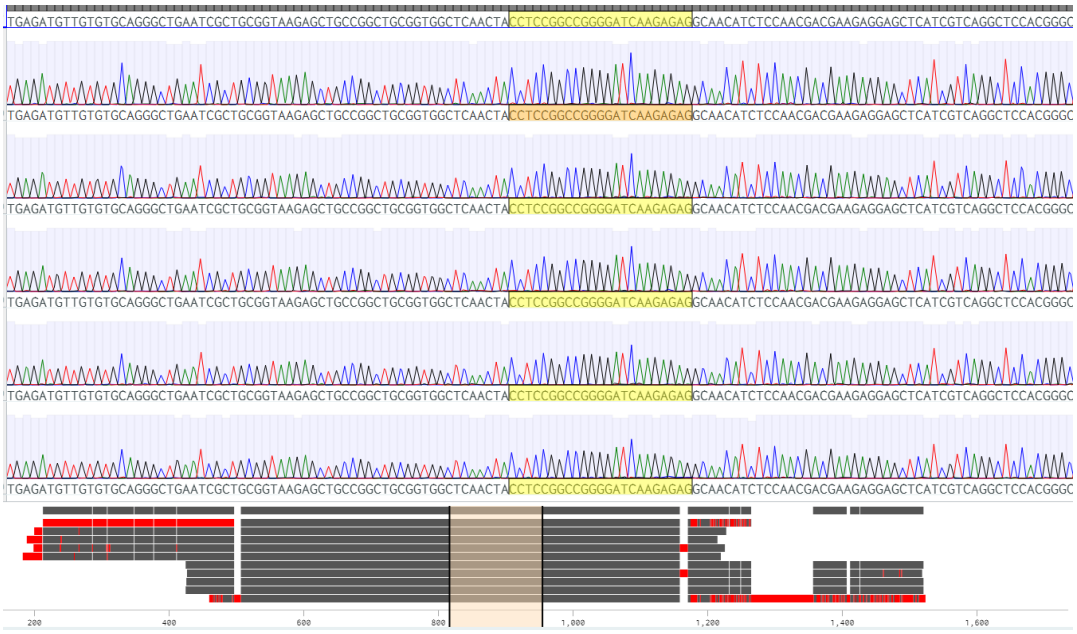


Figure 5.4. Sequences Consensus of five TAM 114 *Tamyb10*-B obtained from multiple alignments of DNA inserted into a *E. Coli*.

Figure 5.5. Sequences Consensus of five TAM 114 *Tamyb10*-D obtained from multiple alignments of DNA inserted into a *E. Coli*.

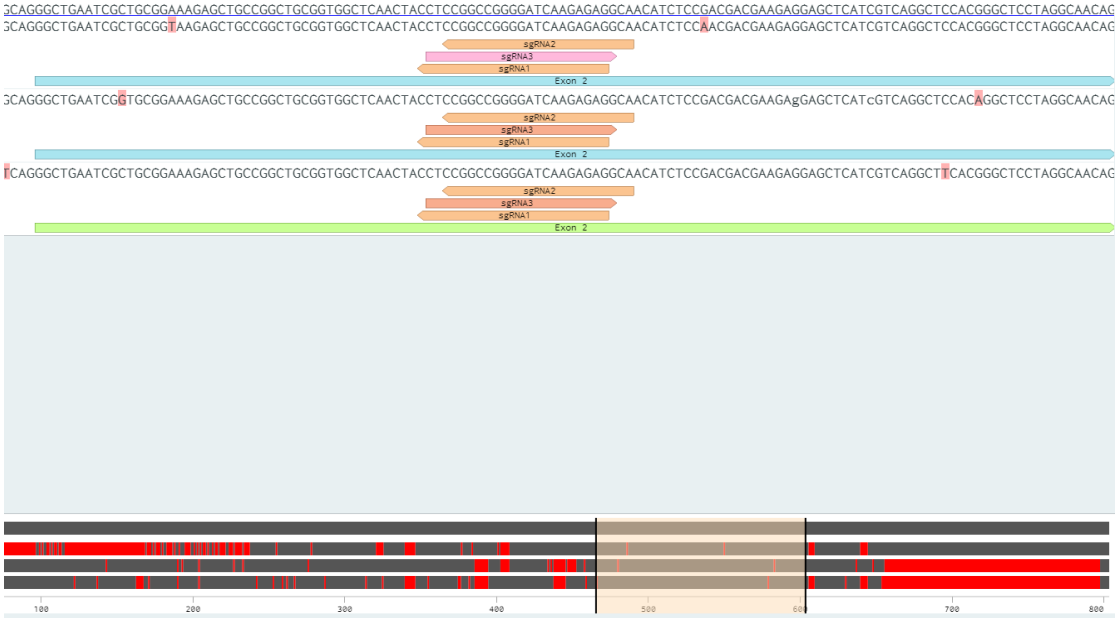


Figure 5.6. The consensus of TAM 114 *Tamyb10*-A, B, and D second exon multiple consensuses, and sgRNA target regions.

Table 5.3. sgRNAs sequence, target and strand sequence, and potential off-targets.

Name	Target sequence	sgRNA sequence	Strand	Off-target
SgRNA -1	5'-CTCTCTTGATCCCCGGCCGG**AGG**-3'	3'-**CCU**CCGGCCGGGGAUCAAGAGAG-5'	reverse	TraesCS5D02G234800, TraesCS5A02G227400, TraesCS5B02G226100. Chr 7B:657056032-657056048 & Chr 2B:118774754-118774770 bp
SgRNA -2	5'-TGCCTCTCTTGATCCCCGGC**CGG**-3'	3'-**CCG**GCCGGGGAUCAAGAGAGGCA-5'	reverse	

PAM region in bold blue. Chr = Chromomosome. Bp = base pairs.

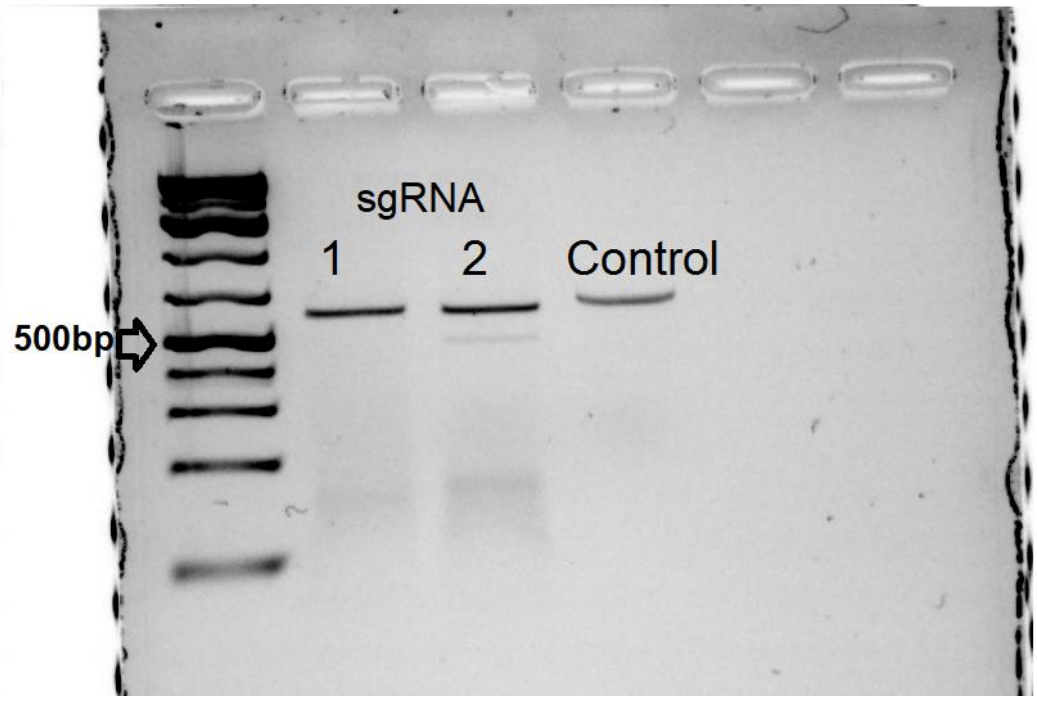


Figure 5.7. Agarose gel with the *in-vitro* digestion of the second exon of genome B using sgRNA1, sgRNA2, and the control.

5.4.2 Immature embryos transformation and tissue culture

The obtaining of immature embryos starts with harvesting the spikes after flowering. Liang et al. (2018) recommend harvesting the grains 12-14 days after flowering, but this step is highly genotype and environment-dependent. In our case, 10 and 11 days after flowering yielded good quality immature embryos (Figure 5.8). After kernels were detached from the spike and rinsed, immature embryos were excised, a total of 266 immature embryos were obtained with 144 and 122 for 10 and 11 days after flowering, respectively. Every immature embryo was placed in the center of a high osmotic medium for incubation in a density of 30 embryos per plate. Simultaneously, the mixture with the CRISPR/Cas9 complex and RNP was prepared with the concentrations, as suggested in the protocol. For this experiment, five combinations of flowering times and sgRNA were tested as in Table 5.4.

Figure 5.8. Excised immature embryo under the microscope before transformation.

One day after transformation, the treated embryos (30 per petri dish) were placed into the recovery medium. During this stage, embryos from treatments one and four were monitored for the expression of GFP under the microscope. However, we were not able to detect it. We hypothesize that the protein was not expressed in the tissue given the promoter (OsU3) was not

appropriate for wheat. The auxin in the medium favored the induction of calli at a high rate (Figure 5.9A). Once the calli were in the regeneration medium and under light conditions, green tips started to develop in some calli. Nevertheless, this process was not simultaneous in all of the calli. One-third of the calli developed plantlets two weeks after and were transferred to a Magenta GA-7 containing rooting medium, at a density of two/three plants per box. This process allowed the plantlets to harden before vernalization since they were never exposed to cold temperatures before. The plantlets were 10 weeks, more time than required to avoid possible vernalization issues. The remaining two-thirds were transferred to a new regeneration medium until plantlets started developing (~30 days) and continued the tissue culture process, as with the other batch. This sequential order of how the plantlets were transferred continued until the plants were harvested.

Table 5.4. Treatments information, immature embryos transformed and total number of regenerated plants.

Treatment No.	Treatment	No. of immature embryos	No. of plant regenerated
1	10 days + GFP plasmid + sgRNA-1	27	6
2	10 days + sgRNA-1	59	9
3	10 days + sgRNA-2	58	7
4	11days + GFP plasmid	73	23
5	11 days + sgRNA-1	49	18
Total		266	63

10 and 11 days = refers to the number of days after flowering when the immature embryos were excised. GFP plasmid = plasmid DNA containing the with Green Fluorescence Protein transgene. One shot per treatment.

