DISSERTATION

BI-PARENTAL MAPPING AND GENOME-WIDE ASSOCIATION STUDIES FOR GRAIN QUALITY TRAITS IN WINTER WHEAT UNDER CONTRASTING SOIL MOISTURE CONDITIONS

Submitted by

Hung Quoc Dao

Department of Soil and Crop Sciences

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2015

Doctoral Committee:

Advisor: Patrick F. Byrne

Mark A. Brick
Scott D. Haley
Courtney E. Jahn

ABSTRACT

BI-PARENTAL MAPPING AND GENOME-WIDE ASSOCIATION STUDIES FOR GRAIN QUALITY TRAITS IN WINTER WHEAT UNDER CONTRASTING SOIL MOISTURE CONDITIONS

Wheat grain quality is characterized by parameters such as grain protein concentration (Gpc), grain ash concentration (Gac), kernel weight (Kw), kernel diameter (Kd), and kernel hardness (Kh). Grain protein determines dough strength and loaf volume, while kernel hardness and size impact milling efficiency. Drought stress at flowering time can cause floral organ necrosis, thus, decreasing the number of grains per spike and filled grain percentage, while drought stress during grain filling reduces kernel weight and size, but increases grain protein concentration.

A previous study reported three chromosomal regions (1B, 6B, and 7B) associated with many quantitative trait loci (QTL) co-located for grain quality traits in a doubled haploid (DH) population derived from the cross CO940610/Platte. To validate those QTL, three objectives of this study were (1) QTL mapping in a CO940610/Platte recombinant inbred line (RIL) population, (2) transferring alleles of interest from CO940610 to the recurrent parent Platte by marker-assisted backcross (MABC), and (3) genome-wide association studies for grain yield (Gy), Gpc, grain protein deviation (Gpd), Gac, and test weight (Tw) in an association mapping panel.

A population of 186 CO940610/Platte RIL was grown in the Akron rainfed and Greeley fully irrigated environments in 2009/10. The same set of RIL was grown in a CSU Plant Sciences greenhouse for DNA extraction, and genotypes were obtained for 18 simple sequence repeat and

sequence tagged site markers in three chromosome regions of interest. JoinMap 4.0 was used to construct linkage maps from the molecular marker data. Marker-trait associations (MTA) were detected by single-factor analysis of variance (ANOVA). Linkage maps constructed in the CO940610/Platte RIL and DH populations were mostly consistent. Most of the grain quality traits investigated were associated with the three chromosome regions on 1B, 6B, and 7B in at least one environment, confirming findings in the CO940610/Platte DH population.

Five selected DH lines and the recurrent parent Platte were used during MABC, resulting in 35 BC_3F_2 lines for field trials. These lines were classified into 8 allelic combinations at the selective marker loci *Glu-B1*, *Xwmc182a*, and *Xwmc182b*, representing of the regions of interest on chromosomes 1B, 6B, and 7B, respectively. Of these allelic combinations, lines having PL-PL-CO and CO-CO-PL at *Glu-B1*, *Xwmc182a* and *Xwmc182b*, respectively, were hypothesized to have the lowest and highest Gpc. Experiments for the 35 MABC lines were conducted in Fort Collins fully irrigated (sprinkler irrigation), Greeley irrigated (drip irrigation), and Greeley water deficit (severe stress during grain filling) environments. Marker-trait associations for Gpc detected at *Xwmc182a* and *Xwmc182b* in the BC_3F_2 backcross population were consistent with findings in the CO940610/Platte DH population. The MTA for Gpc and Gac at locus *Xwmc182a* were robust across two of three environments. In the Fort Collins fully irrigated environment, Gpc of the allelic combination CO-CO-PL was significantly higher than the combination PL-PL-CO, confirming the hypothesized results.

A collection of 299 hard winter wheat cultivars and breeding lines representative of the U.S. Great Plains germplasm was evaluated for Gy, Gpc, grain protein deviation (Gpd), Gac and Tw. Experiments were designed as side-by-side moisture treatments in Greeley 2011/12 (drip irrigation, stress began pre-flowering) and Fort Collins 2012/13 (sprinkler irrigation, severe

stress during grain filling). Each treatment was arranged as an augmented design with two check varieties, each check having 15 replicates. Grain protein concentration and Gpd were highly correlated (0.72 to 0.87, $P<0.001$) in all four environments. The panel was characterized using a high-density 90,000 gene-associated single nucleotide polymorphism (SNP) genotyping platform. After removing SNP that did not meet data quality criteria, 16,052 filtered SNP were used to perform the genome-wide association studies (GWAS) conducted in the R programming environment using the 'GAPIT' package. Principal components and a kinship matrix were incorporated to correct for population structure and relatedness among individuals. A total of 40 significant MTA (according to the significance threshold of $P<1.67x10^{-4}$, suggested by Gao *et al.* 2008) were detected for the five evaluated traits (Gy, Gpc, Gpd, Gac, and Tw). Of these, two SNP (*BS00021704_51* and *Excalibur_c4518_2931*) on chromosome 6A were associated with Gy. The same SNP (*BS00064369_51*) on 4A was associated with both Gpc and Gpd. Test weight had the most MTA (17). In particular, two SNP, *BS00047114_51* and *BS00065934_51*, both associated with Tw on chromosome 3B, were robust across three of four environments investigated.

In conclusion, two narrow regions (~2 cM each) around *Xwmc182a* on 6B and *Xwmc182b* on 7B are of potential value for breeding programs. The incorporation of favorable allele combinations into a uniform background (Platte) was successful, but further investigation is needed for the MABC lines. Grain protein deviation appears to be a useful metric for increasing both Gpc and Gy. Five SNP (*BS00021704_51*, *Excalibur_c4518_2931*, *BS00064369_51*, *BS00047114_51*, and *BS00065934_5*) should be investigated further to detect candidate genes in their respective chromosome regions.

ACKNOWLEDGEMENTS

I would like to express my profound gratitude to my advisor, Dr. Patrick F. Byrne, for his guidance, support, and advice throughout my Ph.D. program. I also would like to thank my committee members, Dr. Mark A. Brick, Dr. Scott D. Haley, and Dr. Courtney E. Jahn, for their guidance and suggestions for my projects.

I am thankful to Scott D. Reid and Dr. Judy Harrington who were of great help to me during my research program. I am also grateful for all the members (current and past) of Dr. Byrne's lab and their wonderful attitudes, and great work ethic.

I would like to thank Dr. Nora L.V. Lapitan and her lab members for their guidance and support during the time I worked at the CSU Crop Genomics lab. I am also thankful to Dr. John Stromberger for his guidance during the time I worked at the CSU Wheat Quality lab.

I am grateful to Victoria Anderson and Scott Seifert (CSU wheat breeding crew), Tom Trout and Jerry Buchleiter (CSU & USDA-ARS Limited Irrigation Research Farm) for their help during the fieldwork, and Eduard Akhunov (Kansas State University) for genotyping and data management. I also would like to thank my friends, who were so supportive, encouraging, and knowledgeable.

Finally, I would like to thank my parents, my wife, my daughters, siblings, and family members for their compassion, encouragement and support.

Additionally, funding for my Ph.D. program was supported by the Vietnamese Government and the United States Department of Agriculture.

DEDICATION

This work is dedicated to my father, Tao Xuan Dao, my mother, Tan Thi Pham, and my mother-in-law, Lan Thi Tran.

TABLE OF CONTENTS

ABSTRACT ii

ACKNOWLEDGEMENTS v

DEDICATION vi

TABLE OF CONTENTS vii

LIST OF TABLES xii

LIST OF FIGURES xvi

CHAPTER 1: LITERATURE REVIEW 1

1.1. Wheat 1

1.2. Drought stress and wheat 3

1.3. Molecular markers 4

1.4. Construction of linkage maps 7

1.4.1. Linkage maps 7

1.4.2. Mapping populations 8

1.4.3. Identification of marker polymorphism 9

1.4.4. Genotyping of polymorphic markers 9

1.4.5. Linkage analysis of markers 10

1.5. QTL analysis 11

1.6. Investigated traits and their genetic control 13

1.6.1. Grain protein concentration 13

1.6.2. Test weight 16

1.6.3. Kernel characteristics 20

1.6.4. Grain yield 23

1.7. Genome-wide association study 25

1.8. Marker assisted selection 27

CHAPTER 2: VALIDATION OF QUANTITATIVE TRAIT LOCI FOR GRAIN QUALITY TRAITS IN WINTER WHEAT USING A CO940610/PLATTE RECOMBINANT INBRED LINE POPULATION 29

SUMMARY 29

2.0. INTRODUCTION 30

2.1. MATERIALS AND METHODS 33

2.1.1. Mapping population 33

2.1.2. Experimental design and trial management 34

2.1.3. Phenotypic evaluation 37

2.1.4. Statistical analysis 38

2.1.5. DNA extraction 41

2.1.6. Molecular marker evaluation 43

2.1.7. Linkage map construction 43

2.1.8. QTL analysis 45

2.2. RESULTS 45

2.2.1. Trait distribution and means 45

2.2.2. Correlation among traits 48

2.2.3. Heritability 51

2.2.4. Marker analysis 53

2.2.5. Construction of linkage map 55

2.2.6. Marker-trait associations 57

2.3. DISCUSSION 68

2.3.1. Trait means 68

2.3.2. Correlation among traits 69

2.3.3. Heritability estimates 70

2.3.4. Marker analysis and genetic map construction 71

2.3.5. Marker-trait associations 72

CHAPTER 3: VALIDATION OF QUANTITATIVE TRAIT LOCI FOR GRAIN QUALITY TRAITS IN WINTER WHEAT USING A CO940610/PLATTE BACKCROSS POPULATION 78

SUMMARY 78

3.0. INTRODUCTION 79

3.1. MATERIALS AND METHODS 82

3.1.1. Mapping population 82

3.1.2. Genotyping 86

3.1.2. Experimental design and trial management 87

3.1.3. Phenotypic evaluation 89

3.1.4. Statistical analysis 90

3.2. RESULTS 91

3.2.1. Trait distribution and means 91

3.2.2. Correlation among traits 94

3.2.3. Heritability 95

3.2.4. Detection of significant marker-trait associations 97

3.2.5. Epistatic interactions 98

3.2.6. Trait mean comparisons among eight genotype classes 99

3.3. DISCUSSION 104

3.3.1. Trait means, correlation and heritability 104

3.3.2. Marker-trait-associations 105

3.3.3. Combined genotype trait means 106

CHAPTER 4: GENOME-WIDE ASSOCIATION STUDY FOR GRAIN QUALITY TRAITS OF A WINTER WHEAT ASSOCIATION MAPPING PANEL UNDER TWO WATER REGIMES 109

SUMMARY 109

4.1. MATERIALS AND METHODS 114

4.1.1. Association mapping panel 114

4.1.2. Experimental design and management 115

4.1.3. Phenotypic evaluation 116

4.1.4. Phenotypic data analysis 118

4.1.5. Genome-wide association study 119

4.2. RESULTS 124

4.2.1. Trait distribution and means 124

4.2.2. Correlation among traits 127

4.2.3. Heritability of traits 129

4.2.4. Marker-trait associations 130

4.3. DISCUSSION 162

Trait means, correlation, and heritability 162

Marker-trait associations 165

CHAPTER 5: GENERAL CONCLUSIONS 169

REFERENCES 173

APPENDIX 193

LIST OF TABLES

Table 1. Expected segregation ratios for markers in different population types. 9

Table 2. QTL for grain protein concentration, test weight, kernel weight, kernel diameter, kernel hardness, and grain yield from published literature. 17

Table 3. Allelic or phenotypic variation for selected major genes or traits of CO940610 and Platte winter wheat† 35

Table 4. Markers and primer sequences used in this study..... 44

Table 5. Means, standard errors (SE) and ranges for nine traits of the CO940610/Platte population (n=186) in Akron and Greeley in the 2009/10 growing season..... 47

Table 6. Means for nine traits of the two parents, CO940610 and Platte, at Akron and Greeley, CO in the 2009/10 growing season. 48

Table 7. Pearson correlation coefficients among traits of the CO940610/Platte population (n=186) at Akron and Greely in the 2009/10 growing season. Correlations for Akron are below the diagonal and those for Greeley are above the diagonal. 50

Table 8. Heritability estimates (H^2) and 90% confidence intervals for nine traits of CO940610/Platte RIL population in Akron and Greeley in the 2009/10 growing season. 52

Table 9. The goodness of fit of observed marker data for the CO940610/Platte RIL population based on deviation from expected segregation for the $F_{5:6}$ generation. 54

Table 10. Markers associated with traits of the CO940610/Platte RIL population in Akron and Greeley, CO in the 2009/10 growing season. 58

Table 11. Number of significant marker-trait associations by chromosome. Stable associations were those detected in both environments and unstable ones were detected only in one environment. 66

Table 12. Environments† in which QTL were detected in the doubled haploid (DH) (El-Feki 2010) and recombinant inbred line (RIL) populations derived from the cross CO940610/Platte. 67

Table 13. Marker associations with grain protein concentration in the CO940610/Platte DH population in three environments (El-Feki 2010). 82

Table 14. Allelic constitution of five selected CO940610/Platte DH lines used for backcrossing to Platte as the recurrent parent. 83

Table 15. Eight combinations of alleles at the loci *Glu-B1, Xwmc182a, and Xwmc182b* on chromosomes 1B, 6B, and 7B, respectively. 84

Table 16. The 35 selected plants and their genotypes of the CO940610/Platte DH (BC_3) backcross population. 85

Table 17. Means, standard errors, and ranges for traits of the CO940610/Platte backcross population (n=35) for the ARDEC wet treatment and the LIRF wet and dry treatments in the 2012/13 growing season. 93

Table 18. Pearson correlation coefficients among traits of the CO940610/Platte backcross population (n=35) at LIRF, Greeley, CO in the 2012/13 growing season. Correlations for the wet treatment are above the diagonal and those for the dry treatment are below the diagonal. 94

Table 19. Pearson correlation coefficients among traits of the CO940610/Platte backcross population (n=35) in the ARDEC wet treatment, Fort Collins, CO in the 2012/13 growing season. 95

Table 20. Broad-sense heritability estimates (H^2) and 90% confidence intervals for seven traits of the CO940610/Platte BC_3F_2 backcross population in the ARDEC wet, LIRF wet, and LIRF dry environments in the 2012/13 growing season. 96

Table 21. Significance of loci detected with analysis of variance for single factors or digenic epistatic interactions for seven traits in the CO940610/Platte BC_3F_2 population in the 2012/13 growing season. 98

Table 22. Genotype class means for the epistatic interaction of loci *Glu-B1* and *Xwmc182a* in the CO940610/Platte BC_3F_2 population. 99

Table 23. Genotype class means for the epistatic interaction of loci *Glu-B1* and *Xwmc182b* in the CO940610/Platte BC_3F_2 population. 99

Table 24. Least squares means of seven traits in each genotype class of the CO940610/Platte BC_3F_2 population in the ARDEC wet treatment in the 2012/13 growing season. Genotype classes are defined in Table 15. 101

Table 25. Least squares means of seven traits in each genotype class of the CO940610/Platte BC_3F_2 population in the LIRF wet treatment in the 2013 growing season. Genotype classes are defined in Table 15. 102

Table 26. Least squares means of seven traits in each genotype class of the CO940610/Platte BC_3F_2 population in the LIRF dry treatment in the 2013 growing season. Genotype classes are defined in Table 15. 103

Table 27. Distribution of markers across genomes and chromosome in the hard winter wheat association-mapping panel (HWWAMP), provided by Mary Guittieri (University of Nebraska-Lincoln, personal communication). 120

Table 28. Positive R^2 values obtained with different model combinations used for marker-trait association (MTA) identification. 121

Table 29. Means, standard errors (SE) and ranges for five evaluated traits of the Hard Winter Wheat Association Mapping Panel in LIRF 2011/12 and ARDEC 2012/13 growing seasons. . 126

Table 30. Pearson correlation coefficients (n = 295 to 299) among traits of the Hard Winter Wheat Association Mapping Panel at LIRF, Greeley, CO in the 2011/12 growing season. Correlations for the wet treatment are above the diagonal and those for the dry treatment are below the diagonal. 128

Table 31. Pearson correlation coefficients among traits of the Hard Winter Wheat Association Mapping Panel at ARDEC, Fort Collins, CO in the 2012/13 growing season. Correlations for the wet treatment are above diagonal and those for the dry treatment are below the diagonal. 128

Table 32. Broad-sense heritability estimates (H^2) for five evaluated traits of the Hard Winter Wheat Association Mapping Panel in LIRF 2011/12 and ARDEC 2012/13 growing seasons. . 129

Table 33. Marker-trait associations detected in the Hard Winter Wheat Association Mapping Panel at the unadjusted *P*-value < 0.001 for five traits in four environments. 131

Table 34. Strong marker-trait associations in the Hard Winter Wheat Association Mapping detected at $P < 1.67 \times 10^{-4}$ for five traits in four environments. 159

Table 35. Number of MTA detected at $P < 0.001$ and at $P < 1.67 \times 10^{-4}$ 161

Table 36. Precipitation and irrigation from January to July 15 for 13 environments 193

Table 37. Monthly maximum and minimum temperature (^{o}C) and precipitation (mm) from January to July 15 for seven location-years, where experiments involved were conducted. 194

LIST OF FIGURES

Figure 1. The evolutionary and genomic relationships between cultivated bread and durum wheats and related wild diploid grasses (Shewry 2009). 2

Figure 2. An example of the output of SIM and CIM methods of chromosome 1 of maize for silk maysin concentration (Dr. Patrick Byrne, Plant and Soil Science eLibrary). 12

Figure 3. Classification of wheat grain proteins (Tazzini 2015). 14

Figure 4. The response of wheat grain yield and grain protein to increasing N (Jones & Olson-Rutz 2012). 15

Figure 5. Wheat grain classes (http://www.uswheat.org/wheatClasses). 22

Figure 6. The steps for performing AM and identifying candidate genes (Abdurakhmonov & Abdukarimov 2008). 26

Figure 7. Linkage map constructed in JoinMap for the CO940610/Platte RIL population. Cumulative distances between markers are given in cM, calculated from recombination frequencies according to the Haldane mapping function. 56

Figure 8. QTL maps of the CO940610/ Platte/ RIL population with location of marker-trait associations indicated. Cumulative distances between markers are given in cM. A = detected in Akron, G = detected in Greeley. Gy, grain yield; Sl, plant height; Sl, spike length; Gpc, grain protein concentration; Gac, grain ash concentration; Tw, test weight; Kw, kernel weight; Kd, kernel diameter; Kh, kernel hardness. 61

Figure 9. Scheme of selection and developing backcross populations. 83

Figure 10. Manhattan plots for Gpc in the LIRF Dry with the different kinship matrices used for analysis. The X-axis is the genomic position of each SNP; the Y-axis is the negative logarithm of

the P-value obtained from the GWAS model. The lines identify the two threshold lines of significance. A, Loiselle kinship matrix used; B, IBS kinship matrix used; C, rrBLUP kinship matrix used. UM, unmapped SNP. 122

Figure 11. Manhattan plots for five traits of the Hard Winter Wheat Association Mapping Panel evaluated in four environments. The X-axis is the genomic position of each SNP; the Y-axis is the negative logarithm of the *P*-value obtained from the GWAS model. The lower line represents the significance threshold proposed by Gao et al. (2008) and the upper line is the Bonferroni significance threshold. UM, unmapped SNP. 144

Figure 12. QQ plots for five traits investigated in four environments: A, ARDEC Dry; B, ARDEC wet; C, LIRF Dry; D, LIRF Wet. 154

Figure 13. Frequency distributions for the traits for CO940610/Platte RIL population in 2009/10 growing season. *P*-values are for the Shapiro-Wilk test of normality, with * ($P<0.05$) indicating deviation for normality. 195

Figure 14. Frequency distributions for seven traits for the CO940610/Platte//Platte backcross populations in the 2013 growing seasons. *P*-values are for the Shapiro-Wilk Test of normality, with * ($P<0.05$) indicating deviation from normality. 200

Figure 15. Frequency distributions for five traits investigated in the Hard Winter Wheat Association Mapping Panel. *P*-values are for the Shapiro-Wilk Test of normality, with * ($P<0.05$) indicating deviation from normality. 207

CHAPTER 1: LITERATURE REVIEW

1.1. Wheat

Common wheat (*Triticum aestivum* L.) and durum wheat (*Triticum turgidum* L.) are members of the family Poaceae (http://plants.usda.gov). The species of *Triticum* are grouped into three ploidy classes: diploid (2n = 2x = 14), tetraploid (2n = 4x = 28), and hexaploid (2n = 6x = 42). Common wheat is an allohexaploid (AABBDD), and has six copies of each of its seven chromosomes, with the complete set numbering 42 chromosomes (Sears 1954). *Triticum turgidum* (AABB) evolved as an alloploid combining genomes from the diploid species *T. urartu* (AA, 2n = 2x = 14) and an unknown and possibly extinct diploid species related to *Aegilops speltoides* (2n = 2x = 14, BB) containing the B genome (Matsuoka 2011). Subsequently, bread wheat was formed through hybridization between cultivated tetraploid emmer wheat (AABB, *T. dicoccoides*) and diploid goat grass (DD, *Ae. tauschii*) approximately 8,000 years ago (Daud & Gustafson 1996; Haider 2013; Gale & Devos 1997). Spikes and grains of these species are shown in Figure 1.

Common wheat has a large genome about 17 Gb, with three complete genomes A, B, and D in the nucleus of its cells (Paux et al. 2006). Its genome size is approximately 5 times, 35 times, and 110 times larger than that of humans (*Homo sapiens*), rice (*Oryza sativa* L.) and *Arabidopsis thaliana*, respectively (Syed & Rivandi 2007). The genome is also highly repetitive and complex. Repetitive DNA accounts for approximately 90% of the genome, of which transposable elements constitute 60-80% (Wanjugi et al. 2009).

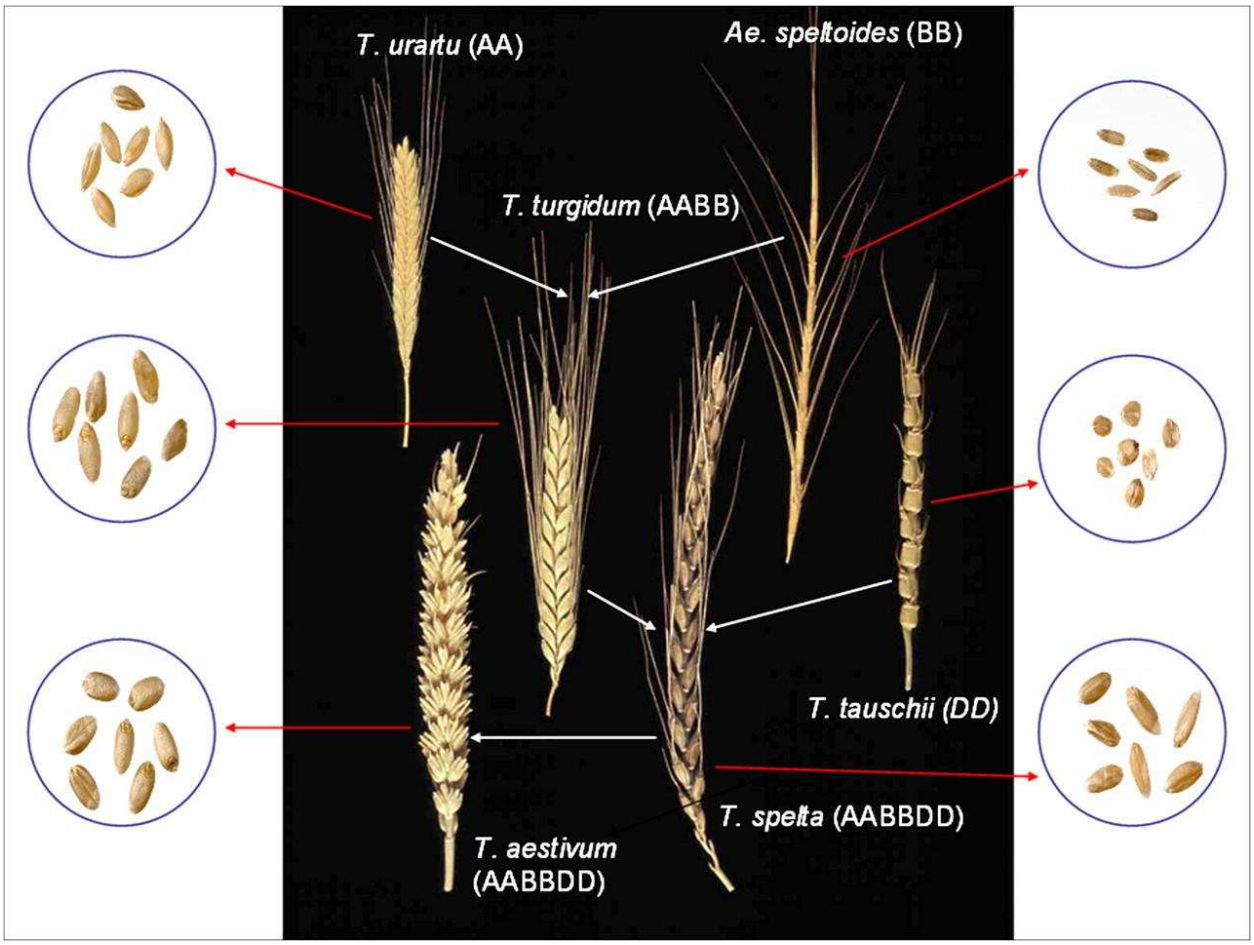


Figure 1. The evolutionary and genomic relationships between cultivated bread and durum wheats and related wild diploid grasses (Shewry 2009).

Cultivated common wheat is exceptionally diverse in the physiological characteristics that adapt different wheat cultivars for a wide range of climatic environments. It is also diverse in the chemical and physical characteristics of the gluten proteins that contribute to the wide use of wheat grain for many different food products. Wheat's physiological characteristics are generally related to vernalization requirement, winter hardiness or cold tolerance, and photoperiod response (Sleper & Poehlman 2006).

Wheat is the most widely grown crop in the world, with 219 million ha in 2013 (FAO 2013). Approximately 20% of the daily protein and about the same proportion of the total calories of the world's population come from wheat (FAO 2013). Its annual production was the third highest of cereal crops after only maize (*Zea mays* L.) and rice (*Oryza* spp.) (FAO 2013). In addition to serving as a vehicle for carbohydrates and protein, wheat also is a source of vitamins, minerals, fiber, magnesium, folic acid, antioxidants, and other phytochemicals. Therefore, it is an important component of food security globally.

1.2. Drought stress and wheat

In agricultural terms, drought is insufficient soil moisture to meet the needs of a particular crop at a particular time (FAO 2013). As a result, drought stress reduces crop productivity and sometimes increases plant disease severity. The mechanisms of drought tolerance are classified into three categories. *Drought escape* is the ability of a plant to complete its life cycle before serious soil and plant water deficits develop; an example is earlier flowering in annual species before the onset of severe drought (Turner 1986). *Drought avoidance* is the ability of plants to maintain relatively high tissue water potential despite a shortage of soil moisture. This is accomplished, for example, by developing deep root systems, and reducing stomatal density and leaf area (Blum 1988). *Drought tolerance* is the ability of a plant to withstand water deficit with low tissue water potential, for instance, improving osmotic adjustment ability and increasing cell wall elasticity to maintain tissue turgidity (Fleury et al. 2010).