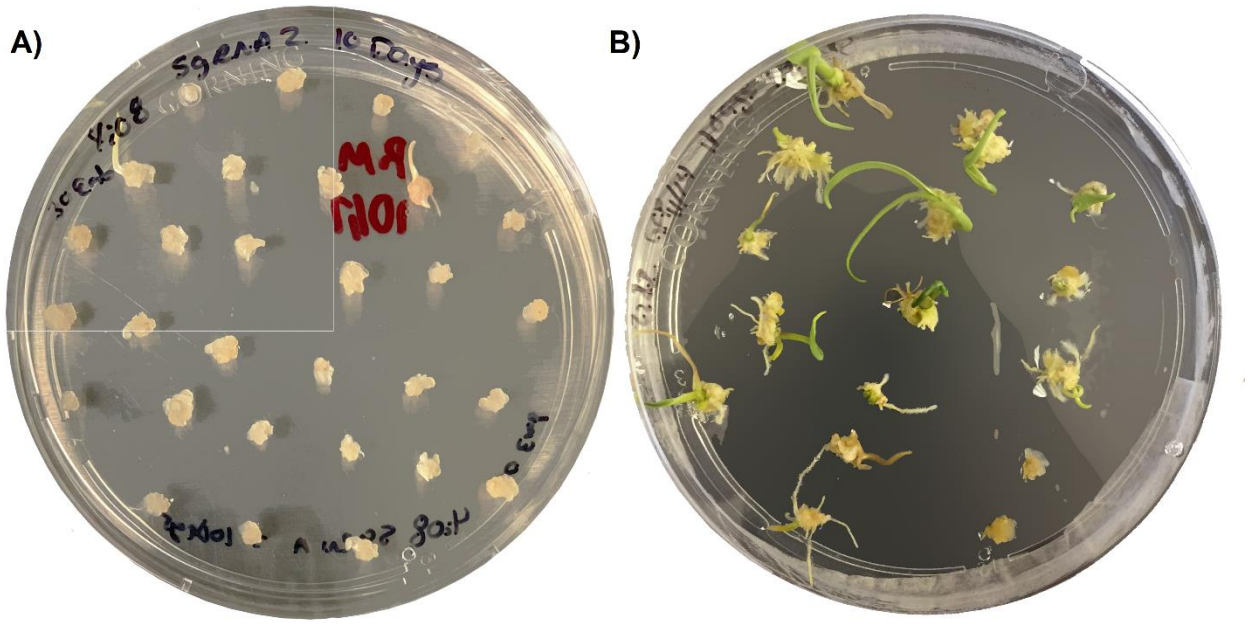


Figure 5.9. Calli formed in recovery medium (A) and Green tips formed in calli under regeneration medium (B).

After the vernalization finished, every box was moved to a growth chamber at the Institute for Plant Genomics and Biotechnology. The plantlets were planted in a density of two per pot to reduce inter-plant competition and favor their development. In addition to that, the selected temperature and humidity conditions permitted an excellent establishment of the experiment (Figure 5.10A), given only three plants died during the transferring process. Figure 4.10B shows an excellent adaptation of the recovered plants. Due to the significant development of tillers per plant, the spikes were covered with glassine bags to prevent potential cross-contamination, given the high density in the growth chamber. During harvesting, the grains were examined for phenotypic changes; however no variation was observed in the grain color.

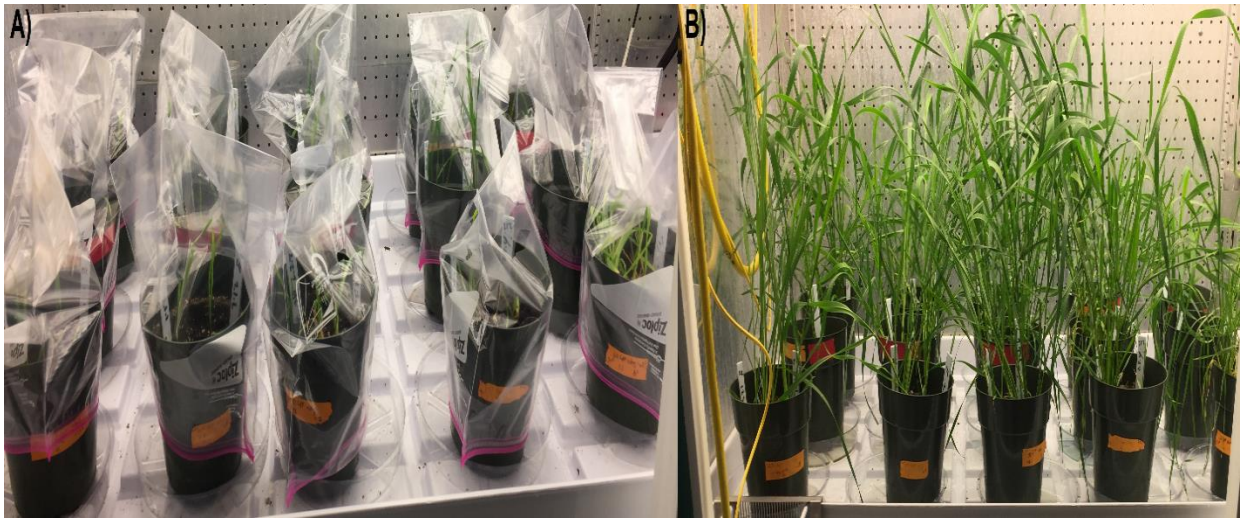


Figure 5.10. Regenerated plants protected with bags to maintain humidity (A). Plants at growing stage Feekes 9 (B).

Table 5.4 also shows that a higher number of plants were regenerated when the immature embryos were obtained from spikes harvested 11 days after heading, with efficiencies of more than 30 %. For 10 days after heading, it varied from 12-22 %. For treatment 3, a petri dish with 30 embryos was discarded due to fungal contamination. Although three plants failed to adapt from tissue culture conditions to the growth chamber, their DNA was collected and will be genotyped to test the transformation efficiency of TAM 114. In total, a set of 60 plants were successfully regenerated (Supplemental 7). As a result, TAM 114 showed an acceptable regeneration efficiency; however, not as high Fielder or Bobwhite (Zale et al., 2004; Hayta et al., 2019).

5.4.3 Genotyping of the regenerated plants

The primers that amplified the B genome described in Table 4.2 were used to amplify the target region in the regenerated plants. Initially, to reduce the genotyping cost, PCR products of four plants treated with CRISPR/Cas9 were pooled into one sample in a total of 10 pools, as well as

an extra pool of four plants treated with only plasmid DNA (Supplemental Table 7). These eleven pools were sent for Sanger sequencing but no results were obtained despite we followed the same procedure and using adequate quality DNA, similar to previous sequencing submissions (Figure 5.11).Therefore, to solve this problem, we decided to use a single plant as representative of each pool and send it to sequencing, however, the result was the same.

We initially hypothesized that there was allelic variation in each plant's genomic DNA due to the treatments and therefore, the Sanger sequencer could not generate a consistent call. However, the controls (only treated with plasmid DNA) could not be sequenced either. The next step was to recur to next-generation sequencing, which given its robustness would allow us to get an idea of the genetic variation present.

Figure 5.11. Amplicons of the Tamyb-10 second exon of eleven pools of four plants.
P= refers to the pool of PCR products number described in Supplemental Table 7.

5.4.5 Shoot apical meristem transformation in TAM 114

In this project, we also explored different alternative tissues for transformation. The previously described experiment is labor-intensive and the bottleneck of the process is the tissue culture steps. Hence, we decided to test the protocol of Hamada et al. (2017) in TAM 114. In the first step of mature embryo preparation, this variety showed reasons why its genetic background was selected to establish the transformation protocol. Compared to what was tested in the protocol, TAM 114 took only 48 hours to germinate, proving its good PHS resistance. Before the transformation, we tested the regeneration of green tissues and plants using the previously used medium to see whether they would grow. The coleoptiles were removed and the SAM were exposed in imbibed seeds (Figure 5.12A). The mature embryos were excised from the grain (Figure 5.12B) and placed upright in the growing media to test the plant regeneration capacity. We found plantlets were successfully regenerated out of the mature embryos with SAM exposed using a regeneration medium. Also, we tested the number of shots required for transformation using biolistics, with two and four shots. Similarly, as in the other experiment, the mature embryos with excised SAM were placed in a high osmotic medium for incubation before transformation to enhance transformation efficiency. YFP expression was detected in both shot rate treatments 5 hours after transformation (Figure 5.13A and B). The mature embryos were transferred to a regeneration medium to develop plantlets and further evaluate the expression of YFP during the process. Mature embryos were selected to keep track of the YFP expression in the excised region.

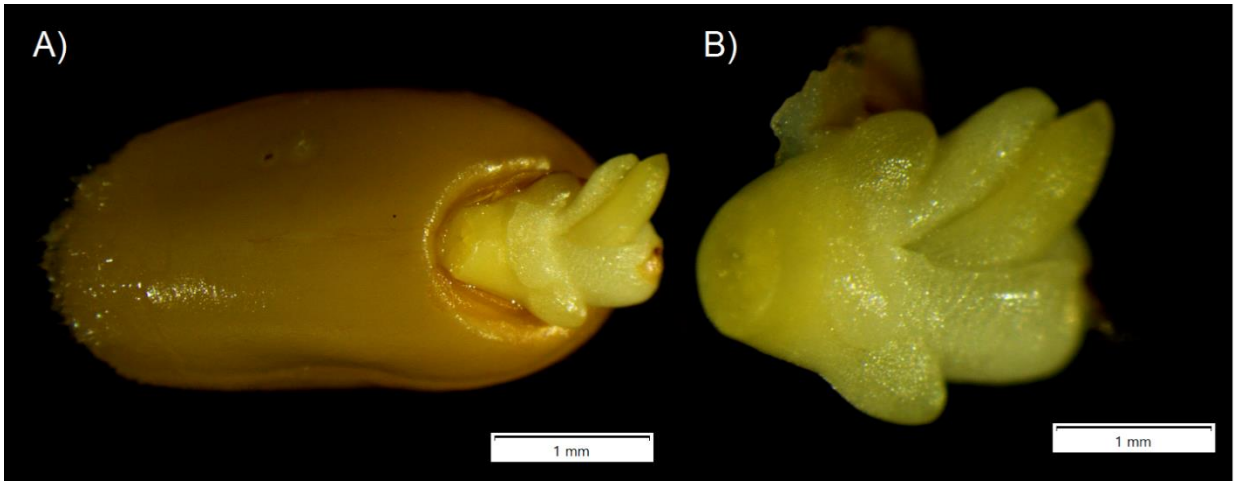


Figure 5.12. TAM 114 germinated seed with Shoot Apical Meristem (SAM) exposed (A) and excised mature embryo with SAM exposed (B).

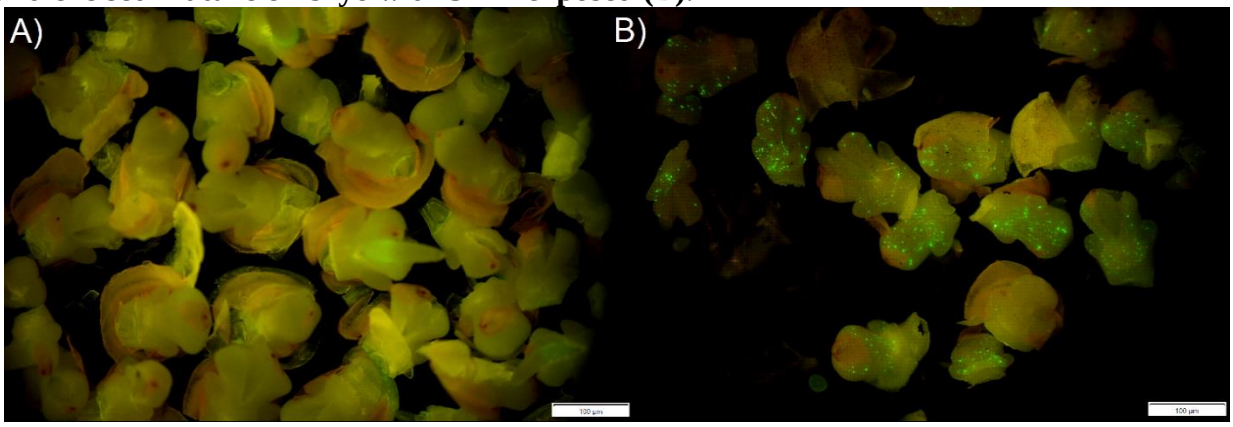


Figure 5.13. Mature embryos with exposed Shoot Apical Meristem (SAM) after transformation, control (A), and transformed expressing yellow fluorescent protein (B).

Fluorescent protein expression remained expressed in the embryos up to two weeks after transformation (Figure 5.14B). As the apical meristem grew, it was possible to observe that the cells developed in it also continued to express the fluorescent protein when multiplying (Figure 5.15). Hence, we hypothesized the possibility that the expression of YFP was due to the insertion of the transgene in the genome.

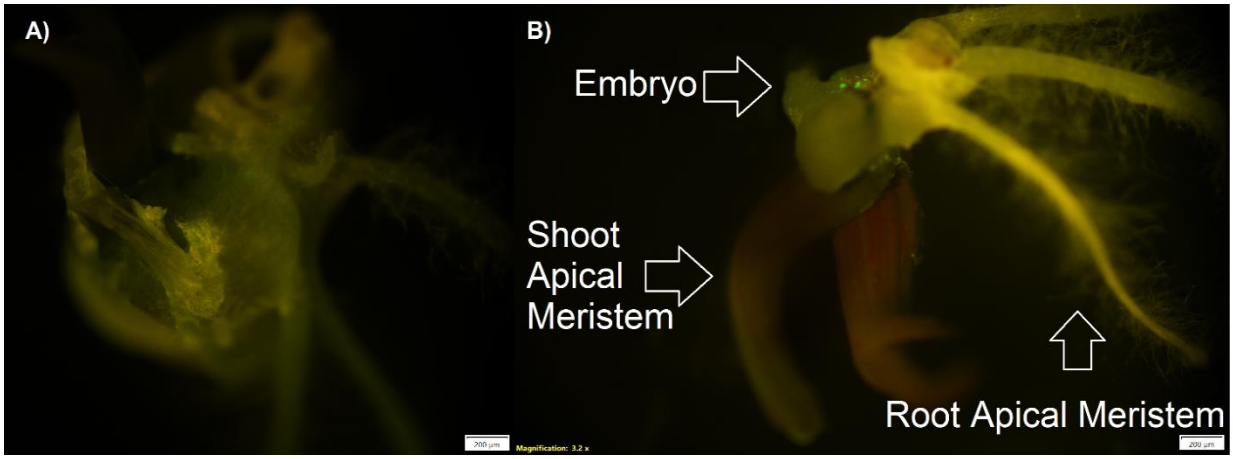


Figure 5.14. Plantlets developed from mature embryos after two weeks of transformation control (A) and transformed expressing yellow fluorescent protein (B).

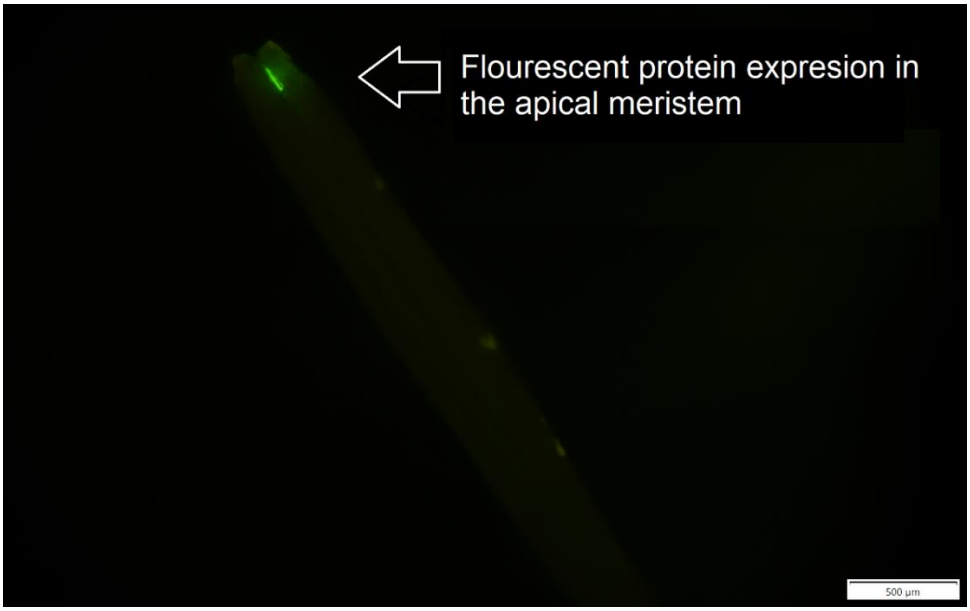


Figure 5.15. Yellow fluorescent protein expression in the apex a growing meristem.

Therefore, the PCR test was performed in order to amplify the transgene. For this, the plants were transferred to a rooting medium and were allowed to grow until they reach a considerable developmental stage. After that, the transgene amplification test was carried out using fresh tissue from two plants, one from each shot treatment, which presented YFP in the meristematic tissue, using TAM 114 as a negative control and positive control (plasmid DNA). The results of the PCR test with two repetitions demonstrated the positive control showed a fragment consistently of 230-250bp (expected size 232bp) (Figure 5.16A and B). Regarding the negative control, for the first replication, a long-faded smear with no clear band was obtained but in the second one, the sample seemed to be contaminated. For the treated plants, plant 1 (P1) and plant 2 (P2) appeared to have similarly obtained fragments ~ 2000bp. However, these were false signals given that they did not match the size of the positive control. The third replication of transgene amplification was not possible due to the plants dying after multiple tissue cuts, contamination or chlorosis.

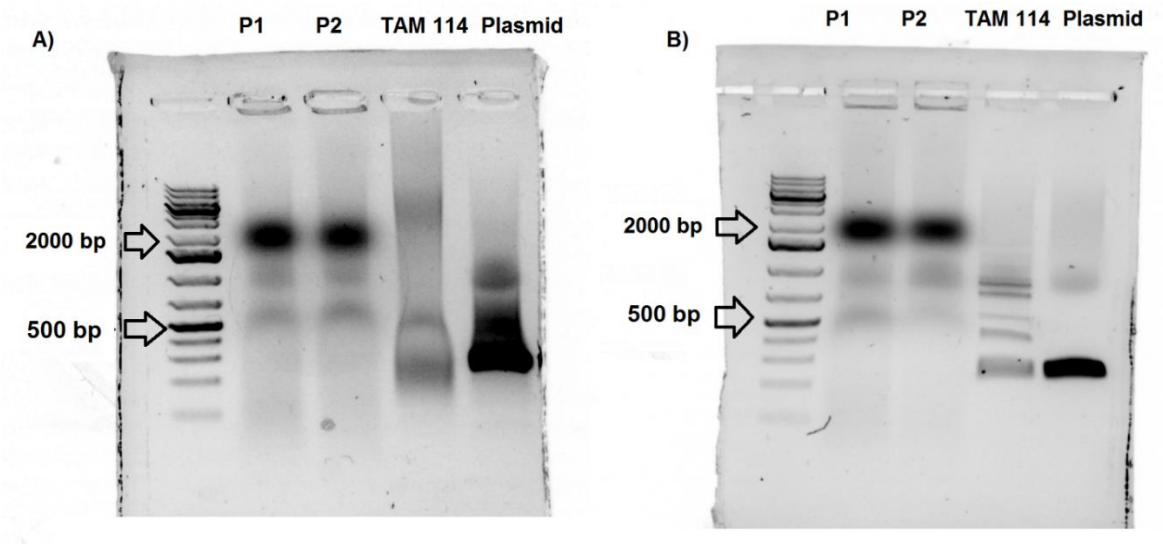


Figure 5.16. Replication one (A) and two (B) of the amplification of the transgene in two treated plants (P1 and P2), TAM 114 (negative control) and plasmid DNA (positive control).

5.5 Conclusion

We evaluated the genetic variations in TAM 114 Tamyb10 homoeologous genes to induce a precise modification in order to develop a white grain isoline. There was no conserved region in the first exon with the characteristics required to design a sgRNA; however, we found regions in the second exon, for which two highly specific sgRNA were designed for multiplexed editing. The biolistic transformation of immature embryos using the CRISPR/Cas9 complex coated with ribonucleoproteins was performed. TAM 114 proved to be an excellent choice for the establishment of the transformation protocol, given its modest regeneration efficiency. Nevertheless, its mutant production efficiency is yet to be proven. Alternative transformation tissues proved to be an alternative to avoid tissue culture and speed up the development of isolines or candidate gene validation. The establishment of the protocol for wheat transformation in SAM significantly reduces the time and labor associated with tissue culture, permitting screening for mutant plants in a shorter period of time. These steps further establish the transformation and tissue culture protocol in TAM 114 will permit the development of superior isolines that meet the current grain yield performance and end-use quality requirements. In contrast to model cultivars as Bobwhite and Fielder with poor agronomical importance.

CHAPTER V

SUMMARY AND CONCLUSION

This dissertation focuses on two molecular breeding techniques that have been widely used in plant breeding. In the first three chapters, the major objectives were to map QTL associated with favorable alleles for insect resistance and agronomical trait using a bi-parental population. The reliability of the high-density genetic map constructed was validated by mapping highly heritable traits. The awns suppressor *B1* and the glume color *Rg1* genes were mapped in the same genomic region where were previously described. In chapter II, QTL associated with insect resistance were identified. For Hessian fly, two QTL explaining 12.69-15.3 % of the phenotypic variation were detected on chromosome 1A and 3B. Within our QTL flanking region, two major genes are known as *H9* and *H35*, respectively, have been identified. A high confidence gene that belongs to the resistance kinase *Lr-10* related family is a promising candidate for the QTL identified on 1A. For Greenbug, a QTL explaining up to 88 % of the phenotypic variation was identified in Chromosome 7D. The major gene, *Gb3* has been identified within our QTL and it is known to be present in Texas A&M wheat germplasm.

In chapter III, the QTL analysis showed a moderate number of QTL in the single environment analysis, across environments and mega-environments for grain yield, days to heading, plant height, test weight and thousand kernel weight. Among the detected QTL 38 were consistent, elucidating that those genomic regions affected the trait in more than one environment. Ten consistent QTL were detected for grain yield and were found on chromosome 2A, 2B, 2D, 3D, 5A and 6A. Among these QTL, the one located at 29 Mb on 2D had the highest coefficient of determination and additive effects with 38.31 % and 40 g m^{-2}, respectively. This QTL was inherited from Iba. In the case of days to heading, four consistent QTL were identified

chromosomes 2B, 2D and 5B, also all of them were found in at least three environments, denoting that they were affecting the plant phenotype regarding the environment. The additive effect of the QTL for Days to heading varied from 1.59-6.2 days. Eleven consistent QTL were identified on chromosome 2A, 2B, 4A, 4B, 4D, 5A, 6A and 6D for plant height. The additive effect of these QTL varied from 1.2-2.45 cm. Eight consistent QTL on chromosome 1A, 2B, 2D, 4A, 5A and 5B were detected for test weight, however, only one was detected in more than one single environment analysis. This QTL detected in two environments and across environments had the larger additive effect with increments associated with it in an order of 12.80 kg/m^{-3}. Lastly, for thousand kernel weight, no consistent QTL was detected when considering individual environments, however, three of them were detected in the across environment analysis. These consistent QTL coefficient of determination and additive effects ranged from 16.09-29.65% and 1.4-2.5 gr, respectively.