Plants respond to drought by complex mechanisms, including molecular expression, biochemical metabolism, individual plant physiological processes, and crop canopy behavior (Chaves et al. 2003; Xu et al. 2009; Kadam et al. 2012). Drought tolerance is a quantitative and complex trait, as drought induces the up- or down-regulation of thousands of drought-responsive genes

according to growth stage, plant organ, and time of day (Blum 2011). Drought tolerance traits typically are multi-genic, and have low heritability and high genotype by environment interaction (Fleury et al. 2010). Drought stress may also occur simultaneously with other abiotic stress, such as high temperatures, high irradiance, and nutrient toxicities or deficiencies.

Drought stress drastically reduces wheat grain yield, up to a 100% reduction in comparison to yield under full irrigation in several recent reports (Edae et al. 2014; El-Feki et al. 2013; Nezhad et al. 2012). However, drought stress usually increases grain protein concentration in wheat. Drought stress reduces starch accumulation and increases protein concentration due to smaller grain size (Balla et al. 2011). Drought conditions rapidly increase the quantity of insoluble protein in the wheat grain (Daniel & Triboı 2002). Drought stress also reduces green leaf area and the plant's ability to fix dry matter during the grain filling, thus decreasing starch accumulation in grain (Foulkes et al. 2002), and increasing grain protein concentration (Weightman et al. 2008).

1.3. Molecular markers

A deoxyribose nucleic acid (DNA) marker, also called a molecular marker, is a particular sequence of DNA that reveals sites of variation in DNA within the context of an entire genome (Jones et al. 1997; Winter & Kahl 1995). Markers are formed by point mutations, rearrangements, or errors in replication of tandemly repeated DNA (Paterson 1996), and are usually located in non-coding regions of DNA. Molecular markers are widely used because of their abundance. They are practically unlimited in number and not affected by environmental conditions and/or the developmental stage of the plant (Winter & Kahl 1995). The numerous applications of molecular markers include the construction of linkage maps, assessment of the level of genetic diversity within germplasm, and establishing cultivar identity (Jahufer et al.

2003; Winter & Kahl 1995). Based on the method of detection, there are three classes of molecular markers, including hybridization-based, polymerase chain reaction-based, and DNA sequence-based (Gupta et al. 1999; Jones et al. 1997; Winter & Kahl 1995). Genetic polymorphism revealed by molecular markers can be visualized by gel electrophoresis and staining with ethidium bromide or silver or by radioactive or colorimetric probes. Based on whether markers can discriminate between homozygotes and heterozygotes, there are two types of polymorphic markers, that is, dominant and codominant. Dominant markers are expressed as either presence or absence of a DNA fragment, while codominant markers indicate difference in fragment size, discriminating between homozygous and heterozygous genotypes. The most commonly used molecular markers, either historically or currently, are restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), simple sequence repeats (SSR), amplified fragment length polymorphism (AFLP), and single nucleotide polymorphism (SNP). Of these, SSR and SNP are the two types of marker most commonly used in wheat studies today.

SSR

The wheat genome is large (about 17 Gb) (Paux et al. 2006) and has an abundance of repetitive DNA (Chantret et al. 2005; Paux et al. 2008; Wanjugi et al. 2009; Lagercrantz et al. 1993; Tautz & Renz 1984). There are repetitive elements that are dispersed throughout the genome, including transposable elements and tandemly repeated DNA (Sehgal et al. 2012). The tandemly repeated DNAs with a repeat length of up to 13 bases are known as SSR (Jacob et al. 1991) or microsatellites (Litt & Luty 1989) or short tandem repeats (Edwards et al. 1991), whereas those with longer repeats are termed minisatellites (Ramel 1997). The most common classes of microsatellites are dinucleotide, trinucleotide, and tetranucleotide repeats (Ramel 1997). SSR are

present both in coding and noncoding regions (Katti et al. 2001) and are distributed throughout the nuclear genome (Cavagnaro et al. 2010), as well as in the chloroplast (Bryan et al. 1999) and mitochondrial genomes (Zhao et al. 2014). SSR tend to be highly polymorphic and codominant (S. Zhang et al. 2014). The high length polymorphism is caused by a different number of repeats in the microsatellite region (Tautz & Renz 1984). Therefore, SSR can be easily and reproducibly detected by PCR followed by separation on agarose or polyacrylamide electrophoresis gels (Cosson et al. 2014). Alternatively, if fluorescent dyes are incorporated into the primers, SSR can be detected by a capillary sequencer (Hayden et al. 2008). The major disadvantage of SSR is the large amount of time and effort required to detect SSR sequences, then design and test primers (Zane et al. 2002; Stackelberg et al. 2006). Various applications of SSR in wheat breeding and research programs have included quantitative trait locus (QTL) mapping (Chu et al. 2008; El-Feki et al. 2013), tagging of resistant genes (Alam et al. 2011), marker-assisted selection (MAS) (Parisod et al. 2013), and assessing diversity (Rector et al. 2013; Salem et al. 2008).

Single nucleotide polymorphisms

The acronym SNP, pronounced "snip", stands for single nucleotide polymorphism. A SNP is a variation between individuals in a single base at the same position in a DNA sequence. SNP can be changed from one base to another through transitions and transversions. They can also be single-base insertions and deletions, called "indels". SNP are classified into noncoding SNP and coding SNP. Noncoding SNP may be located in a 5' or 3' nontranscribed region, a 5' or 3' untranslated region, an intron, or an intergenic region. Meanwhile, coding SNP are located in coding regions or exons. Therefore, coding SNP may change the amino acid that is encoded, known as replacement polymorphism, or change the codon but not the amino acid, called synonymous polymorphisms (Gibson & Muse 2009).

There is great potential for the use of SNP in the detection of association between alleles and traits, and for the use for identification in the vicinity of virtually every gene (Rafalski 2002; Mammadov et al. 2012). Mammadov et al. (2012) reported that a vast majority of publications used SNP for quantitative trait loci (QTL) mapping. The combination of three key phrases ("marker-assisted" AND "SNP" AND "plant breeding" showed about 4,560 articles from 2006 to 2012 indicating the applications of SNP in marker-assisted selection (MAS). Several studies on association mapping in plants have been published and reviewed (Zhu et al. 2008; Rafalski 2010; Abdurakhmonov & Abdukarimov 2008). Sukumaran et al. (2015) conducted a genome-wide association study (GWAS) using 18,704 SNP to identify 31 loci associated with yield and related traits in wheat, explaining 5-14% of phenotypic variation. Plessis et al. (2013) used SNP and other markers to conduct GWAS and identify candidate genes for grain protein concentration and composition in wheat.

1.4. Construction of linkage maps

1.4.1. Linkage maps

A genetic map or linkage map is considered a road map of the chromosomes derived from the cross of two genetically distinct parents (Paterson 1996). It is a diagrammatic representation of the linear order and relative genetic distance between genes or markers along chromosomes derived from frequencies of recombination. The markers are analogous to signs or landmarks along a road. One of the most important uses for linkage maps is to identify chromosomal locations of genes and QTL associated with traits of interest (Collard et al. 2005).

Genetic mapping is based on the principle that genes and marker loci segregate via chromosome recombination (also called cross-over or strand exchange) during meiosis (i.e., sexual reproduction), thus allowing their analysis in the progeny (Paterson 1996). When two genes or

markers are close together or tightly-linked on the same chromosome, they will be transmitted together from parent to progeny more frequently than those genes or markers located far apart. Linked genes or markers have a recombination frequency of less than 50%, while unlinked genes or markers have a recombination frequency of 50% or more, located further apart on the same chromosome or on different chromosomes (Hartl & Jones 2001). The lower the frequency of recombination between markers the closer they are situated on a chromosome, and conversely, the higher the recombination frequency between genes or markers, the more distant are their chromosome locations. Requirements for linkage map construction are (1) development of a mapping population; (2) identification of marker polymorphisms; (3) genotyping of polymorphic markers; and (4) linkage analysis of markers.

1.4.2. Mapping populations

The first step in creating a mapping population is selection of genetically divergent parents for one or more traits of interest. Population size used in preliminary genetic mapping studies varied from 50 to 250 individuals (Mohan et al. 1997), however a larger population is necessary for high resolution mapping. Different types of populations utilized for mapping are F_2, recombinant inbred line (RIL), near isogenic line (NIL), doubled haploid (DH), and backcross populations. An F_2 population provides maximum genetic information by using codominant markers, while RIL, NIL, and DH populations have nearly the same genetic information for both codominant and dominant markers (Table 1). A backcross population using either codominant or dominant markers has less genetic information than an F_2 population.

The following examples demonstrate the types of populations used for linkage map construction and QTL studies in wheat. Two $F_{8:9}$ RIL populations derived from crosses between three common Chinese wheat varieties, consisting of 229 and 302 lines were conducted under two

water conditions for analysis of 12 traits of wheat seedlings (Zhang et al. 2013). El-Feki et al. (2013) used 185 DH lines of hard white winter wheat for a grain quality trait QTL study, and Barakat et al. (2011) developed an F_2 wheat population (n=162) to identify new microsatellite markers linked to the grain filling rate. Ibrahim et al. (2012) used a $BC_2F_{4:6}$ population of 223 lines for advanced backcross QTL analysis of drought tolerance in spring wheat.

Table 1. Expected segregation ratios for markers in different population types.

	Markers	
Population types	Codominant	Dominant
F_2	1:2:1 (AA:Aa:aa)	3:1 (B_:bb)
Backcross	1:1 (Cc:cc)	1:1 (Dd:dd)
Recombinant inbred or doubled haploid	1:1 (EE:ee)	1:1 (FF:ff)

Source: (Collard et al. 2005).

1.4.3. Identification of marker polymorphism

Besides differing for phenotypic traits, the parents of a mapping population must also be sufficiently polymorphic at the molecular level in order to construct a linkage map (Young 1994). Therefore, identification of molecular markers that reveal differences between parents is an essential step in the construction of a linkage map. Wheat is an inbreeding species, which results in lower levels of DNA polymorphism than cross-pollinating species, so selection of parents that are more distantly related is often required.

1.4.4. Genotyping of polymorphic markers

Once polymorphic markers have been identified, the entire mapping population, including parents must be genotyped (Collard et al. 2005). DNA from each individual of the mapping population is extracted, then genotyped by polymorphic markers. Genotyping approaches are based on characteristics and availability of markers chosen for mapping, which differ among

species (Young 1994). The key factor is efficient use of time, labor, and supplies. Since markers are screened and scored for a whole mapping population, the "goodness of fit" of observed maker data to the expected segregation ratios can be tested with a chi-square statistic (Collard et al. 2005) using the following formula:

Chi-square = $\Sigma[(\text{observed} - \text{expected})^2/\text{expected}]$. If degree of freedom is 1, this adjustment is sometimes used: $\Sigma[(|\text{observed} - \text{expected}| - 0.5)^2/\text{expected}]$. Calculated values are then compared to values in the chi-square table to determine if observed ratios conform to expected values (Griffiths et al. 2000).

Expected ratios of genotypes depend on the type of marker and population (Table 1). As an example, populations of 47% AA or BB and 6% AB may be described for the expected segregation ratios for SSR in a RIL F_5-derived population. However, distorted segregation ratios may be encountered (Sayed et al. 2002; Xu et al. 1997).

1.4.5. Linkage analysis of markers

Two major outputs from the marker evaluation are a file of marker scores and a genetic linkage map. Linkages between markers are usually calculated using the ratio of the probability of linkage versus no linkage, also called odds ratios. The alternative expression of the odds ratios is as the logarithm of the ratio, also known as a logarithm of odds (LOD) value (Risch 1992). LOD values of greater than 3 (typically used to construct linkage maps), between two markers indicate that linkage is 10^3 times more likely than no linkage. Linear arrangements of markers are clustered into "linkage groups". Linkage groups represent chromosomal segments or entire chromosomes, while markers represent signposts or landmarks on them.

A linkage map can be constructed manually or by using a computer program. Manual construction is feasible for a few markers, but computer programs are required for larger numbers of markers. Commonly-used software programs include Map-maker/EXP (Lander et al. 1987; Lincoln et al. 1993), MapManager QTX (Manly et al. 2001), R/QTL (Broman et al. 2003) and JoinMap (Stam 1993).

1.5. QTL analysis

Quantitative traits are characterized by a continuous distribution of phenotypic variation as a result of the combined effects of many genes interacting with the environmental factors (Falconer & Mackay 1996). For example, grain yield and grain protein concentration are quantitative traits. The genetic loci controlling quantitative traits are referred to as quantitative trait loci (QTL). QTL analysis is a statistical method that connects phenotypic data and genotypic data (usually molecular markers) in an attempt to explain the genetic basis of variation in quantitative traits (Falconer & Mackay 1996; Kearsey 1998). Single-marker analysis (SMA), simple interval mapping (SIM), and composite interval mapping (CIM) are three widely used methods for detecting QTL (Tanksley 1993; Liu 1997).

The statistical methods used for SMA include *t*-tests, analysis of variance (ANOVA), and linear regression. This method does not require a complete linkage map and needs only a basic statistical software programs, e.g., QGene (Nelson 1997) and MapManager QTX (Manly et al. 2001). However, the further a QTL is from a marker, the less likely it is to be detected. QTL are usually reported in a table with chromosome or linkage group, markers, probability value (*P*-value), and the percentage of phenotypic variation explained by the QTL (R^2). The SIM method uses linkage maps and analyzes intervals between adjacent pairs of linked markers along chromosomes. This method is considered statistically more powerful than SMA (Lander &

Botstein 1989; Liu 1997). Software programs MapMaker/EXP (Lincoln et al. 1993) and QGene (Nelson 1997) have been used to conduct SIM. The CIM method combines interval mapping with linear regression and uses additional markers besides an adjacent pair of linked markers in the statistical model (Jansen 1993; Jansen & Stam 1994; Zeng 1994). The CIM approach is more precise and effective at QTL mapping than SMA and SIM. The CIM method can be performed with QTL Cartographer (Basten et al. 2005), MapManager/EXP (Manly et al. 2001), R/QTL (Broman et al. 2003) and PLABQTL (Utz & Melchinger 1996). QTL detected using interval mapping are located with respect to a linkage map. The test statistic for SIM and CIM is a logarithm of odds (LOD) score or likelihood ratio statistic (LRS). A typical output from interval mapping is a graph with linkage distance on the x-axis and the test statistic on the y-axis (Figure 2). In order to avoid an excessive number of false positive results and ensure that true indications of linkage are not missed, there may be suggestive, significant, and highly significant QTL (Lander & Kruglyak 1995).