Furthermore, the QTL analyses revealed seven genomic regions on chromosomes 2B, 2D, 5A and 6D affecting more than one trait. This suggests that there is either a pleiotropic effect of the gene within the QTL region or a linkage between multiple genes. The QTL located on at 29 Mb on 2D was found to be affecting all five evaluated traits. Besides, we found that the *B1* QTL was pleiotropic to test weight, similar findings were obtained in other studies where the presence of awns was associated with increments in test weight and grain size. The large number of epistatic interactions between multiple genomic regions and even consistent QTL revealed the complex genetic architecture of quantitative traits evaluated.

The fourth chapter focused on the development of a white grain isoline of TAM 114 using gene-editing. This project also paved the way for the establishment of the CRISPR/Cas9 transformation protocol for further use in the Texas A&M wheat program given could be used

for the creation of new allelic variation in the wheat germplasm, by knocking-out or inserting genes. In this experiment, we designed two highly specific sgRNA that target the second exon of the three homeologous Tamyb10 genes. Followed, we delivered the CRISPR/CAS9 system into immature embryos using biolistics. Subsequently, the treated immature embryos were subjected to different growing media to develop plantlets. TAM 114 showed to be a good choice for transformation due to its modest regeneration efficiency. However, the transformation efficiency has not been tested yet despite the multiple genotyping attempts. The mature embryos protocol showed to be an excellent choice for transformation given significantly reduces the development of plant time by eliminating some tissue culture steps. This established transformation protocol could be used to validate candidate genes identified in QTL analysis as the ones described in the previous chapters.

REFERENCES

Abdullah, R., E.C. Cocking, and J.A. Thompson. 1986. Efficient Plant Regeneration from Rice Protoplasts Through Somatic Embryogenesis. Nat. Biotechnol. 4(12): 1087–1090. doi: 10.1038/nbt1286-1087.

Alexandratos, N., and J. Bruinsma. 2012. World agriculture towards 2030/2050: The 2012 revision. ESA Work. Pap. No.12-03. doi: 10.1002/jso.2930300113.

Alvarado, G., F.M. Rodríguez, A. Pacheco, J. Burgueño, J. Crossa, M. Vargas, P. Pérez-Rodríguez, and M.A. Lopez-Cruz. 2020. META-R: A software to analyze data from multi-environment plant breeding trials. Crop J. 8(5): 745–756. doi: 10.1016/j.cj.2020.03.010.

Ambawat, S., P. Sharma, N.R. Yadav, and R.C. Yadav. 2013. MYB transcription factor genes as regulators for plant responses: An overview. Physiol. Mol. Biol. Plants 19(3): 307–321. doi: 10.1007/s12298-013-0179-1.

Anderson, S.L., A.L. Mahan, S.C. Murray, and P.E. Klein. 2018. Four Parent Maize (FPM) Population: Effects of Mating Designs on Linkage Disequilibrium and Mapping Quantitative Traits. Plant Genome 11(2): 0. doi: 10.3835/plantgenome2017.11.0102.

Appels, R., K. Eversole, C. Feuillet, B. Keller, J. Rogers, N. Stein, C.J. Pozniak, F. Choulet, A. Distelfeld, J. Poland, G. Ronen, O. Barad, K. Baruch, G. Keeble-Gagnère, M. Mascher, G. Ben-Zvi, A.A. Josselin, A. Himmelbach, F. Balfourier, J. Gutierrez-Gonzalez, M. Hayden, C.S. Koh, G. Muehlbauer, R.K. Pasam, E. Paux, P. Rigault, J. Tibbits, V. Tiwari, M. Spannagl, D. Lang, H. Gundlach, G. Haberer, K.F.X. Mayer, D. Ormanbekova, V. Prade, T. Wicker, D. Swarbreck, H. Rimbert, M. Felder, N. Guilhot, G. Kaithakottil, J. Keilwagen, P. Leroy, T. Lux, S. Twardziok, L. Venturini, A. Juhasz, M. Abrouk, I. Fischer, C. Uauy, P.

Borrill, R.H. Ramirez-Gonzalez, D. Arnaud, S. Chalabi, B. Chalhoub, A. Cory, R. Datla, M.W. Davey, J. Jacobs, S.J. Robinson, B. Steuernagel, F. Van Ex, B.B.H. Wulff, M. Benhamed, A. Bendahmane, L. Concia, D. Latrasse, M. Alaux, J. Bartoš, A. Bellec, H. Berges, J. Doležel, Z. Frenkel, B. Gill, A. Korol, T. Letellier, O.A. Olsen, H. Šimková, K. Singh, M. Valárik, E. Van Der Vossen, S. Vautrin, S. Weining, T. Fahima, V. Glikson, D. Raats, H. Toegelová, J. Vrána, P. Sourdille, B. Darrier, D. Barabaschi, L. Cattivelli, P. Hernandez, S. Galvez, H. Budak, J.D.G. Jones, K. Witek, G. Yu, I. Small, J. Melonek, R. Zhou, T. Belova, K. Kanyuka, R. King, K. Nilsen, S. Walkowiak, R. Cuthbert, R. Knox, K. Wiebe, D. Xiang, A. Rohde, T. Golds, J. Čížkova, B.A. Akpinar, S. Biyiklioglu, L. Gao, A. N'Daiye, J. Číhalíková, M. Kubaláková, J. Šafář, F. Alfama, A.F. Adam-Blondon, R. Flores, C. Guerche, M. Loaec, H. Quesneville, A.G. Sharpe, J. Condie, J. Ens, R. Maclachlan, Y. Tan, A. Alberti, J.M. Aury, V. Barbe, A. Couloux, C. Cruaud, K. Labadie, S. Mangenot, P. Wincker, G. Kaur, M. Luo, S. Sehgal, P. Chhuneja, O.P. Gupta, S. Jindal, P. Kaur, P. Malik, P. Sharma, B. Yadav, N.K. Singh, J.P. Khurana, C. Chaudhary, P. Khurana, V. Kumar, A. Mahato, S. Mathur, A. Sevanthi, N. Sharma, R.S. Tomar, K. Holušová, O. Plíhal, M.D. Clark, D. Heavens, G. Kettleborough, J. Wright, B. Balcárková, Y. Hu, N. Ravin, K. Skryabin, A. Beletsky, V. Kadnikov, A. Mardanov, M. Nesterov, A. Rakitin, E. Sergeeva, H. Kanamori, S. Katagiri, F. Kobayashi, S. Nasuda, T. Tanaka, J. Wu, F. Cattonaro, M. Jiumeng, K. Kugler, M. Pfeifer, S. Sandve, X. Xun, B. Zhan, J. Batley, P.E. Bayer, D. Edwards, S. Hayashi, Z. Tulpová, P. Visendi, L. Cui, X. Du, K. Feng, X. Nie, W. Tong, and L. Wang. 2018. Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science (80-.). 361(6403). doi: 10.1126/science.aar7191.

Assanga, S.O., M. Fuentealba, G. Zhang, C. Tan, S. Dhakal, J.C. Rudd, A.M.H. Ibrahim, Q. Xue, S. Haley, J. Chen, S. Chao, J. Baker, K. Jessup, and S. Liu. 2017. Mapping of quantitative trait loci for grain yield and its components in a US popular winter wheat TAM 111 using 90K SNPs. : 1–21. doi: 10.1371/journal.pone.0189669.

Azhaguvel, P., J.C. Rudd, Y. Ma, M.C. Luo, and Y. Weng. 2012. Fine genetic mapping of greenbug aphid-resistance gene Gb3 in Aegilops tauschii. Theor. Appl. Genet. 124(3): 555–564. doi: 10.1007/s00122-011-1728-z.

Bage, S.A., T.J. Barten, A.N. Brown, J.H. Crowley, M. Deng, R. Fouquet, J.R. Gomez, T.W. Hatton, J.C. Lamb, J.R. LeDeaux, B.M. Lemke, S. Manjunath, M.S. Marengo, E.Y. Morales, M.O. Garcia, J.M. Peevers, J.L. Pellet, A.R. Avendano, L.A. Rymarquis, K. Sridharan, M.F. Valentine, D.H. Yang, and E.J. Cargill. 2019. Genetic characterization of novel and CRISPR-Cas9 gene edited maize brachytic 2 alleles. Plant Gene (January): 100198. doi: 10.1016/j.plgene.2019.100198.

Barrera, G.N., G.T. Pérez, P.D. Ribotta, and A.E. León. 2007. Influence of damaged starch on cookie and bread-making quality. Eur. Food Res. Technol. 225(1): 1–7. doi: 10.1007/s00217-006-0374-1.

Belhaj, K., A. Chaparro-Garcia, S. Kamoun, N.J. Patron, and V. Nekrasov. 2015. Editing plant genomes with CRISPR/Cas9. Curr. Opin. Biotechnol. 32: 76–84. doi: 10.1016/j.copbio.2014.11.007.

Bennett, D., A. Izanloo, M. Reynolds, H. Kuchel, P. Langridge, and T. Schnurbusch. 2012. Genetic dissection of grain yield and physical grain quality in bread wheat (Triticum aestivum L.) under water-limited environments. Theor. Appl. Genet. 125(2): 255–271. doi: 10.1007/s00122-012-1831-9.

Borlaug, N.E. 1968. Wheat Breeding and its Impact on World Food Supply. Third Int. Wheat Genet. Symp.: 1–36. http://libcatalog.cimmyt.org/download/borlaug/66179.pdf.

Braun, H.-J., S. Rajaram, and M. Ginkel. 1996. CIMMYT's approach to breeding for wide adaptation. Euphytica 92(1–2): 175–183. doi: 10.1007/BF00022843.

Browning, B.L., and S.R. Browning. 2016. Genotype Imputation with Millions of Reference Samples. Am. J. Hum. Genet. 98(1): 116–126. doi: 10.1016/j.ajhg.2015.11.020.

Burd, J.D., and D.R. Porter. 2006. Biotypic diversity in greenbug (Hemiptera: Aphididae): characterizing new virulence and host associations. J. Econ. Entomol. 99(3): 959–965. doi: 10.1603/0022-0493-99.3.959.

Burton, R.L. 1986. Effect of greenbug (Homoptera: Aphidae) damage on root and shoot biomass of wheat seedlings. J. Econ. Entomol. 79: 633–636.

Byers, R.A., and R.L. Gallun. 1972. Ability of Hessian fly to stunt winter wheat. 1. Effect of larval feeding on elongation of leaves. J Econ Entomol 65: 955–958. doi: 10.1093/jee/65.4.955.

Cabral, A.L., M.C. Jordan, C.A. McCartney, F.M. You, D.G. Humphreys, R. MacLachlan, and C.J. Pozniak. 2014. Identification of candidate genes, regions and markers for pre-harvest sprouting resistance in wheat (*Triticum aestivum* L.). BMC Plant Biol. 14: 340. doi: 10.1186/s12870-014-0340-1.

Chen, M.-S., E. Echegaray, R.J. Whitworth, H. Wang, P.E. Sloderbeck, A. Knutson, K.L. Giles, and T.A. Royer. 2009. Virulence analysis of Hessian fly populations from Texas, Oklahoma, and Kansas. J. Econ. Entomol. 102(2): 774–780. doi: 10.1603/029.102.0239.

Chen, K., Y. Wang, R. Zhang, H. Zhang, and C. Gao. 2019. CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. Annu. Rev. Plant Biol. 70(1): 667–697. doi:

10.1146/annurev-arplant-050718-100049.

Crossa, J., P. Pérez-Rodríguez, J. Cuevas, O. Montesinos-López, D. Jarquín, G. de los Campos, J. Burgueño, J.M. González-Camacho, S. Pérez-Elizalde, Y. Beyene, S. Dreisigacker, R. Singh, X. Zhang, M. Gowda, M. Roorkiwal, J. Rutkoski, and R.K. Varshney. 2017. Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. Trends Plant Sci. 22(11): 961–975. doi: 10.1016/j.tplants.2017.08.011.