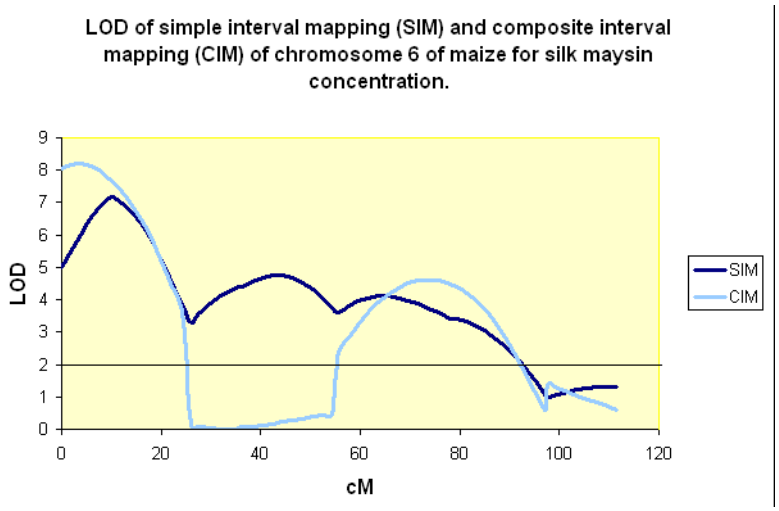


Figure 2. An example of the output of SIM and CIM methods of chromosome 1 of maize for silk maysin concentration (Dr. Patrick Byrne, Plant and Soil Science eLibrary).

The detection of QTL segregating in a population is affected by many factors (Tanksley 1993; Asins 2002). The genetic properties of QTL that control the trait, environmental effects, population size, and experimental errors are the main factors. QTL with sufficiently large phenotypic effects can be detected routinely, but small effect QTL are more difficult to detect. Closely-linked loci affecting a trait are usually detected as a single QTL with typical population sizes (Tanksley 1993). The expression of quantitative traits is influenced by environmental conditions, so conducting experiments in multiple environments allows investigation of the effect of environment on QTL of interest. The larger the population, the more accurate the mapping and the more likely is detection of QTL with small effects (Tanksley 1993). The order and distance between markers in linkage maps can be influenced by mistakes in marker genotyping and detected QTL positions can be affected by errors in phenotypic evaluation. Only reliable genotypic and phenotypic data can produce a reliable QTL map (Collard et al. 2005).

1.6. Investigated traits and their genetic control

1.6.1. Grain protein concentration

Wheat protein concentration and composition are important for end-use quality, with different products requiring different protein amounts and patterns. Wheat proteins are comprised of gluten proteins, about 80-85% of total wheat grain protein, and a highly heterogeneous group of non-gluten proteins, about 15-20% of the total (Veraverbeke & Delcour 2002) (Figure 3). All non-gluten proteins are considered to be soluble in salt solutions. In contrast, gluten proteins have low solubility in water or salt solutions (Veraverbeke & Delcour 2002). The gluten proteins form the major class of wheat storage proteins and have extreme importance because they are responsible for the unique visco-elastic properties of wheat flour dough (Payne 1987).

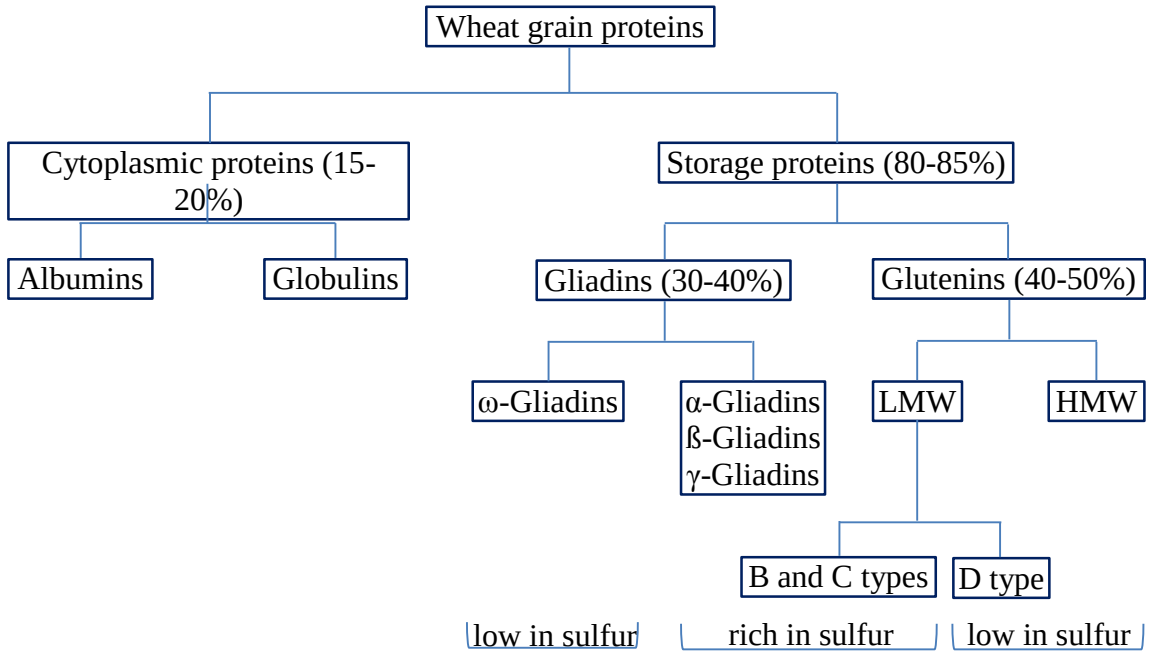


Figure 3. Classification of wheat grain proteins (Tazzini 2015).

Gliadins and glutenins are two recognized storage protein groups in the endosperm. Gliadin molecules are small, about 35 KDa, and are separated into four groups, α, ß, γ, and ω, by gel electrophoresis at low pH (Wall 1979). Glutenins are large, heterogeneous molecules, and fall predominantly into the low-molecular-weight subunit (LMW-GS) (Jackson et al. 1983) and the high-molecular-weight subunit (HMW-GS) classes (Payne et al. 1982).

In general, grain protein concentration of wheat ranges from 8 to 16% of the dry weight (Payne 1987; El-Feki et al. 2013), but higher values are sometimes observed (Dr. Scott Haley, personal communication). El-Feki et al. (2013) reported that Gpc ranged from 10.0 to 16.7% in the CO940610/Platte DH population. Wheat Gpc is a quantitative trait that is controlled by multiple genes and thus, is significantly affected by environmental conditions. Soil moisture and nitrogen nutrition are two major environmental factors influencing Gpc. For example, drought stress was associated with an increase in Gpc (El-Feki et al. 2013; Zheng et al. 2009) because under drought

conditions less starch is accumulated (Ahmadi & Baker 2001), thus, increasing the proportion of protein in the kernel (Weightman et al. 2008). Wheat Gpc at harvest can be either increased or decreased with increased available N during vegetative growth, depending on the severity of the N deficiency (Jones & Olson-Rutz 2012; Brown et al. 2005) (Figure 4).

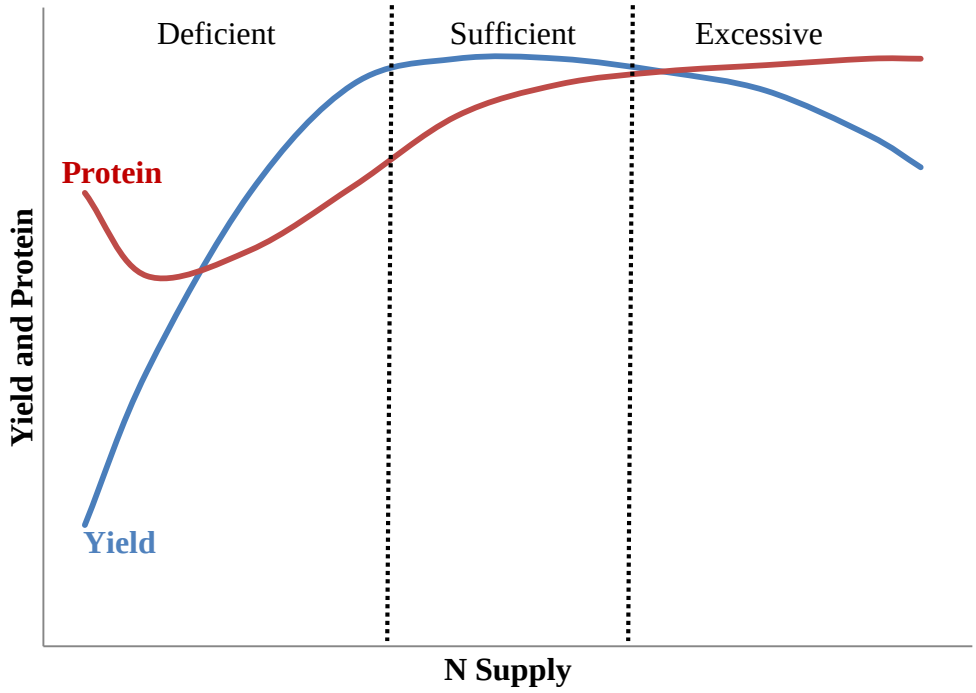


Figure 4. The response of wheat grain yield and grain protein to increasing N (Jones & Olson-Rutz 2012).

Increased grain yield (Gy) and an appropriate Gpc are two main goals of wheat breeding programs (Bogard et al., 2010). Unfortunately, these traits have an inverse relationship (El-Feki et al. 2013; Kibite & Evans 1984; Pleijel 1999; Wang et al. 2012), that is, the higher the grain yield, the lower the Gpc. Under water stress Gy decreases, but Gpc increases (Edae et al. 2014; El-Feki et al. 2013).

Multiple loci that control Gpc are distributed throughout the wheat genome in many different chromosome regions. Many researchers have identified QTL for wheat Gpc in different mapping populations (mostly RIL and DH) and various environments (Table 2).

1.6.2. Test weight

Test weight is a measure of grain bulk density, the weight of wheat kernels at a given grain moisture (12%) in a specific volume. It is usually reported as kg/hL. Test weight is used as an indicator of general grain quality, and provides a rough estimate of potential flour yield. Higher Tw normally means higher quality grain. Test weight generally increases as grain is dried. With drying, the proportion of water in the kernels decreases, so bulk density of the kernel increases. In fact, the type of dryer and drying methods can affect Tw during drying. Cleaning also increases the Tw by removing small and damaged kernels. In contrast, Tw decreases as grain deteriorates. Poor growing conditions, especially at grain filling, and sprout damage can cause lower Tw. The lower Tw may require a greater volume of grain storage and transportation for the same grain weight, therefore, adding more expense to the growers (El-Feki 2010; Shelton et al. 2008).

Test weight is also affected by genotype. Several researchers have reported QTL that control Tw of wheat grain (Table 2) in different populations (mostly RIL, DH) and environments. El-Feki et al. (2013) found four QTL for Tw on chromosomes 1B, 6B, 7A and 7D in a DH population in a single environment. The phenotypic variation explained by the QTL ranged from 5.6 to 7.9%.

Table 2. QTL for grain protein concentration, test weight, kernel weight, kernel diameter, kernel hardness, and grain yield from published literature.

Trait	Population	QTL	No. of lines	No. of env.	References
Gpc†	Messapia x MG4343 (RIL)	4S, 5AL, 6AS, 6BS, 7AS, 7BS	65	8	Blanco et al. 2002
	Latino x MG29896 (BIL)	2AS, 6AS, 7BL	92	4	Blanco et al. 2006
	PH132 X WL711 (RIL)	2BL, 7AS	106	2	Dholakia et al. 2001
	CO940610 x Platte (DH)	5B1, 6A1, 6B1, 7B, 7D2	185	4	El-Feki et al. 2013
	Récital x Renan (RIL)	1A, 2A, 3A, 3B, 4A, 4D, 5B, 6A, 7A, 7D	194	6	Groos et al. 2003
	Récital x Renan (RIL)	3A, 5B	165	3	Groos et al. 2004
	ACKama x 87E03-S2B1 (DH)	2D, 4B, 4D, 7B	185	3	Huang et al. 2006
	Neixiang188 x Yanzhan1 (RIL)	1B, 2A, 2B, 2D, 3A, 3B, 4D, 5B, 5D, 7B, 7D	198	2	Li et al. 2009
	Kukri x Janz (DH)	1B, 3A, 7A	160	5	Mann et al. 2009
	Sunco x Tasman (DH)	1B, 2B, 5B	163	4	Mares & Campbell 2001
	WP219 x Opata85 (RIL)	2A, 2D, 6D	114	5	Nelson et al. 2006
	PDW233 x Bhalegaon4 (RIL)	7B	140	5	Patil et al. 2009
	Courtot x CV (DH)	1B, 6A	187	2	Perretant et al. 2000
	PH132 x WL711 (RIL)	2D	100	1	Prasad et al. 1999
	WL711 x PH132 (RIL)	2AS, 2BL, 2DL, 3DS, 4AL, 6BS, 7AS, 7DS	100	5	Prasad et al. 2003
	Chara x WW2449 (DH)	4A	190	2	Raman et al. 2005
	Courtot x Chinese Spring (DH)	1BL, 6AS	217	5	Sourdille et al. 2003
	Ning7840 x Clark (RIL)	3AS, 4B	132	7	X. Sun et al. 2010
	DT695 X Strongfield (DH)	1A, 1B, 2A, 2B, 5B, 6B, 7A, 7B	185	6	Suprayogi et al. 2009
	Beaver x Soissons (DH)	1B, 3A, 3B, 4D, 5D, 7A, 7D	46	2	Weightman et al. 2008
Tw	Wichita x Cheyenne (RIL)	3A	98	7	Campbell et al. 2003
	CO940610 x Platte (DH)	1B1, 6B1, 7A2, 7D2	185	4	El-Feki et al. 2013
	ACKama x 87E03-S2B1 (DH)	2D, 4A, 4D, 5A, 7A	185	3	Huang et al. 2006
	SeriM82 x Babax (RIL)	2B, 3B, 4D, 7A	194	8	McIntyre et al. 2010

Table 2. (Continued)

Trait	Population	QTL	No. of lines	No. of env.	References
Tw	Karl92 x TA4152 (AB)	2D	190	2	Narasimhamoorthy et al. 2006
Kw	Rye111 x Chinese Spring (RIL)	1A, 1D, 2D, 6B	113		Ammiraju et al. 2001
	AC Reed x Grandin (DH)	2BL, 2DS	101	2	Breseghello and Sorrells, 2007
	PH132 x WL711 (RIL)	2BL, 2DL	106	2	Dholakia et al. 2003
	CO940610 x Platte (DH)	1A1, 1B1, 2B1, 2D2, 3B1, 6A1, 7D2	185	4	El-Feki et al. 2013
	ACKama x 87E03-S2B1 (DH)	2B, 2D, 3B, 4B, 4D, 6A, 7A	185	3	Huang et al. 2006
	Kukri x Janz (DH)	4B, 4D	160	5	Mann et al. 2009
	Sunco x Tasman (DH)	2B, 4D	163	4	Mares & Campbell 2001
	Ning7840 x Clark (RIL)	1BS, 2A, 3A, 3B, 4A, 4D, 5B, 6A, 7A, 7D	132	7	X. Sun et al. 2010
Kd	W7984 x Opata 85 (RIL)	1B	115	2	Igrejas et al. 2002
	NY6432-18 x Clark's Cream (RIL)	1A, 2A, 2B, 2DL	78	6	Campbell et al. 1999
	PH132 x WL711 (RIL)	2DL	106	2	Dholakia et al. 2003
	CO940610 x Platte (DH)	1A1, 2B1, 2D2, 3B1, 6A1, 7B, 7D2	185	4	El-Feki et al. 2013
	Kukri x Janz (DH)	4B, 4D	160	5	Mann et al. 2009
	Chuan35050 x Shannong483 (RIL)	2A, 5D, 6A	131	4	Sun et al. 2009
	Ning7840 x Clark (RIL)	4AL, 5AL, 5AS, 6AS	132	7	X. Sun et al. 2010
Kh	CO940610 x Platte (DH)	1D, 2B1, 3B1, 6B2, 7A2, 7D2	185	4	El-Feki et al. 2013
	Récital x Renan (RIL)	1A, 1B, 2A, 2B, 3A, 3B, 4A, 5A, 5B, 5D, 6A, 6B, 6D	165	3	Groos et al. 2004
	W7984 x Opata 85 (RIL)	5D	115	2	Igrejas et al. 2002
	Neixiang188 x Yanzhan1 (RIL)	1BL, 3B, 4B, 4D, 5A, 5B, 5D	198	2	Li et al. 2009
	Courtot x CV (DH)	1A, 5D, 6D	187	2	Perretant et al. 2000
	Opata 85 x W7984	5D, 6A, 3A	63	2	Pshenichnikova et al. 2008

Table 2. (Continued)

Trait	Population	QTL	No. of lines	No. of env.	References
Kh	W7984 x Opata 85 (RIL)	2AL, 2DL, 5BL, 5DS, 6DS	114	2	Sourdille et al. 1996
	Ning7840 x Clark (RIL)	1DL, 5B, 5DS, 5DL	132	7	Sun et al. 2010
	Beaver x Soissons (DH)	2A, 2D, 3A, 4A, 5A, 5D, 6D	46	2	Weightman et al. 2008
	PH82-2 x Neixiang 188 (RIL)	5D	214	6	Zhang et al. 2009
Gy	Chinese Spring x Kanto107 (RIL)	4A	98	2	Araki et al. 1999
	Wichita x Cheyenne (RIL)	3A	98	7	Campbell et al. 2003
	Superb x BW278 (DH)	1A, 2D, 3B, 5A	178	12	Cuthbert et al. 2008
	CO940610 x Platte (DH)	2D, 5A, 5B, 7B	185	4	El-Feki 2010
	Récital x Renan (RIL)	2B, 3B, 4A, 4B, 5A, 5B, 7D	194	4	Groos et al. 2003
	Prinz x W-7984 (AB)	1B, 2A, 2D, 5B	72	4	Huang et al. 2003
	ACKarma x 87E03-S2B1 (DH)	5A, 7A, 7B	185	3	Huang et al. 2006
	Dharwar x Sitta (RIL)	4A	127	7	Kirigwi et al. 2007
	Trident x Molineux (DH)	1B, 2D, 3D, 4D, 6A, 6D	182	18	Kuchel et al. 2007
	WL711 x PH132 (RIL)	1DL, 2DL, 3BL, 4AS, 4DL, 7AS, 7AL	100	6	Kumar et al. 2007
	Opata85 x W7984 (RIL)	1AL, 2AS, 2DS, 4BL, 6DL	110	6	Kumar et al. 2007
	Chuang35050 x Shanong483 (RIL)	1D, 2D, 3B, 6A	131	6	Li et al. 2007
	Kofa x Svevo (RIL)	2B, 3B, 7B	249	16	Maccaferri et al. 2008
	Sunco x Tasman (DH)	2B, 4D	163	4	Mares & Campbell 2001
	Ning7840 x Clark (RIL)	1AL, 1B, 2BL, 4AL, 4B, 5A, 5B, 6B, 7A, 7DL	132	5	Marza et al. 2006
	RL4452 x AC Domain (DH)	2A, 2B, 3D, 4A, 4D	182	8	McCartney et al. 2005
	SeriM82 x Babax (RIL)	6D, 7A	194	8	McIntyre et al. 2010
	Chinese Spring x SQ1 (DH)	1AS, 1BL, 2BS, 4AS, 4AL, 4BS, 4BL, 4DL, 5AL, 5BS, 5BL, 5DS, 5DL, 6BL, 7AL, 7BS, 7BL	96	24	Quarrie et al. 2005

† Gpc, grain protein concentration; Tw, test weight; Kw, kernel weight; Kd, kernel diameter; Kh, kernel hardness; Gy, grain yield.

1.6.3. Kernel characteristics

Characteristics of a single kernel include kernel weight (Kw), kernel diameter (Kd), and kernel hardness (Kh). These kernel characteristics may be measured using a single kernel characterization system (SKCS) (Perten Instruments, Springfield, IL). Kernel weight is the weight of a single kernel, expressed in mg. Kernel diameter is the diameter of a single kernel, expressed in mm. Kernel hardness is the hardness of a single kernel, expressed as an index of 0 to 100. These traits are quantitative traits that are controlled by both genetic and environmental factors.

1.6.3.1. Kernel weight

Kernel weight is one of the most important components of grain yield (Cui et al. 2011),which has a relationship with milling quality. An increase in Kw results in an increase in flour yield (Wiersma et al. 2001). Selection for increased Kw could result in an increase in grain yield (Alexander et al., 1984). Kernel weight has been reported to be both positively and negatively correlated with grain yield (Fjell et al., 1985). Single kernel weight was highly correlated with Tw (r=0.89-0.91, P<0.01) (El-Feki et al. 2013) and 1,000-kernel weight (r=0.94, P<0.001) (Tsilo et al. 2010).

Several QTL controlling Kw have been detected in different populations and environments, as summarized in Table 2. El-Feki et al. (2013) investigated a DH population (CO9406610/Platte) in four environments and found two QTL for Kw on chromosomes 1A and 7D in a single environment, two QTL on 1B and 2D repeated in two environments, and three QTL on 2B, 3B and 6A detected in three environments.

1.6.3.2. Kernel diameter

Kernel size is an important factor in wheat and is related to grain yield and quality (Tsilo et al. 2010). Kernel size and shape may influence milling and baking quality (Breseghello and Sorrells, 2007). Changes in kernel shape and size may increase flour yield of up to 5% (Marshall et al. 1984). Kernel volume and flour yield were significantly correlated (r=0.64) (Berman et al. 1996). Kernel size, reported as either kernel diameter or kernel width, is the trait most highly correlated with flour yield (Giura & Saulescu 1996).

Breseghello & Sorrells (2007) identified QTL for Kd on chromosome 1B and for both Kd and Kw on 2DS. Campbell et al. (1999) evaluated a population of 78 RIL over six environments and found that QTL for kernel width were located on chromosomes 1A, 2A, 2B, 2DL, and 3DL, and were detected in more than one environment (P<0.01). Dholakia et al. (2003) detected only one QTL for kernel width on chromosome 2DL in a population of 106 RIL. El-Feki et al. (2013) conducted experiments with 185 DH lines in four environments. The results showed three QTL for Kd on chromosomes 2D, 7A and 7D in a single environment, one QTL on 1A repeated in two environments, one QTL on 6A reproduced in three environments, and two stable QTL on 2B and 3D in all four environments. Some other studies also found QTL for kernel size in different populations and environments, and are summarized in Table 2.

1.6.3.3. Kernel hardness

Kernel hardness or texture is used as a grading factor to determine the type of wheat (Morris 2002), evaluate end product quality (Campbell et al. 1999), and classify wheat into soft and hard types (Figure 5) (Campbell et al. 1999). The major methods for determining softness and hardness are particle size index (Osborne et al. 2001), energy required for grinding a sample (Kosmolak 1978), pearling value (Chung et al. 1977) and near infrared reflectance (Manley

1995). Kernel hardness has a profound effect on milling and baking qualities of wheat (Bettge et al. 1995). The endosperm texture influences tempering requirements, flour particle size, flour density, starch damage, water absorption, milling yield and rheological properties of dough (Martin et al. 2001; J. M. Martin et al. 2007; Chen et al. 2007; Cane et al. 2004; Branlard et al. 2001).

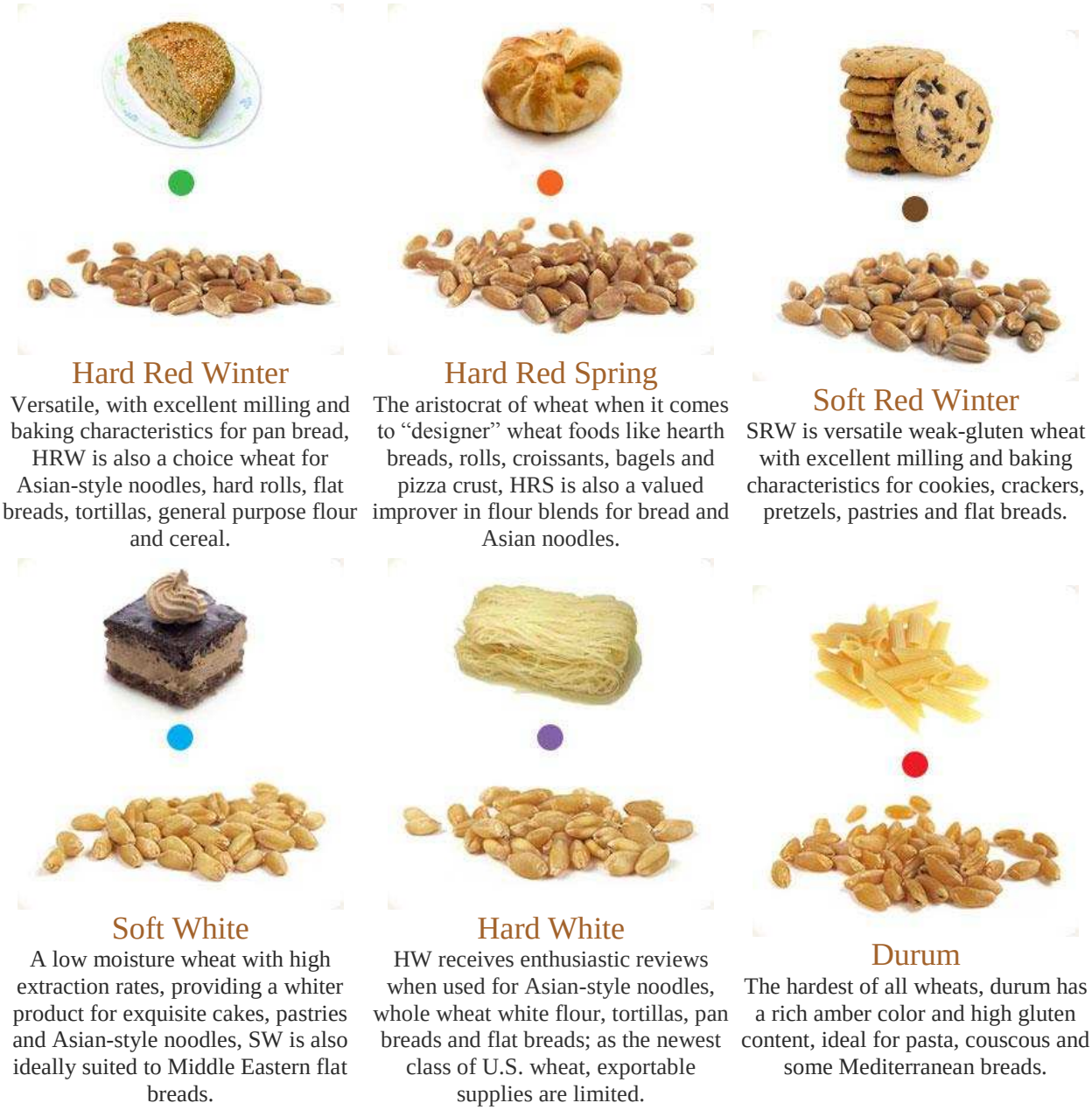


Figure 5. Wheat grain classes (http://www.uswheat.org/wheatClasses).