Darrier, B., J. Russell, S.G. Milner, P.E. Hedley, P.D. Shaw, M. Macaulay, L.D. Ramsay, C. Halpin, M. Mascher, D.L. Fleury, P. Langridge, N. Stein, and R. Waugh. 2019. A comparison of mainstream genotyping platforms for the evaluation and use of barley genetic resources. Front. Plant Sci. 10(April): 1–14. doi: 10.3389/fpls.2019.00544.

Demirer, G.S., H. Zhang, N.S. Goh, E. González-Grandío, and M.P. Landry. 2019a. Carbon nanotube–mediated DNA delivery without transgene integration in intact plants. Nat. Protoc. 14(10): 2954–2971. doi: 10.1038/s41596-019-0208-9.

Demirer, G.S., H. Zhang, J.L. Matos, N.S. Goh, F.J. Cunningham, Y. Sung, R. Chang, A.J. Aditham, L. Chio, M.J. Cho, B. Staskawicz, and M.P. Landry. 2019b. High aspect ratio nanomaterials enable delivery of functional genetic material without DNA integration in mature plants. Nat. Nanotechnol. 14(5): 456–464. doi: 10.1038/s41565-019-0382-5.

DeWitt, N., M. Guedira, E. Lauer, M. Sarinelli, P. Tyagi, D. Fu, Q. Hao, J.P. Murphy, D. Marshall, A. Akhunova, K. Jordan, E. Akhunov, and G. Brown-Guedira. 2019. Sequence-based mapping identifies a candidate transcription repressor underlying awn suppression at the B1 locus in wheat . New Phytol. doi: 10.1111/nph.16152.

Dhakal, S., C.T. Tan, L. Paezold, M.P. Fuentealba, J.C. Rudd, B.C. Blaser, Q. Xue, C.M. Rush, R.N. Devkota, and S. Liu. 2017. Wheat curl mite resistance in hard winter wheat in the US

great plains. Crop Sci. 57(1): 53–61. doi: 10.2135/cropsci2016.02.0121.

Doudna, J.A., and E. Charpentier. 2014. The new frontier of genome engineering with CRISPR-Cas9. Science (80-.). 346. doi: 10.1126/science.1258096.

Doyle, JJ; Doyle, J. 1990. Doyle&Doyle_Focus_1990_CTAB.pdf. Focus (Madison). 12(1): 13–15.

Edwards, J.T. 2013. Iba. : 2013.

Fagerlund, R.D., R.H.J. Staals, and P.C. Fineran. 2015. The Cpf1 CRISPR-Cas protein expands genome-editing tools. Genome Biol. 16(1): 15–17. doi: 10.1186/s13059-015-0824-9.

FAO. 2020. FAOSTAT wheat production. http://www.fao.org/faostat/en/#data/QC.

Ferrante, A., J. Cartelle, R. Savin, and G.A. Slafer. 2017. Yield determination, interplay between major components and yield stability in a traditional and a contemporary wheat across a wide range of environments. F. Crop. Res. 203: 114–127. doi: 10.1016/j.fcr.2016.12.028.

Feuillet, C., S. Travella, N. Stein, L. Albar, A. Nublat, and B. Keller. 2003. Map-based isolation of the leaf rust disease resistance gene Lr10 from the hexaploid wheat (Triticum aestivum L.) genome. Proc. Natl. Acad. Sci. U. S. A. 100(25): 15253–15258. doi: 10.1073/pnas.2435133100.

Gao, F., W. Wen, J. Liu, A. Rasheed, G. Yin, X. Xia, X. Wu, and Z. He. 2015. Genome-wide linkage mapping of QTL for yield components, plant height and yield-related physiological traits in the Chinese wheat cross Zhou 8425B/Chinese spring. Front. Plant Sci. 6(DEC): 1–17. doi: 10.3389/fpls.2015.01099.

Garcés-Carrera, S., A. Knutson, H. Wang, K.L. Giles, F. Huang, R.J. Whitworth, C.M. Smith, and M.-S. Chen. 2014. Virulence and biotype analyses of Hessian fly (Diptera: Cecidomyiidae) populations from Texas, Louisiana, and Oklahoma. J. Econ. Entomol.

107(1): 417–23. doi: 10.1603/EC13372.

Gu, X.Y., M.E. Foley, D.P. Horvath, J. V. Anderson, J. Feng, L. Zhang, C.R. Mowry, H. Ye, J.C. Suttle, K.I. Kadowaki, and Z. Chen. 2011. Association between seed dormancy and pericarp color is controlled by a pleiotropic gene that regulates abscisic acid and flavonoid synthesis in weedy red rice. Genetics 189(4): 1515–1524. doi: 10.1534/genetics.111.131169.

Hamada, H., Q. Linghu, Y. Nagira, R. Miki, N. Taoka, and R. Imai. 2017. An in planta biolistic method for stable wheat transformation. Sci. Rep. 7(1): 2–9. doi: 10.1038/s41598-017-11936-0.

Hamada, H., Y. Liu, Y. Nagira, R. Miki, N. Taoka, and R. Imai. 2018. Biolistic-delivery-based transient CRISPR/Cas9 expression enables in planta genome editing in wheat. Sci. Rep. 8(1): 6–12. doi: 10.1038/s41598-018-32714-6.

Hao, Y., S.E. Cambron, C. Zhenbang, Y. Wang, D.E. Bland, G.D. Butin, and J.W. Johnson. 2013. Characterization of new loci for Hessian fly resistance in common wheat. Theor Appl Genet 126: 1067–1076. doi: 10.1007/s00122-012-2037-x.

Hayta, S., M.A. Smedley, S.U. Demir, R. Blundell, A. Hinchliffe, N. Atkinson, and W.A. Harwood. 2019. An efficient and reproducible Agrobacterium-mediated transformation method for hexaploid wheat (Triticum aestivum L.). Plant Methods 15(1): 1–15. doi: 10.1186/s13007-019-0503-z.

Himi, E., M. Maekawa, H. Miura, and K. Noda. 2011. Development of PCR markers for Tamyb10 related to R-1, red grain color gene in wheat. Theor. Appl. Genet. 122(8): 1561–1576. doi: 10.1007/s00122-011-1555-2.

Himi, E., and K. Noda. 2004. Isolation and location of three homoeologous dihydroflavonol-4-

reductase (DFR) genes of wheat and their tissue-dependent expression. J. Exp. Bot. 55(396): 365–375. doi: 10.1093/jxb/erh046.

Himi, E., and K. Noda. 2005. Red grain colour gene (R) of wheat is a Myb-type transcription factor. Euphytica 143(3): 239–242. doi: 10.1007/s10681-005-7854-4.

Jamil, M., A. Ali, A. Gul, A. Ghafoor, A.A. Napar, A.M.H. Ibrahim, N.H. Naveed, N.A. Yasin, and A. Mujeeb-Kazi. 2019. Genome-wide association studies of seven agronomic traits under two sowing conditions in bread wheat. BMC Plant Biol. 19(1): 1–18. doi: 10.1186/s12870-019-1754-6.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, and E. Charpentier. 2012. A Programmable Dual-RNA – Guided. 337(August): 816–822. doi: 10.1126/science.1225829.

Juliana, P., R.P. Singh, P.K. Singh, J.A. Poland, G.C. Bergstrom, J. Huerta-Espino, S. Bhavani, J. Crossa, and M.E. Sorrells. 2018. Genome-wide association mapping for resistance to leaf rust, stripe rust and tan spot in wheat reveals potential candidate genes. Theor. Appl. Genet. (0123456789): 1–18. doi: 10.1007/s00122-018-3086-6.

Kato, K., H. Miura, M. Akiyama, M. Kuroshima, and S. Sawada. 1998. RFLP mapping of the three major genes, Vrn1, Q and B1, on the long arm of chromosome 5A of wheat. Euphytica 101(1): 91–95. doi: 10.1023/A:1018372231063.

Kosambi, D.D. 2016. The Estimation of Map Distances from Recombination Values. p. 125–130. *In* D.D. Kosambi. Springer India, New Delhi.

Kottearachchi, N.S., N. Uchino, K. Kato, and H. Miura. 2006. Increased grain dormancy in white-grained wheat by introgression of preharvest sprouting tolerance QTLs. Euphytica 152(3): 421–428. doi: 10.1007/s10681-006-9231-3.

Kuang, C.H., X.F. Zhao, K. Yang, Z.P. Zhang, L. Ding, Z.E. Pu, J. Ma, Q.T. Jiang, G.Y. Chen,

J.R. Wang, Y.M. Wei, Y.L. Zheng, and W. Li. 2020. Mapping and characterization of major QTL for spike traits in common wheat. Physiol. Mol. Biol. Plants 26(6): 1295–1307. doi: 10.1007/s12298-020-00823-0.

Kuzay, S., Y. Xu, J. Zhang, A. Katz, S. Pearce, Z. Su, M. Fraser, J.A. Anderson, G. Brown-Guedira, N. DeWitt, A. Peters Haugrud, J.D. Faris, E. Akhunov, G. Bai, and J. Dubcovsky. 2019. Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. Theor. Appl. Genet. 132(9): 2689–2705. doi: 10.1007/s00122-019-03382-5.

Langmead, B., and S.L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9(4): 357–359. doi: 10.1038/nmeth.1923.

Larson, M.H., L.A. Gilbert, X. Wang, W.A. Lim, J.S. Weissman, and L.S. Qi. 2013. CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nat. Protoc. 8(11): 2180–2196. doi: 10.1038/nprot.2013.132.

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. doi: 10.1093/bioinformatics/btp352.

Li, G., Y. Wang, M.-S. Chen, E. Edae, J. Poland, E. Akhunov, S. Chao, G. Bai, B.F. Carver, and L. Yan. 2015. Precisely mapping a major gene conferring resistance to Hessian fly in bread wheat using genotyping-by-sequencing. BMC Genomics 16(1): 108. doi: 10.1186/s12864-015-1297-7.

Li, J., Z. Wang, G. He, L. Ma, and X.W. Deng. 2020. CRISPR/Cas9 mediated disruption of TaNP1 genes results in complete male sterility in bread wheat. J. Genet. Genomics. doi: 10.1016/j.jgg.2020.05.004.

Li, F., W. Wen, J. Liu, Y. Zhang, S. Cao, Z. He, A. Rasheed, H. Jin, C. Zhang, J. Yan, P. Zhang, Y. Wan, and X. Xia. 2019. Genetic architecture of grain yield in bread wheat based on genome-wide association studies. BMC Plant Biol. 19(1): 1–19. doi: 10.1186/s12870-019-1781-3.

Liang, Z., K. Chen, T. Li, Y. Zhang, Y. Wang, Q. Zhao, J. Liu, H. Zhang, C. Liu, Y. Ran, and C. Gao. 2017. Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. Nat. Commun. 8: 1–5. doi: 10.1038/ncomms14261.

Liang, Z., K. Chen, Y. Zhang, J. Liu, K. Yin, J. Qiu, and C. Gao. 2018. Genome editing of bread wheat using biolistic delivery of CRISPR/Cas9 in vitro transcripts or ribonucleoproteins. Nat. Protoc. 13(3): 413–430. doi: 10.1038/nprot.2017.145.

Lin, M., D. Zhang, S. Liu, G. Zhang, J. Yu, A.K. Fritz, and G. Bai. 2016. Genome-wide association analysis on pre-harvest sprouting resistance and grain color in U.S. winter wheat. BMC Genomics 17(1). doi: 10.1186/s12864-016-3148-6.