Kernel hardness is affected both by major genes and QTL. Major genes coding for puroindoline a (*PinA*) and puroindoline b (*PinB*) are tightly linked to the *Ha* locus on chromosome 5D (Jolly et al. 1993; Sourdille et al. 1996). El-Feki et al. (2013) conducted experiments using a population of 185 DH lines, developed from the cross CO940610/Platte, in four environments. The results showed that four QTL for Kh on chromosomes 1D, 3B, 7A and 7D were detected in a single environment, and two QTL for Kh on chromosomes 2B and 6B were found in three environments. Phenotypic variation explained by the QTL ranged from 5.7 to 16.5%. Several other QTL for Kh have been detected in different populations and environments (Table 2).

1.6.4. Grain yield

Improvement of Gy is the primary goal of all wheat breeding programs in the Great Plains of North America (Graybosch & Peterson 2010) and around the world (Wu et al. 2012). Grain yield is the biological and mathematical product of the yield components, and can be expressed as in the following equation (Chastain 2003):

$$\text{Grain yield (kg ha}^{-1}) = \frac{\#\text{ of plants}}{\text{hectare}} \text{ x } \frac{\#\text{ of tillers}}{\text{plant}} \text{ x } \frac{\#\text{ of spikes}}{\text{tiller}} \text{ x } \frac{\#\text{ of grains}}{\text{spike}} \text{ x } \frac{\text{weight (mg)}}{\text{kernel}}$$

Grain yield is a complex quantitative trait controlled by multiple genes and highly influenced by environmental conditions (Jiaqin et al. 2009). Drought stress considerably reduces Gy. El-Feki et al. (2013) reported Gy reduction of 18.7 – 21.4% in the limited soil moisture treatment compared to the fully watered treatment. Kilic & Tacettin (2010) reported that the average Gy reduction due to drought conditions was 61.4%; they suggested that reduced grain filling period, fewer spikes per square meter, lighter grains, and shorter plant cycle caused lower Gy under drought stress. Water stress during grain filling decreases sucrose and starch accumulation, thus, reducing harvested Gy (Ahmadi & Baker 2001).

Although Gy and Gpc are two major targets of most wheat breeding programs, they are inversely correlated (Daniel & Triboı 2002; Guttieri et al. 2000; Weightman et al. 2008; El-Feki et al. 2013). Therefore, the effort to improve both these traits simultaneously is challenging, particularly in semi-arid or arid regions. In order to improve Gy, it is possible to select Gy components and related traits because Gy is directly or multilaterally determined by its component traits, and indirectly influenced by other yield-related traits i.e., plant architecture (Wu et al. 2012). Wu et al. (2012) found Gy per plant was significantly correlated with number of spikes per plant (r=0.16, P<0.05), number of grains per spike (r=0.39, P<0.0001), 1000-grain weight (r=0.48, P<0.0001), total number of spikelets per spike (r=0.21, P<0.01), proportion of fertile spikelets per spike (r=0.27, P<0.005), spike length (r=0.29, P<0.005), and plant height (r=0.52, P<0.0001), but negatively correlated with number of sterile spikelets per spike (r=-0.22, P<0.01) and number of spikelets per spike (r=-0.19, P<0.05).

Several studies have reported QTL for Gy in wheat (Table 2). All 21 wheat chromosomes have been reported to be involved in controlling Gy. Five major QTL for Gy in a population of 402 DH lines were detected on chromosomes 1A, 2D, 3B, and 5A, particularly the one on 5AL, which explained 17.4% of the phenotypic variation (Cuthbert et al. 2008). Huang et al. (2003) identified QTL for Gy on chromosomes 1AL, 1BL, 2BL, 2DL, 3AS, 3BL, 4DS and 5BS in an advanced backcross population of 72 lines using 210 SSR markers. Maccaferri et al. (2008) detected one QTL for Gy on chromosome 2BL in eight environments and another QTL on chromosome 3BS over seven environments. The average phenotypic variation explained by the 2B QTL was 21.5% and for the 3B QTL was 13.8%. These QTL overlapped extensively with plant height QTL. Bennett et al. (2012) found nine loci on chromosomes 3A, 3BS, 3BL, 3D, 4A, 4D, 5B, 7A.1, 7A.2 that were associated with Gy. Two QTL for Gy on 3BS and 3BL had a large

effect with phenotypic variation of up to 22%. These two QTL co-located with QTL for canopy temperature. J. Zhang et al. (2014) evaluated a mapping population of 159 $F_{8:10}$ RIL over six location-year environments and identified 17 QTL for Gy located on 14 chromosomal regions 1A.1, 1B.1, 2B.1, 2B.2, 2D, 3B.1, 3B.2, 4B, 5A.1, 5B.2, 6B.2, 7A.4, 7A.5, and 7B.1. The phenotypic variation explained by the QTL ranged from 6 to 22%.

1.7. Genome-wide association study

Association mapping (AM), also called LD mapping, refers to the analysis of statistical associations between genotypes and the phenotypes of the same individuals (Rafalski 2010). While QTL mapping typically uses a bi-parental mapping population, the progeny of parents having contrasting trait(s) of interest, AM utilizes a diverse collection of individuals derived from wild populations, germplasm collections, or subsets of breeding germplasm. A bi-parental mapping population requires a much longer time to be developed compared to an AM population. In the former type of population, only two alleles at a locus can be evaluated, while the latter type allows evaluation of a broader range of alleles. An AM population has much higher mapping resolution than the bi-parental one because of the limited numbers of recombination events in a typical QTL population. Therefore, increased mapping resolution, reduced research time, and broader allele number (Yu & Buckler 2006) are three advantages of AM.

Two AM approaches are in general use: (1) candidate-gene association mapping, and (2) whole genome scan or GWAS. Candidate-gene association mapping relates polymorphism in selected candidate genes controlling phenotypic variation for specific traits, while GWAS surveys genetic variation in the whole genome to identify signals of association for various complex traits (Risch & Merikangas 1996).

Performing an AM study consists of the following steps (Figure 6): (1) selection of a group of individuals with wide coverage of genetic diversity; (2) measuring the phenotypic characteristics; (3) genotyping the mapping population; (4) quantification of the extent of LD for a chromosome and/or genome; (5) assessment of the population structure and kinship; and (6) identification of association of phenotypic and genotypic data (Abdurakhmonov & Abdukarimov 2008).

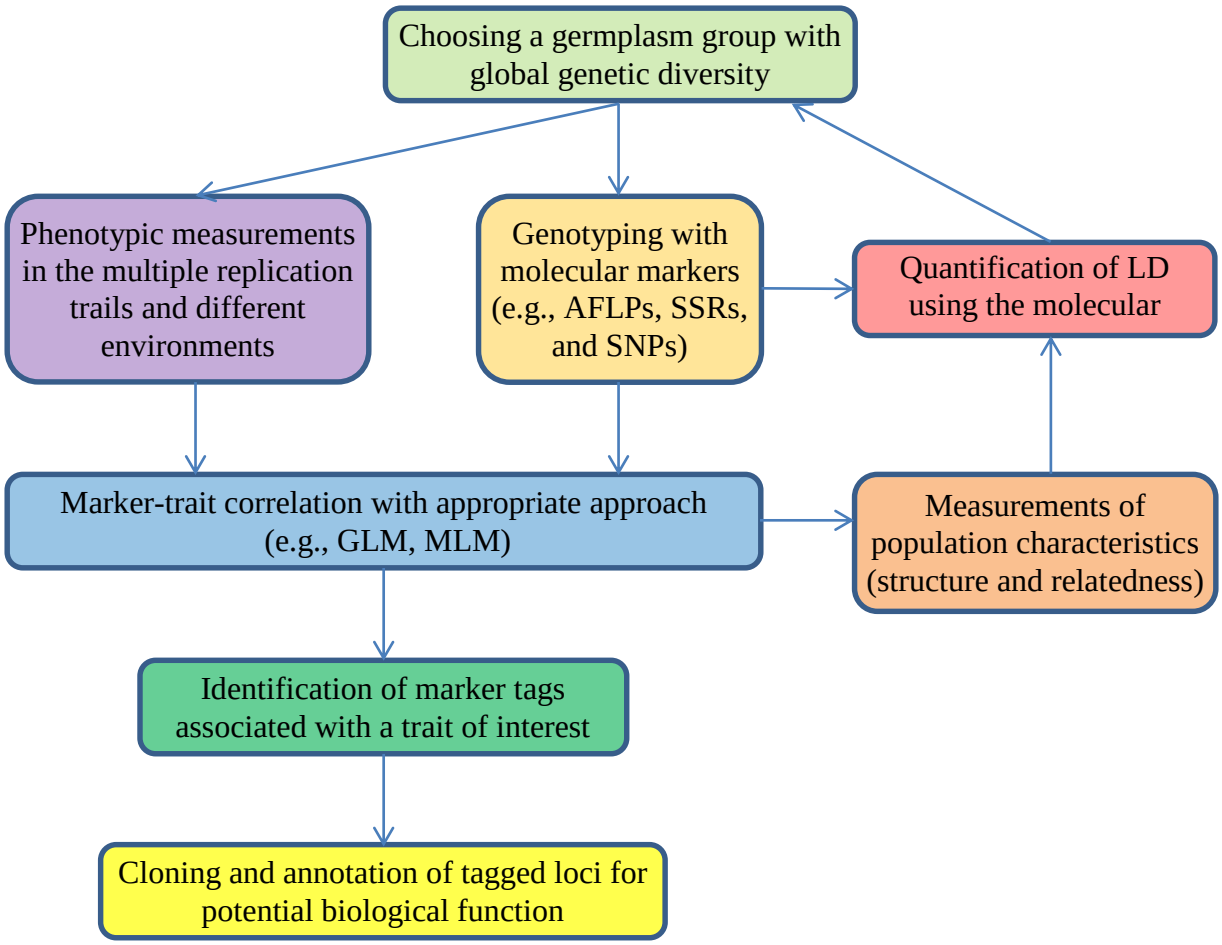


Figure 6. The steps for performing AM and identifying candidate genes (Abdurakhmonov & Abdukarimov 2008).

However, AM requires a large number of molecular markers and powerful statistical methods. Advances in the development of high throughput genotyping technology (Ansorge 2009) and statistical methodology have enabled AM analysis of complex traits and the subsequent identification of causal genes (Rafalski 2010). Genomic locations of marker-trait associations (MTA) detected by the AM analysis are necessarily inferred from a consensus genetic map and/or physical map for the crop investigated. Drawbacks of the AM are the high risk of type I error (false positives) and the high sampling variance of rare alleles. False discoveries are a major concern (Abiola et al. 2003). If the rare allelic effect is not very large, the rare alleles cannot be detected with good confidence (Rafalski 2010; Visscher 2008).

The AM method was first applied to plant research by Thornsberry et al. (2001), who studied maize flowering time. AM has been used successfully to detect QTL in wheat for end-use quality traits (Breseghello & Sorrells 2006; Plessis et al. 2013; Zheng et al. 2009), grain yield and yield components (Edae et al. 2014; Maccaferri et al. 2011; Neumann et al. 2011; Dodig et al. 2012; Sukumaran et al. 2015), disease resistance (Adhikari et al. 2011; Ghavami et al. 2011; Maccaferri et al. 2010; Crossa et al. 2007; Yu et al. 2011; Peng et al. 2009; Maccaferri et al. 2015), and root traits (Canè et al. 2014).

1.8. Marker assisted selection

With the availability of more sophisticated tools, the art of plant breeding has expanded to include technology of molecular plant breeding (Xu 2010). The advent of molecular technology has allowed development of QTL mapping and its follow-up, marker-based selection for trait(s) of interest. QTL mapping or bulk segregant analysis is a necessary precursor to marker-assisted selection, also called “marker-assisted breeding” or “marker-aided selection” (Collard et al. 2005). Marker-assisted selection is a breeding method in which a phenotype is selected based on

the genotype of a marker (Collard et al. 2005) by selecting desired allele(s). With recent progress in genomic research, the MAS approach enables the selection of plants with the targeted traits as early as the F_2.

There are several advantages of MAS compared to conventional plant breeding (Collard et al. 2005): (1) time saved by replacing complex field trials with molecular tests; (2) elimination of unreliable phenotypic evaluation associated with field trials due to environmental effects; (3) selection of genotypes at seedling stage; (4) gene pyramiding or combining multiple genes simultaneously; (5) avoiding the transfer of undesirable genes via 'linkage drag'; (6) selecting for traits with low heritability; (7) testing for specific traits where phenotypic evaluation is not feasible. However, inheritance of the trait, method of phenotypic evaluation, field/greenhouse and labor costs, and the cost of resources are factors that influence the cost of utilizing MAS (Collard et al. 2005). In some cases, phenotypic screening is cheaper than MAS (Bohn et al. 2001; Dreher et al. 2003). Although markers may be cheap to genotype, a large initial cost is required in their development (Langridge et al. 2001).

MAS has been successfully applied to practical breeding in wheat. Vishwakarma et al. (2014) used marker-assisted backcrossing (MABC) to improve Gpc in the wheat cultivar HUW468 by selecting for the appropriate allele at marker *Xucw108*, which is linked to the high-value allele of the *Gpc-B1* gene. The result showed significantly higher Gpc in improved lines by MABC, and no yield penalty. Kumar et al. (2011) improved Gpc of seven MAS-derived progenies by 14.8 to 17.9% compared to their recipient parents using MABC. The authors suggested that combining MAS and phenotypic selection is useful to improve wheat genotypes for high Gpc without yield loss.