Liu, S., C.A. Griffey, M.D. Hall, A.L. McKendry, J. Chen, W.S. Brooks, G. Brown-Guedira, D. Van Sanford, and D.G. Schmale. 2013a. Molecular characterization of field resistance to Fusarium head blight in two US soft red winter wheat cultivars. Theor. Appl. Genet. 126(10): 2485–2498. doi: 10.1007/s00122-013-2149-y.

Liu, S., J.C. Rudd, G. Bai, S.D. Haley, A.M.H. Ibrahim, Q. Xue, D.B. Hays, R.A. Graybosch, R.N. Devkota, and P. St. Amand. 2014. Molecular markers linked to important genes in hard winter wheat. Crop Sci. 54(4): 1304–1321. doi: 10.2135/cropsci2013.08.0564.

Liu, S., S.K. Sehgal, J. Li, M. Lin, H.N. Trick, J. Yu, B.S. Gill, and G. Bai. 2013b. Cloning and characterization of a critical regulator for preharvest sprouting in wheat. Genetics 195(1): 263–273. doi: 10.1534/genetics.113.152330.

Lopes, M.S., M.P. Reynolds, Y. Manes, R.P. Singh, J. Crossa, and H.J. Braun. 2012. Genetic yield gains and changes in associated traits of CIMMYT spring bread wheat in a “Historic” set representing 30 years of breeding. Crop Sci. 52(3): 1123–1131. doi: 10.2135/cropsci2011.09.0467.

Lowder, L.G., A. Malzahn, and Y. Qi. 2018. Plant Gene Regulation Using Multiplex CRISPR-dCas9 Artificial Transcription Factors. p. 197–214. *In* Lagrimini, L.M. (ed.), Methods in Molecular Biology. Springer New York, New York, NY.

Lozada, D.N., R.E. Mason, S. Sukumaran, and S. Dreisigacker. 2018. Validation of grain yield QTLs from soft winter wheat using a CIMMYT spring wheat panel. Crop Sci. 58(5): 1964–1971. doi: 10.2135/cropsci2018.04.0232.

Lu, H., J.C. Rudd, J.D. Burd, and Y. Weng. 2010. Molecular mapping of greenbug resistance genes Gb2 and Gb6 in T1AL . 1RS wheat-rye translocations. Plant Soil 129: 472–476. doi: doi:10.1111/j.1439-0523.2009.01722.x.

Mali, P., L. Yang, K.M. Esvelt, J. Aach, M. Guell, J.E. DiCarlo, J.E. Norville, and G.M. Church. 2013. RNA-guided human genome engineering via Cas9. Science (80-.). 339(6121): 823–826. doi: 10.1126/science.1232033.

Mayer, K.F.X., T. Marcussen, S.R. Sandve, L. Heier, M. Pfeifer, K.G. Kugler, B. Zhan, M. Spannagl, M. Pfeifer, K.S. Jakobsen, B.B.H. Wulff, B. Steuernagel, and O.-A. Olsen. 2014. A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome Ancient hybridizations among the ancestral genomes of bread wheat Genome interplay in the grain transcriptome of hexaploid bread wheat Structural and functional pa. Science 345(6194): 1250092. doi: 10.1126/science.1251788.

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(9): 1297–1303. doi: 10.1101/gr.107524.110.

Meng, L., H. Li, L. Zhang, and J. Wang. 2015. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. Crop J. 3(3): 269–283. doi: 10.1016/j.cj.2015.01.001.

Minkenberg, B., K. Xie, and Y. Yang. 2017. Discovery of rice essential genes by characterizing a CRISPR-edited mutation of closely related rice MAP kinase genes. Plant J. 89(3): 636–648. doi: 10.1111/tpj.13399.

Mishra, R., R.K. Joshi, and K. Zhao. 2018. Genome editing in rice: recent advances, challenges, and future implications. Front. Plant Sci. 9(September). doi: 10.3389/fpls.2018.01361.

Mishra, R., and K. Zhao. 2018. Genome editing technologies and their applications in crop improvement. Plant Biotechnol. Rep. 12(2): 57–68. doi: 10.1007/s11816-018-0472-0.

Mo, Y., L.S. Vanzetti, I. Hale, E.J. Spagnolo, F. Guidobaldi, J. Al-Oboudi, N. Odle, S. Pearce, M. Helguera, and J. Dubcovsky. 2018. Identification and characterization of Rht25, a locus on chromosome arm 6AS affecting wheat plant height, heading time, and spike development. Theor. Appl. Genet. 131(10): 2021–2035. doi: 10.1007/s00122-018-3130-6.

Ogbonnaya, F.C., A. Rasheed, E.C. Okechukwu, A. Jighly, F. Makdis, T. Wuletaw, A. Hagras, M.I. Uguru, and C.U. Agbo. 2017. Genome-wide association study for agronomic and physiological traits in spring wheat evaluated in a range of heat prone environments. Theor. Appl. Genet. 130(9): 1819–1835. doi: 10.1007/s00122-017-2927-z.

Van Ooijen, J.W. 2006. JoinMap 4 , Software for the calculation of genetic linkage maps in experimental populations.

Osa, M., K. Kato, M. Mori, C. Shindo, A. Torada, and H. Miura. 2003. Mapping QTLs for seed dormancy and the Vp1 homologue on chromosome 3A in wheat. Theor. Appl. Genet. 106(8): 1491–1496. doi: 10.1007/s00122-003-1208-1.

Pascual, L., E. Albert, C. Sauvage, J. Duangjit, J.P. Bouchet, F. Bitton, N. Desplat, D. Brunel, M.C. Le Paslier, N. Ranc, L. Bruguier, B. Chauchard, P. Verschave, and M. Causse. 2016. Dissecting quantitative trait variation in the resequencing era: Complementarity of bi-parental, multi-parental and association panels. Plant Sci. 242: 120–130. doi: 10.1016/j.plantsci.2015.06.017.

Poland, J., and J. Rutkoski. 2016. Advances and Challenges in Genomic Selection for Disease Resistance. Annu. Rev. Phytopathol. 54(1): 79–98. doi: 10.1146/annurev-phyto-080615-100056.

Rasheed, A., W. Wen, F. Gao, S. Zhai, H. Jin, J. Liu, Q. Guo, Y. Zhang, S. Dreisigacker, X. Xia, and Z. He. 2016. Development and validation of KASP assays for genes underpinning key economic traits in bread wheat. Theor. Appl. Genet. 129(10): 1843–1860. doi: 10.1007/s00122-016-2743-x.

Ray, D.K., N.D. Mueller, P.C. West, and J.A. Foley. 2013. Yield Trends Are Insufficient to Double Global Crop Production by 2050. PLoS One 8(6). doi: 10.1371/journal.pone.0066428.

Rebetzke, G.J., R.A. Richards, N.A. Fettell, M. Long, A.G. Condon, R.I. Forrester, and T.L. Botwright. 2007. Genotypic increases in coleoptile length improves stand establishment, vigour and grain yield of deep-sown wheat. F. Crop. Res. 100(1): 10–23. doi: 10.1016/j.fcr.2006.05.001.

RStudio Team. 2015. Integrated development for R. Inc. : 2015. http://www.rstudio.com/.

Rudd, J.C., R.N. Devkota, A.M. Ibrahim, J.A. Baker, S. Baker, R. Sutton, B. Simoneaux, G. Opena, D. Hathcoat, J.M. Awika, L.R. Nelson, S. Liu, Q. Xue, B. Bean, C.B. Neely, R.W. Duncan, B.W. Seabourn, R.L. Bowden, Y. Jin, M. Chen, and R.A. Graybosch. 2019. 'TAM 204' Wheat, Adapted to Grazing, Grain, and Graze-out Production Systems in the Southern High Plains. 382: 377–382. doi: 10.3198/jpr2018.12.0080crc.

Sadras, V.O. 2007. Evolutionary aspects of the trade-off between seed size and number in crops. F. Crop. Res. 100(2–3): 125–138. doi: 10.1016/j.fcr.2006.07.004.

Sadras, V.O., and G.A. Slafer. 2011. Environmental modulation of yield components in cereals: Heritabilities reveal a hierarchy of phenotypic plasticities. F. Crop. Res. 127: 215–224. doi: 10.1016/j.fcr.2011.11.014.

Saltz, J.B., F.C. Hessel, and M.W. Kelly. 2017. Trait Correlations in the Genomics Era. Trends Ecol. Evol. 32(4): 279–290. doi: 10.1016/j.tree.2016.12.008.

Sánchez-León, S., J. Gil-Humanes, C. V. Ozuna, M.J. Giménez, C. Sousa, D.F. Voytas, and F. Barro. 2017. Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9. Plant Biotechnol. J.: 902–910. doi: 10.1111/pbi.12837.

Shokat, S., D. Sehgal, P. Vikram, F. Liu, and S. Singh. 2020. Molecular markers associated with agro-physiological traits under terminal drought conditions in bread wheat. Int. J. Mol. Sci. 21(9). doi: 10.3390/ijms21093156.

Shorinola, O., N. Bird, J. Simmonds, S. Berry, T. Henriksson, P. Jack, P. Werner, T. Gerjets, D. Scholefield, B. Balcárková, M. Valárik, M.J. Holdsworth, J. Flintham, and C. Uauy. 2016. The wheat Phs-A1 pre-harvest sprouting resistance locus delays the rate of seed dormancy loss and maps 0.3 cM distal to the PM19 genes in UK germplasm. J. Exp. Bot. 67(14): 4169–4178. doi: 10.1093/jxb/erw194.

Singh, M., M. Kumar, M.C. Albertsen, J.K. Young, and A.M. Cigan. 2018. Concurrent modifications in the three homeologs of Ms45 gene with CRISPR-Cas9 lead to rapid generation of male sterile bread wheat (Triticum aestivum L.). Plant Mol. Biol. 97(4–5): 371–383. doi: 10.1007/s11103-018-0749-2.

Slafer, G.A., R. Savin, and V.O. Sadras. 2014. Coarse and fine regulation of wheat yield components in response to genotype and environment. F. Crop. Res. 157: 71–83. doi: 10.1016/j.fcr.2013.12.004.

Sukumaran, S., S. Dreisigacker, M. Lopes, P. Chavez, and M.P. Reynolds. 2014. Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. Theor. Appl. Genet. 128(2): 353–363. doi: 10.1007/s00122-014-2435-3.

Tan, C.T., H. Yu, Y. Yang, X. Xu, M. Chen, J.C. Rudd, Q. Xue, A.M.H. Ibrahim, L. Garza, S. Wang, M.E. Sorrells, and S. Liu. 2017. Development and validation of KASP markers for the greenbug resistance gene Gb7 and the Hessian fly resistance gene H32 in wheat. Theor. Appl. Genet.: 1–18. doi: 10.1007/s00122-017-2930-4.

Velu, G., R.P. Singh, M.E. Cardenas, B. Wu, C. Guzman, and I. Ortiz-Monasterio. 2017. Characterization of grain protein content gene (GPC-B1) introgression lines and its potential use in breeding for enhanced grain zinc and iron concentration in spring wheat. Acta Physiol. Plant. 39(9): 1–9. doi: 10.1007/s11738-017-2509-3.

Vetch, J.M., R.N. Stougaard, J.M. Martin, and M.J. Giroux. 2019. Review: Revealing the genetic mechanisms of pre-harvest sprouting in hexaploid wheat (Triticum aestivum L.). Plant Sci. 281(January): 180–185. doi: 10.1016/j.plantsci.2019.01.004.