CHAPTER 2: VALIDATION OF QUANTITATIVE TRAIT LOCI FOR GRAIN QUALITY TRAITS IN WINTER WHEAT USING A CO940610/PLATTE RECOMBINANT INBRED LINE POPULATION

SUMMARY

Wheat is among the three most important cereal crops that provide the global human food supply. Drought has serious effects on the productivity and grain quality of wheat. Grain quality traits affect the end uses of wheat that include grain protein concentration (Gpc), grain ash concentration (Gac), test weight (Tw), kernel weight (Kw), kernel diameter (Kd), and kernel hardness (Kh).

A previous study reported that three chromosomal regions (1B, 6B, and 7B) were associated with QTL grain quality traits in a doubled haploid (DH) population derived from the cross CO940610/Platte. To validate those QTL, a recombinant inbred line (RIL) population was developed using the same parental lines (CO940610 and Platte) as in the previous DH mapping population.

The objectives of this study were to (1) evaluate grain yield and grain quality traits in the mapping population of 186 F_5-derived RIL grown in Akron (rainfed) and Greeley (fully-irrigated) environments in the 2009/10 growing season, (2) genotype the RIL population for 18 molecular markers, and (3) conduct QTL analysis to find associations for the evaluated traits.

Near-infrared spectroscopy (NIRS) was used to determine Gpc, Gac and Tw. The Single Kernel Characterization System (SKCS) was employed to measure Kw, Kd, and Kh. JoinMap 4.0 was used to construct linkage maps from the molecular marker data. Correlation coefficients were

calculated for each pair of traits. Heritability estimates were calculated for each trait based on variance components on a plot mean basis. Marker-trait associations (MTA) were detected by single-factor analysis of variance (ANOVA).

Grain yield was negatively correlated with Gpc and Gac, but positively correlated with Tw. Grain protein concentration was also positively correlated with Gac, but negatively correlated with Tw. Test weight, Kw and Kd were positively correlated with each other, with a particularly tight association between Kw and Kd. Grain yield and Gpc had low (<0.50) heritability estimates, while Gac, Tw, Kw, Kd and Kh had moderate (0.50-0.75) to high (>0.75) heritability estimates, and stability of heritability across environments.

The linkage group constructed on 6B for the RIL and DH populations was 9.5 and 12.2 cM, respectively, but marker orders are slightly different. The linkage group constructed on 7B for the RIL and DH populations was 39.3 and 19.5 cM, respectively, but marker orders are exactly the same. There were 64 significant MTA detected for nine traits in either one or two environments. Some MTA detected in the RIL population coincided with those detected in the DH population, while other MTA were located close to, far away from, or contradictory to QTL detected in the DH population. Therefore, results of this study only partially confirmed results of the DH population. Tightly linked markers for grain quality traits located around *Xwmc182a* on 6B and *Xwmc182b* on 7B may be useful for selection.

2.0. INTRODUCTION

Wheat (*Triticum aestivum* L.) is one of the most highly produced cereal crops after maize (*Zea mays* L.) and rice (*Oryza sativa* L.), and is important in meeting global food demands. The 2013

global production was 1,018 million tons (MT) for maize, 741 MT for rice, and 716 MT for wheat (FAO 2013). Wheat supplies about 20% of the total food calories consumed worldwide, and is the major source of dietary proteins (Wen et al. 2012). The global demand for wheat is predicted to increase as the world's population increases to an estimated 9.6 billion by 2050 (UN 2013).

Wheat grain can be used as livestock feed or various types of human food. It is used primarily to produce flour for bread, but is used widely in the production of many other baked goods, cooked products, and alcoholic beverages. These end-use products are determined by several wheat grain characteristics. Wheat breeding programs put considerable effort into improving grain quality in addition to increasing yield. Grain quality is characterized by several quality parameters such as grain protein concentration (Gpc), grain protein composition, grain ash concentration (Gac), flour color, test weight (Tw), polyphenol oxidase activity, dough-mixing properties, and single kernel characteristics.

Wheat productivity is mainly limited by biotic and abiotic stresses. Abiotic factors are considered the main source (70%) of yield reductions (Boyer 1982). Among the abiotic stresses, drought is the most limiting factor, reducing productivity of wheat and other crops in many parts of the world (Ahuja et al. 2010; Rampino et al. 2006; Mohammadi et al. 2015). Drought stress may occur from early in the growing season to terminally at grain filling and maturation, particularly in arid and semi-arid regions of the world. The climates of these regions usually have inadequate precipitation, a shortage of irrigation water, and high crop evapotranspiration demand.

Drought conditions also influence grain quality traits of wheat. Under drought stress Gpc usually increases, and is negatively correlated with grain yield (Daniel & Triboı 2002; El-Feki et al.

2013; Guttieri et al. 2000; X. Sun et al. 2010; Weightman et al. 2008). Water stress also changes grain protein composition in winter wheat. A study by Saint Pierre et al. (2008) found that under water stress, the percentage of soluble protein decreased, but polymeric protein percenatage was not affected. (Ahmed & Fayyaz-ul-Hassan 2015) found that under drought conditions, proline content increased ten-fold and had a synergistic relationship with grain protein and grain ash. Grain ash content was significantly higher when plants were irrigated fewer times (Seleiman et al. 2011).

With the advent of molecular markers, quantitative trait loci (QTL) mapping became a key approach for understanding the genetic architecture of complex traits in plants (Holland 2007). QTL analysis based on molecular marker linkage maps has been used to detect genomic regions that control wheat grain quality (Breseghello et al. 2005; El-Feki et al. 2013; Kuchel et al. 2006; Li et al. 2009; X. Sun et al. 2010). Recombinant inbred line (RIL) populations have been widely used for QTL mapping in plants, including wheat (Breseghello et al. 2005; Li et al. 2009; Kuchel et al. 2006; X. Sun et al. 2010). However, it takes a long time (three years or more) to produce a RIL population because usually six to eight generations are required, while a doubled haploid (DH) population requires only one to two years. RIL and DH populations produce homozygous or nearly homozygous or 'true breeding' lines. Therefore, seed from individual RIL or DH lines can be grown in multiple environments and shared among different laboratories (Collard et al. 2005).

A previous QTL mapping study conducted in four environments found clusters of QTL for grain quality characteristics in regions of chromosomes 6B and 7B in a population of 185 CO940610/Platte DH lines (El-Feki 2010; El-Feki et al. 2013). A cluster of seven QTL on 6B included two for mixograph peak time, three for Kh, one for Gpc, and one for Tw. Another

cluster of 18 QTL on 7B consisted of three for mixograph peak time, one for mixograph peak height, three for mixograph peak width, two for mixograph right slope, one for Kw, three for Gpc, one for Gac, and four for flour yellowness (b*). Before the QTL can be employed in marker-assisted selection (MAS) programs for improved grain quality traits, their effects should be validated in another population. To validate those QTL, a RIL population was developed using the same parental lines (CO940610 and Platte) as in the previous DH mapping population.

Therefore, The objectives of this study were to (1) evaluate grain yield and grain quality traits in the mapping population of 186 F_5-derived RIL grown in Akron (rainfed) and Greeley (fully-irrigated) environments in the 2009/10 growing season, (2) genotype the RIL population for 18 molecular markers, and (3) conduct QTL analysis to find associations for the evaluated traits.

2.1. MATERIALS AND METHODS

2.1.1. Mapping population

A RIL wheat population was developed by crossing CO940610 as the female parent with 'Platte' (El-Feki et al. 2015). Platte is a hard white winter cultivar developed by HybriTech Seed International, Inc. (a former division of Monsanto Co., St. Louis, MO), with pedigree Tesia 79/Chat'S'//'Abilene'. It is known for its adaptation to irrigated production systems and excellent bread making quality, but is considered to have poor tolerance to pre-harvest sprouting. CO940610 is a hard white winter experimental line developed by the CSU Wheat Breeding Program from the cross KS87H22/MW09. The first parent of CO940610 has a genetic composition that is approximately one-half 'TAM 105' (Porter et al. 1980), which explains at least part of its excellent dryland yield (personal communication, Dr. Scott Haley, Colorado State

University). Clark's Cream constitutes approximately half of the MW09 parent (Anderson et al. 1993), and thus, it is possible that CO940610 has pre-harvest sprouting tolerance from that cultivar. However, this line has poor bread making quality, based on observations in CSU's Wheat Quality Laboratory.

Parental lines were evaluated as part of the multi-state Wheat Coordinated Agriculture Project (CAP) funded by USDA-CSREES. Allelic variation between Platte and CO940610 at selected major genes is presented in Table 3.

The F_1 generation was produced in the CSU greenhouses. The subsequent generations were allowed to self-pollinate (protected from cross-pollination by pollination bags) for four generations, which were advanced by single seed descent. Seed harvested from single plant in the fifth generation ($F_{5:6}$) was grown in the F_6 at the Agricultural Research Development and Education Center (ARDEC), Fort Collins, Colorado to increase seeds for field trials.

2.1.2. Experimental design and trial management

The population for field trials included 186 RIL, the two parents, and four check cultivars. The checks were 'Hatcher' (Haley et al. 2005), 'Ripper' (Haley et al. 2007), 'RonL' (J. T. Martin et al. 2007), and 'Snowmass' (Haley et al. 2011).

The experiments were conducted in the 2009/10 growing season at the USDA-Agricultural Research Service (ARS) Limited Irrigation Research Farm (LIRF) in Greeley, Colorado and at the USDA-ARS Central Great Plains Research Station (CGPRS) in Akron, Colorado. The trial at CGPRS was grown under rainfed conditions, while the trial at LIRF was grown under fully irrigated conditions as shown in Table 36 and 37. The water applied at LIRF was via a drip system, which provided uniform water application.

Table 3. Allelic or phenotypic variation for selected major genes or traits of CO940610 and Platte winter wheat[†]

Characteristic	Locus or trait	Allelic variation	
		CO940610	**Platte**
Grain texture	*PinA*	+	+
	PinB	-	-
Gluten strength	*Glu-A1*	c (null)	b (2*)
	Glu-B1	b (7+8)	e (20x+20y)
	Glu-D1	a (2+12)	d (5+10)
	Glu-A3	c	c
	Glu-B3	h	g
	Glu-D3	b	c
Waxy	*Wx-D1*	+	+
	Wx-A1	+	+
	Wx-B1	+	+
Grain protein	*Gpc-B1*	-	-
Rust resistance	*Lr37/Yr17/ Sr38*	-	-
Leaf rust	*Lr21*	-	-
Hessian fly resistance	*H9*	-	-
Aluminum tolerance	Al tolerance	T	S
Vernalization requirement	Heading	Late	Late
	VRN-A1	*vrn-A1*	*vrn-A1*
	VRN-B1	*vrn-B1*	*vrn-B1*
	VRN-D1	*vrn-D1*	*vrn-D1*
Height	*Rht-B1*	*Rht-B1b*	*Rht-B1b*
	Rht-D1	*Rht-D1a*	*Rht-D1a*
Rye translocation	*1RS*	*non-1RS*	*non-1RS*

[†] For most of these loci and traits, the source of the information is the Wheat CAP website (http://maswheat.ucdavis.edu/wheatcap.htm). Alleles for *Glu-A3*, *Glu-B3*, and *Glu-D3* were determined by Scott Reid, CSU Department of Soil and Crop Sciences, based on protein and polymerase chain reaction (PCR) markers in comparison with a published report (Appelbee et al. 2009).

Notes:

Grain texture: A negative in either *PinA* or *PinB* results in hard texture. Both positives result in soft.

Gluten strength: *Glu-D1a* (2+12) is associated with weak gluten; *Glu-D1d* (5+10) is associated with strong gluten.

Waxy endosperm: One negative *Wx* allele is sufficient for a partial Waxy phenotype.

Grain protein: Positive indicates the presence of the *T. diccocoides* allele for high grain protein concentration.

Lr37/Yr17/Sr38: Positive indicates presence of the 2NS/2AS translocation carrying this set of three resistance genes.

Hessian fly resistance: Positive indicates presence of the *H9* allele for resistance.

Leaf rust resistance: Positive indicates presence of the *Lr21* allele for resistance.

Aluminum tolerance: T indicates the presence of the Atlas tolerant allele; S indicates a susceptible genotype.

VRN: One dominant allele at any of the three loci is sufficient for a dominant spring growth habit.

Dwarfing genes *Rht-B1* and *Rht-D1*: The a allele is for tall, and the b allele is for semi-dwarf.

Rye translocation: The presence of the 1RS translocation on 1AL or 1BL may enhance drought tolerance and pest resistance.

The trial was planted at CGPRS on September 30, 2009, and was harvested on July 13, 2010.

The soil type at LIRF was Nunn clay loam, 0 to 1 percent slopes, while it was Rago silt loam (map unit 58) at CGPRS (http://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm).

Each trial had two replicates of 192 two-row plots. Entries were arranged in a randomized complete block Latinized row-column design, created with CycDesign 3.0 (Cyc software 2009).

Plot size differed in each location. At CGPRS, each plot was 4.27 m long and 0.46 m wide, while at LIRF, plot size was 3.05 m long and 0.46 m wide. Seeding density was the same in both locations, approximately 1,700,000 seed ha^{-1}.

Fertilization and weed control were typical of practices for winter wheat in the area. Disease or insect pest control was not needed. A Hege combine (Wintersteiger, Salt Lake City, Utah) was used for harvesting after plants reached maturity, which occurred in July 2010.