Voss-Fels, K.P., G. Keeble-Gagnère, L.T. Hickey, J. Tibbits, S. Nagornyy, M.J. Hayden, R.K.

Pasam, S. Kant, W. Friedt, R.J. Snowdon, R. Appels, and B. Wittkop. 2019. High-resolution mapping of rachis nodes per rachis, a critical determinant of grain yield components in wheat. Theor. Appl. Genet. 132(9): 2707–2719. doi: 10.1007/s00122-019-03383-4.

Wang, Y., X. Cheng, Q. Shan, Y. Zhang, J. Liu, C. Gao, and J.L. Qiu. 2014. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nat. Biotechnol. 32(9): 947–951. doi: 10.1038/nbt.2969.

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

Wang, Y., X.L. Wang, J.Y. Meng, Y.J. Zhang, Z.H. He, and Y. Yang. 2016. Characterization of Tamyb10 allelic variants and development of STS marker for pre-harvest sprouting resistance in Chinese bread wheat. Mol. Breed. 36(11). doi: 10.1007/s11032-016-0573-9.

Weng, Y., W. Li, R.N. Devkota, and J.C. Rudd. 2005. Microsatellite markers associated with two Aegilops tauschii-derived greenbug resistance loci in wheat. Theor. Appl. Genet. 110(3): 462–469. doi: 10.1007/s00122-004-1853-z.

Williams, C.E., C.C. Collier, N. Sardesai, H.W. Ohm, and S.E. Cambron. 2003. Phenotypic assessment and mapped markers for H31 , a new wheat gene conferring resistance to Hessian fly (Diptera : Cecidomyiidae). TAG. Theor. Appl. Genet. 107: 1516–1523. doi: 10.1007/s00122-003-1393-y.

Woo, J.W., J. Kim, S. Il Kwon, C. Corvalán, S.W. Cho, H. Kim, S.G. Kim, S.T. Kim, S. Choe, and J.S. Kim. 2015. DNA-free genome editing in plants with preassembled CRISPR-Cas9 ribonucleoproteins. Nat. Biotechnol. 33(11): 1162–1164. doi: 10.1038/nbt.3389.

Xu, Y. 2016. Envirotyping for deciphering environmental impacts on crop plants. Theor. Appl.

Genet. 129(4): 653–673. doi: 10.1007/s00122-016-2691-5.

Yan, W., and N.A. Tinker. 2006. Biplot analysis of multi-environment trial data: Principles and applications. Can. J. Plant Sci. 86(3): 623–645. doi: 10.4141/P05-169.

Yang, Y., S. Dhakal, C. Chenggen, S. Wang, Q. Xue, J.C. Rudd, A.M.H. Ibrahim, K. Jessup, J. Baker, M.P. Fuentealba, R. Devkota, S. Baker, C.D. Johnson, R. Metz, and S. Liu. 2020. Genome wide identification of QTL associated with yield and yield components in two popular wheat cultivars TAM 111 and TAM 112. biorxiv: 1–32. doi: 10.1101/2020.07.27.222703.

Yang, Y., Y.Z. Ma, Z.S. Xu, X.M. Chen, Z.H. He, Z. Yu, M. Wilkinson, H.D. Jones, P.R. Shewry, and L.Q. Xia. 2007. Isolation and characterization of Viviparous-1 genes in wheat cultivars with distinct ABA sensitivity and pre-harvest sprouting tolerance. J. Exp. Bot. 58(11): 2863–2871. doi: 10.1093/jxb/erm073.

Yu, J., J.B. Holland, M.D. McMullen, and E.S. Buckler. 2008. Genetic design and statistical power of nested association mapping in maize. Genetics 178(1): 539–551. doi: 10.1534/genetics.107.074245.

Zale, J.M., H. Borchardt-Wier, K.K. Kidwell, and C.M. Steber. 2004. Callus Induction and Plant Regeneration from Mature Embryos of a Diverse Set of Wheat Genotypes. Plant Cell. Tissue Organ Cult. 76(3): 277–281. doi: 10.1023/B:TICU.0000009248.32457.4c.

Zhang, Y., Z. Liang, Y. Zong, Y. Wang, J. Liu, K. Chen, J.L. Qiu, and C. Gao. 2016a. Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. Nat. Commun. 7: 1–8. doi: 10.1038/ncomms12617.

Zhang, Y., Z. Liang, Y. Zong, Y. Wang, J. Liu, K. Chen, J.-L. Qiu, and C. Gao. 2016b. Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9

DNA or RNA. Nat. Commun. 7: 12617. doi: 10.1038/ncomms12617.

Zhang, Y., X. Miao, X. Xia, and Z. He. 2014. Cloning of seed dormancy genes (TaSdr) associated with tolerance to pre-harvest sprouting in common wheat and development of a functional marker. Theor. Appl. Genet. 127(4): 855–866. doi: 10.1007/s00122-014-2262-6.

Zhao, L., N.R. Abdelsalam, Y. Xu, M.S. Chen, Y. Feng, L. Kong, and G. Bai. 2020. Identification of two novel Hessian fly resistance genes H35 and H36 in a hard winter wheat line SD06165. Theor. Appl. Genet. 133(8): 2343–2353. doi: 10.1007/s00122-020-03602-3.

Zhu, C., M. Gore, E.S. Buckler, and J. Yu. 2008. Status and Prospects of Association Mapping in Plants. Plant Genome J. 1(1): 5. doi: 10.3835/plantgenome2008.02.0089.

APPENDIX

Supplemental table 1. Single Nucleotide Polymorphisms markers distribution, genetic distance, physical position and ratio between them in the linkage groups used QTL mapping in the TAM 204/Iba population.

Chromosome	LG number	No. markers	Genetic distance (cM)	Physical distance (Mb)	No. of SNPs/cM	No. of SNPs/Mb
1A	1	431	145.00	533.74	0.34	0.81
1A2	2	130	28.87	41.98	0.22	3.10
1B	3	57	25.53	32.76	0.45	1.74
1B2	4	1033	341.11	528.82	0.33	1.95
1B3	5	88	24.91	27.28	0.28	3.23
1B4	6	12	1.68	2.24	0.14	5.35
1B5	7	36	5.02	13.09	0.14	2.75
1D	8	35	24.29	17.91	0.69	1.95
1D2	9	226	76.45	434.21	0.34	0.52
2A	10	371	121.03	687.99	0.33	0.54
2A2	11	84	42.72	39.58	0.51	2.12
2B	12	66	29.68	34.60	0.45	1.91
2B2	13	488	151.17	704.93	0.31	0.69
2B3	14	74	18.83	19.89	0.25	3.72
2D	15	55	31.67	43.08	0.58	1.28
2D2	16	231	121.93	518.83	0.53	0.45
2D3	17	35	36.53	62.22	1.04	0.56
3A1	18	40	11.19	20.55	0.28	1.95
3A2	19	68	51.81	222.28	0.76	0.31
3A3	20	70	29.16	30.47	0.42	2.30
3B1	21	46	18.00	22.93	0.39	2.01
3B2	22	103	44.82	178.97	0.44	0.58

LG= Linkage groups; cM= centiMorgan; Mb= Mega bases; SNPs= Single Nucleotide Polymorphisms.

Supplemental table 2. Single Nucleotide Polymorphisms markers distribution, genetic distance, physical position and ratio between them in the linkage groups used QTL mapping in the TAM 204/Iba population.

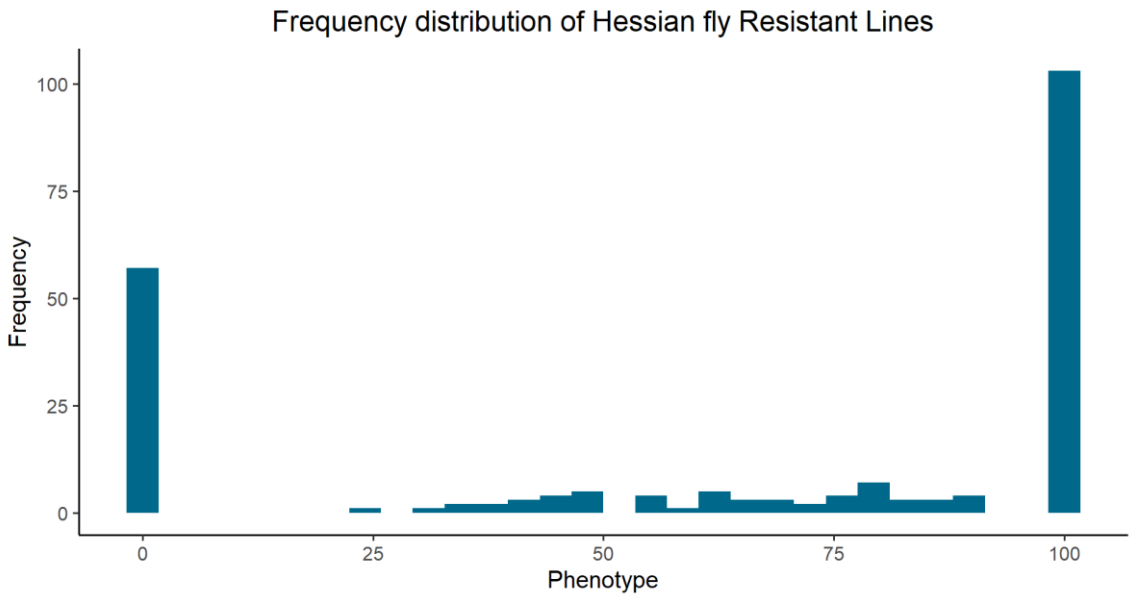
Chromosome	LG number	No. markers	Genetic distance (cM)	Physical distance (Mb)	No. of SNPs/cM	No. of SNPs/Mb
3B3	23	20	17.68	29.40	0.88	0.68
3D	24	131	33.73	348.80	0.26	0.38
4A1	25	222	108.91	534.04	0.49	0.42
4A2	26	22	17.47	13.90	0.79	1.58
4A3	27	219	104.30	116.10	0.48	1.89
4B1	28	21	6.12	5.16	0.29	4.07
4B2	29	220	99.62	471.49	0.45	0.47
4B3	30	73	31.10	24.57	0.43	2.97
4D	31	302	98.41	452.67	0.33	0.67
5A1	32	304	104.46	417.45	0.34	0.73
5A2	33	180	67.63	113.60	0.38	1.58
5A3	34	118	57.06	78.55	0.48	1.50
5A4	35	94	36.22	40.29	0.39	2.33
5B1	36	51	30.61	15.42	0.60	3.31
5B2	37	699	214.94	483.53	0.31	1.45
5B3	38	70	24.57	34.35	0.35	2.04
5B4	39	234	52.53	76.91	0.22	3.04
5B5	40	77	34.38	38.47	0.45	2.00
5D1	41	224	58.84	335.54	0.26	0.67
5D2	42	15	32.64	50.01	2.18	0.30
5D4	43	20	22.02	29.57	1.10	0.68
6A1	44	25	9.11	11.07	0.36	2.26
6A2	45	286	144.08	545.88	0.50	0.52

LG= Linkage groups; cM= centiMorgan; Mb= Mega bases; SNPs= Single Nucleotide Polymorphisms.

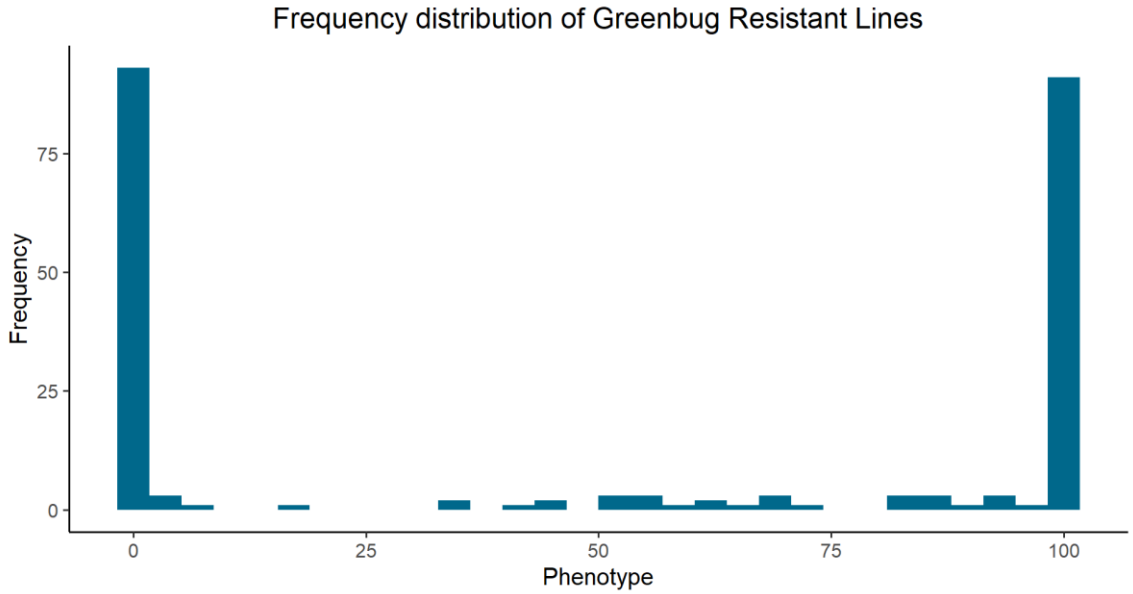
Supplemental table 3. Single Nucleotide Polymorphisms markers distribution, genetic distance, physical position and ratio between them in the linkage groups used QTL mapping in the TAM 204/Iba population.

Chromosome	LG number	No. markers	Genetic distance (cM)	Physical distance (Mb)	No. of SNPs/cM	No. of SNPs/Mb
6A3	46	15	8.04	6.24	0.54	2.40
6A4	47	16	19.23	24.31	0.57	0.99
6B1	48	24	406.93	621.14	0.80	1.76
6B2	49	1096	28.74	21.56	0.37	1.39
6D1	50	30	82.79	364.07	0.96	0.48
6D2	51	175	6.00	15.31	0.47	1.44
6D3	52	22	174.24	593.41	0.27	0.66
7A1	53	393	37.02	47.89	0.44	0.90
7A2	54	43	6.25	10.90	0.86	2.66
7B1	55	29	91.02	610.36	0.22	0.38
7B2	56	229	33.84	40.73	0.40	2.18
7B3	57	89	30.10	40.81	0.38	1.79
7D1	58	73	114.94	465.15	0.41	0.59
7D2	59	273	152.02	54.75	0.42	4.99
7D3	60	273	19.23	24.31	0.56	0.99

LG= Linkage groups; cM= centiMorgan; Mb= Mega bases; SNPs= Single Nucleotide Polymorphisms.



Supplemental figure 1. Frequency distribution of the percentage of Hessian fly resistance lines in the TAM 204/Iba population.



Supplemental figure 2. Frequency distribution of the percentage of Greenbug resistance lines in the TAM 204/Iba population.

Supplemental table 4. Phenotypic values of Iba and TAM 204 for grain yield (GYLD), days to heading (DTH), plant height (PH), thousand kernel weight (TKW) and test weight (TW) in all the tested environments.

Parental line	GYLD (g m^{-2})	DTH (days)	PH (cm)	TKW (g)	TW (Kg m^{-3})
		18CS			
Iba		84.39			
TAM 204		87.07			
		18DMS			
Iba	490.09	135	84.5		
TAM 204	542.19	135	84		
		18MCG			
Iba	419.77	88.10	86.21	36.45	761.89
TAM 204	353.49	89.99	89.95	28.64	699.80
		19BI			
Iba	643.98	130.85	93.29		
TAM 204	581.13	129.71	93.012		
		19CH			
Iba	314.60				
TAM 204	266.88				
		19CS			
Iba	262.24	87.74	87.25		767.30
TAM 204	157.96	87.98	79.5		
		19 EMN			
Iba	497.87		81.71		812.17
TAM 204	480.28		81.65		771.17
		19MCG			
Iba	226.21	96.98	89.35	27.57	662.41
TAM 204	138.17	98.96	87.63	20.83	628.12
		20CS			
Iba	100.55	80.48	83		
TAM 204	90.46	74.99	73.5		
		20EMN			
Iba	256.72				775.5141
TAM 204	157.76				727.3158
		20MCG			
Iba	391.86	98.03	80	36.7	767.06
TAM 204	290.23	105.45	80	29.1	706.99

18CS, College Station 2018; 18DMS, Dumas 2018; 18MCG, McGregor 2018; 19BI, Bushland Irrigated 2019; 19CH, Chillicothe 2019; 19CS, College Station 2019; 19EMN, Emeny Land 2019; 19MCG, McGregor 2019; 20CS, College Station 2020; 20EMN, Emeny Land 2020; 20.MCG, McGregor 2020.

Supplemental table 5. Components of the sgRNA *in-vitro* digestion described by Jinek et al., 2012, Larson et al., 2013 and Mali et al., 2013.

Component	Amount (µL)
Nuclease free water	20
NEBuffer 3.1	3
300 nM sgRNA	3
1 µM Cas9 nuclease	1
30 nM DNA	27

Supplemental table 6. Description and preparation steps of the mediums used for tissue culture described by Liang et al., 2018.

Osmotic medium (1L)	Recovery medium (1L)	Regeneration medium (1L)	Rooting medium (1L)
1)Added 4.4 gr of Murashige and Skoog (MS) salt including vitamins, 30 gr of sucrose, 5 ml of 2,4-D (1 mg/ml), 0and 72.8 gr of mannitol to 800 ml of ddH2O. 2) Adjusted the volume to 1 liter with RNase-free water and the pH to 5.8 with 1 M KOH, and added 3.2 g of phytagel. 3) Autoclaved at 121 °C for 20 min. Poured ~30 ml into sterile Petri dishes after cooling to 50 °C.	1)Added 4.4 gr of Murashige and Skoog (MS) salt including vitamins, 30 gr of sucrose, 2 ml of 2,4-D (1 mg/ml), 0.5 g of N-Z- Amine A (from bovine serum)and 600 µl of CuSO4 (1 mg/ml) to 800 ml of ddH2O. 2) Adjusted the volume to 1 liter with RNase-free water and the pH to 5.8 with 1 M KOH, and added 3.2 g of phytagel. 3) Autoclaved at 121 °C for 20 min. Poured ~30 ml into sterile Petri dishes after cooling to 50 °C.	1) Added 4.4 gr of Murashige and Skoog (MS) salt including vitamins, 30 gr of sucrose, and µl of kinetin (1 mg/ml) to 800 ml of ddH2O. 2) Adjusted the volume to 1 liter with RNase-free water and the pH to 5.8 with 1 M KOH, and added 3.2 g of phytagel. 3) Autoclaved at 121 °C for 20 min. Poured ~30 ml into sterile Petri dishes after cooling to 50 °C.	1) Added 4.4 gr of Murashige and Skoog (MS) salt including vitamins, 30 gr of sucrose, and µl of kinetin (1 mg/ml) to 800 ml of ddH2O. 2) Adjusted the volume to 1 liter with RNase-free water and the pH to 5.8 with 1 M KOH, and added 3.2 g of phytagel. 3) Autoclaved at 121 °C for 20 min. After cooling to ~50 °C, added 100 µl of NAA (0.5 mg/ml). Poured ~30 ml into sterile Petri dishes after cooling to 50 °C.

2,4-D =2,4-Dichlorophenoxyacetic acid.

Supplemental table 7. List of recovered plants and treatments and genotyping information.

Plant name	treatment	Pool for genotyping	Plant name	treatment	Pool for genotyping
Tamyb10-1	10 Days + sgRNA1 + GFP		Tamyb10-33	11 Days + GFP	
Tamyb10-2*	10 Days + sgRNA1 + GFP		Tamyb10-34	11 Days + GFP	
Tamyb10-3	10 Days + sgRNA1 + GFP	1	Tamyb10-35	11 Days + GFP	
Tamyb10-4	10 Days + sgRNA1 + GFP		Tamyb10-36	11 Days + GFP	
Tamyb10-5	10 Days + sgRNA1 + GFP		Tamyb10-37	11 Days + GFP	
Tamyb10-6*	10 Days + sgRNA1 + GFP		Tamyb10-38	11 Days + GFP	
Tamyb10-7	10Days + sgRNA1	2	Tamyb10-39	11days + sgRNA1	4
Tamyb10-8	10Days + sgRNA1		Tamyb10-40	11days + sgRNA1	
Tamyb10-9	10Days + sgRNA1		Tamyb10-41*	11days + sgRNA1	
Tamyb10-10*	10Days + sgRNA1		Tamyb10-42	11days + sgRNA1	5
Tamyb10-11	10Days + sgRNA1	3	Tamyb10-43	11days + sgRNA1	
Tamyb10-12	10Days + sgRNA1		Tamyb10-44	11days + sgRNA1	
Tamyb10-13	10Days + sgRNA1		Tamyb10-45*	11days + sgRNA1	
Tamyb10-14*	10Days + sgRNA1	4	Tamyb10-46	11days + sgRNA1	6
Tamyb10-15	10Days + sgRNA1		Tamyb10-47	11days + sgRNA1	
Tamyb10-16	11 Days + GFP		Tamyb10-48	11days + sgRNA1	
Tamyb10-17	11 Days + GFP		Tamyb10-49*	11days + sgRNA1	
Tamyb10-18	11 Days + GFP		Tamyb10-50	11days + sgRNA1	7
Tamyb10-19*	11 Days + GFP		Tamyb10-51	11days + sgRNA1	
Tamyb10-20	11 Days + GFP	11	Tamyb10-52	11days + sgRNA1	
Tamyb10-21	11 Days + GFP		Tamyb10-53*	11days + sgRNA1	
Tamyb10-22	11 Days + GFP		Tamyb10-54	11days + sgRNA1	8
Tamyb10-23	11 Days + GFP		Tamyb10-55	11days + sgRNA1	
Tamyb10-24	11 Days + GFP		Tamyb10-56	11days + sgRNA1	
Tamyb10-25	11 Days + GFP		Tamyb10-57*	10 Days + sgRNA2	
Tamyb10-26	11 Days + GFP		Tamyb10-58	10 Days + sgRNA2	9
Tamyb10-27	11 Days + GFP		Tamyb10-59	10 Days + sgRNA2	
Tamyb10-28	11 Days + GFP		Tamyb10-60*	10 Days + sgRNA2	
Tamyb10-29	11 Days + GFP		Tamyb10-61	10 Days + sgRNA2	
Tamyb10-30	11 Days + GFP		Tamyb10-62	10 Days + sgRNA2	10
Tamyb10-31	11 Days + GFP		Tamyb10-63	10 Days + sgRNA2	
Tamyb10-32	11 Days + GFP				

*Plants selected for individual genotyping if the second exon. 10 and 11 days = refers to the number of days after flowering when the immature embryos were excised. GFP plasmid = plasmid DNA containing the with Green Fluorescence Protein transgene. Each pool consist of four plants that were selected in sequential order.