2.1.3. Phenotypic evaluation

2.1.3.1. Grain yield, Spike length, and Plant height

The harvested grain samples were dried for about three days at approximately 40^{o}C, weighted, and Gy was converted to kg ha^{-1}. Spike length (Sl) excluding awns, was measured in cm, and averaged from five random spikes per plot. Approximately two weeks before harvest, plant height (Ph), expressed in cm, was measured from the ground to the tip of the spike, excluding awns.

2.1.3.2. Grain protein concentration, grain ash concentration, and test weight

Three traits, Gpc, Gac and Tw, were measured by the Near Infrared Reflectance System (NIRs) in the CSU Wheat Quality Laboratory. A Foss-Tecator NIR systems Model 6500 instrument (Foss North America, Eden Prairie, MN) was used to measure these traits in reflectance mode with a wavelength of 450–2498 nm.

Wheat kernels were cleaned before measurement. A sample of approximately 100 g was added into a transport module in a rectangular quartz cup (NIRSystems, part 0IH–0379). The estimates

of the traits were obtained with calibrations from the manufacturer for Gpc and Gac, and a calibration developed in the CSU Wheat Quality Laboratory for Tw. The output measured by the instrument was recorded on a computer. The units of traits were converted to g kg^{-1} for Gpc and Gac, and kg hL^{-1} for Tw.

2.1.3.3. Single kernel characteristics

Three single kernel traits measured included kernel weight (Kw), kernel diameter (Kd) and kernel hardness (Kh). These traits were measured by the Single Kernel Characterization System (SKCS) in the CSU Wheat Quality Laboratory. Wheat kernels were carefully cleaned. Broken kernels, weed seeds, metal staples, and any other foreign materials were manually removed. The cleaned wheat kernels were placed in the access hopper of the SKCS instrument (model 4100, Perten Instruments, Springfield, IL), then the hopper was placed in the SKCS instrument for kernel measurement. The instrument analyzed 100 kernels individually, and measured Kw, Kd and Kh. Kernel weight was measured by a load cell, and reported in mg. Kernel diameter was measured by electrical current, and expressed in mm. Kernel hardness was analyzed by pressure force, and expressed as an index scale of 0 (extremely soft) to 100 (extremely hard). These results were recorded on a computer.

2.1.4. Statistical analysis

Best Linear Unbiased Predictions (BLUPs)

The BLUPs were calculated to predict the mean performance of each entry within each location. Entry was considered to be a random effect. The estimation was performed with PROC MIXED of SAS 9.3 (SAS Institute Inc., Cary, NC, 2010), including a spatial adjustment as explained below.

Spatial adjustment

Spatial variation in the field sites was examined to determine whether spatial adjustment was beneficial. As a first step, the phenotypic data for each trait were inspected with a "heat map", which was produced in Excel to determine whether any data values were unreasonably large or small and to visualize spatial trends. Six different spatial models (SAS statements) were applied to adjust the experimental field spatial variation. These included non-adjusted, spherical, exponential, regular power, anisotropic power, and Matérn spatial models. The model with the lowest Akaike's information criterion (AIC) was considered the most appropriate one.

Population distribution

The phenotypic data, including original data, LSMEANs, and BLUPs were evaluated for normality of frequency distributions based on visual inspection and the Shapiro-Wilk test at $\alpha = 0.05$. Random distribution of residuals was also examined visually. These evaluations were implemented in SAS, JMP, and Excel. Based on the frequency distributions, the occurrence of transgressive segregation was determined.

Pearson correlation

The BLUP estimates were used to calculate the Pearson phenotypic correlation coefficient for each pair of traits in each location by using PROC CORR of SAS.

Heritability estimates

The broad sense heritability (H^2) was estimated with SAS PROC GLM and Excel from the analysis of variance of progeny-mean data using the original data and the following formulae:

For a single trait evaluated in a single environment (single location) with two replications ($r = 2$) and 186 RIL in each replication ($t = 186$),

Source	Degree of freedom	Expected mean square
Reps	r-1 = 1	$\sigma^2 + t\sigma^2_R$
Entries	t-1 = 185	$\sigma^2 + r\sigma^2_G$ = MSG (mean square genotype)
Error	(r-1)(t-1) = 185	σ^2 = MSE (mean square error)

$\sigma^2_G = [(\sigma^2 + r\sigma^2_G) - \sigma^2]/r = (MSG - MSE)/r$

$\sigma^2_P = \sigma^2/r + \sigma^2_G = (\sigma^2 + r\sigma^2_G)/r = MSG/r$

$h^2 = \sigma^2_G/\sigma^2_P = \sigma^2_G/(\sigma^2_G + \sigma^2/r) = [(MSG - MSE)/r]/(MSG/r) = 1 - (MSE/MSG)$, where σ^2 is the error mean square, σ^2_G is the entries mean square, and r is number of replications.

For a single trait evaluated in two environments (n = 2) or two locations (Akron and Greeley), two replications (r = 2), and 186 RIL (t = 186),

Source	Degree of freedom	Expected mean square	
Env	n-1 = 1	$\sigma^2 + r\sigma^2_{GxE} + t\sigma^2_{R(E)} + rt\sigma^2_E$	
Rep (Env)	n(r-1) = 2	$\sigma^2 + r\sigma^2_{R(E)}$	
Entries	t-1 = 185	$\sigma^2 + r\sigma^2_{GxE} + rn\sigma^2_G$	= MSG (mean square genotype)
Env x Entries	(n-1)(t-1) = 185	$\sigma^2 + r\sigma^2_{GxE}$	= MSI (mean square interaction)
Error	n(r-1)(t-1) = 370	σ^2	= MSE (mean square error)

When the Env x Entries term was significant ($P<0.05$):

$\sigma^2_G = [(\sigma^2 + r\sigma^2_{GxE} + rn\sigma^2_G) - (\sigma^2 + r\sigma^2_{GE})]/rn = (MSG - MSI)/rn$

$\sigma^2_{GxE} = [(\sigma^2 + r\sigma^2_{GxE}) - \sigma^2]/r = (MSI - MSE)/r$

$\sigma^2_P = \sigma^2/rn + \sigma^2_{GxE}/n + \sigma^2_G = (\sigma^2 + r\sigma^2_{GxE} + rn\sigma^2_G)/rn = MSG/rn$

$h^2 = \sigma^2_G/\sigma^2_P = \sigma^2_G/[(\sigma^2 + r\sigma^2_{GxE} + rn\sigma^2_G)/rn] = [(MSG - MSI)/rn]/(MSG/rn) = 1- (MSI/MSE)$

When the Env x Entries term was not significant (P>0.05):

$\sigma^2_G = [(\sigma^2 + r\sigma^2_{GxE} + rn\sigma^2_G) - \sigma^2]/rn = (MSG - MSE)/rn$

$\sigma^2_{GxE} = [(\sigma^2 + r\sigma^2_{GxE}) - \sigma^2]/r$

$\sigma^2_P = \sigma^2/rn + r\sigma^2_{GxE}/r + \sigma^2_G = (\sigma^2 + r\sigma^2_{GxE} + rn\sigma^2_G)/rn = MSG/rn$

$h^2 = \sigma^2_G/\sigma^2_P = \sigma^2_G/[(\sigma^2 + r\sigma^2_{GxE} + rn\sigma^2_G)/rn] = [(MSG - MSE)/rn]/(MSG/rn) = 1- (MSE/MSG)$

Confidence interval (C.I.) for heritability

Exact confidence intervals for heritability estimates on a progeny mean basis were computed according to Knapp et al. (1985). The 90% confidence intervals (α= 0.05 for each of the upper or lower confidence limits) were calculated.

2.1.5. DNA extraction

Seeds of 186 RIL and two parental lines, CO940610 and Platte, were sown separately in trays in the CSU greenhouse. Each tray cell contained 10 seeds of each RIL or parental line. Two weeks later, when seedlings were at the 1- to 2-leaf stage, a bulked sample of approximately 400 to 500 mg leaf tissue was collected from seedlings of each cell and placed into a 2.0 mL tube. The tubes were immediately placed on ice, and then stored in an -80 ^oC freezer. Two tubes of tissue were collected for each RIL or parent, one for immediate DNA extraction and one for backup.

For DNA extraction, leaf tissue in open tubes was freeze-dried for 48 hours using a freeze drier (Freeze Dry System, Labconco ® 7522900). Freeze drier conditions were temperature of about -50 ^oC and vacuum of -80 x 10^{-3} kPa or lower, causing water to vaporize from the leaf tissues.

After freeze-drying, 2 stainless steel balls were added to each tube of leaf tissue and the tube was tightly capped for grinding. The tubes were put into a shaker (Mini-BeadBeater™, BioSpec Products, Inc.) to grind for 30 to 90 seconds to achieve a fine powder consistency.

DNA was extracted, purified, and diluted using a protocol adapted from that described in Riede & Anderson (1996). A 500 µL quantity of hot extraction buffer [0.5 M NaCl, 0.1 M Tris pH 8.0, 0.05 M ethylenediaminetetra-acetic acid (EDTA), 8.4 g L^{-1} sodium dodecyl sulfate (SDS), 3.8 g L^{-1} sodium bisulfate] was added to each tissue sample. Tubes were vortexed to suspend tissue, and then incubated in a 65ºC water bath for 60 minutes, mixing by inversion every 10 minutes. A quantity of 500 µL 24:1 chloroform:isoamyl alcohol was added, mixed to form an emulsion, then centrifuged at 14,000 rpm for 5 minutes. The aqueous phase was removed to a new 1.5 mL tube and two volumes of 100% EtOH were added. After 15 minutes at room temperature, the tubes were inverted to mix, and left for an additional 15 minutes. After centrifuging at 12,000 rpm for 5 min, the supernatant was decanted. The remaining pellet was washed with 1 mL 70% EtOH, following by centrifuging and decanting the supernatant. The pellet was air dried, then dissolved in 400 µL TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0) and incubated in a 65°C water bath. RNase A was added to a final concentration of 0.1 mg/mL and incubated at 37°C for 30 minutes. One-tenth volume of 3 M sodium acetate and 2 volumes 100% EtOH were added to precipitate DNA. The tubes were placed in a -20°C freezer for about 12 hours (overnight), and then centrifuged at 12,000 rpm for 5 minutes. The supernatant was decanted and the pellet washed with 1 mL 70% EtOH, followed by another round of centrifuging and decanting. The DNA pellet was air-dried and dissolved in 50-200 µL TE. DNA concentrations were quantitated by measuring absorbance at 260 nm on a Nanodrop ND1000 Spectrophotometer (Thermo Scientific, Wilmington, DE), and the DNA quality was evaluated on a 1.0% agarose gel containing one lane

of Lamba DNA/HindIII and stained with ethidium bromide. DNA was diluted to 75 ng/µL in TLE buffer (1.0 mM tris, 0.1 mM EDTA, pH 8.0) for genotyping.

2.1.6. Molecular marker evaluation

Parental DNA was screened for polymorphisms at simple sequence repeat (SSR) loci that had previously been mapped to chromosomes 6B and 7B and at the *Glu-B1* locus on chromosome 1B. Sixteen selected polymorphic markers were evaluated in the whole population of 186 RIL and two parental lines. *Bx7-MAR* is at *Glu-B1* locus, *Xwmc182a, Barc198, Xwmc397, and Barc136* are on chromosome 6B, and *Xgwm569, Xgwm606, Xgwm76, Xwmc182b, Xgwm573, Xgwm46, Xgwm333,* and *Barc278* are on chromosome 7B. These are known from the previous research in the CO940610/Platte DH population (El-Feki 2010; El-Feki et al. 2013). Three additional markers, *Xwmc426, Xwmc17a,* and *Xwmc364* are obtained from other publications (Table 4). Primer sequences for these markers were obtained from Grain Genes 2.0 (http://wheat.pw.usda.gov/GG2/index.shtml), as shown in the Table 4.

2.1.7. Linkage map construction

Data for all marker types for 186 RIL were compiled and tested for goodness of fit to the expected Mendelian segregation ratio of 46.875% Platte or CO940610 allele and 6.25% heterozygotes. The data were analyzed with Join Map 4 (Van Ooijen, 2006) using the regression mapping option to construct linkage groups and the Haldane mapping function to calculate the centiMorgan (cM) distances.

Table 4. Markers and primer sequences used in this study

Marker	Chr.	PCR primer sequence	Reference
Bx7-MAR	1B	CCTCAGCATGCAAACATGCAGC CTGAAACCTTTGGCCAGTCATGTC	Butow et al. 2004
Barc136	6B	GCG AGC TCA CTG CAC ACT TAC CC GCA ACG CAC CTT GAT AAT C	Song et al. 2005
Xwmc397	6B	AGTCGTGCACCTCCATTTTG CATTGGACATCGGAGACCTG	Somers et al. 2004
Barc198	6B	CGCTGAAAAGAAGTGCCGCATTATGA CGCTGCCTTTTCTGGATTGCTTGTCA	Somers et al. 2004
Xwmc182a	6B	GTATCTCACGAGCATAACACAA GAAAGTGTATGGATCATTAGGC	Somers et al. 2004
Xgwm569	7B	GGAAACTTATTGATTGAAAT TCAATTTTGACAGAAGAATT	Roder et al. 1998; Somers et al. 2004
Xwmc606	7B	CCGATGAACAGACTCGACAAGG GGCTTCGGCCAGTAGTACAGGA	Somers et al. 2004
Xwmc76	7B	CTTCAGAGCCTCTTTCTCTACA CTGCTTCACTTGCTGATCTTTG	Somers et al. 2004
Xwmc182b	7B	GTATCTCACGAGCATAACACAA GAAAGTGTATGGATCATTAGGC	Somers et al. 2004
Xwmc426	7B	GACGATCGTTTCTCCTACTTTA ACTACACAAATGACTGCTGCTA	Somers et al. 2004
Xgwm573	7B	AAGAGATAACATGCAAGAAA TTCAAATATGTGGGAACTAC	Paillard et al. 2003; Roder et al. 1998; Somers et al. 2004
Xwmc17a	7B	ACCTGCAAGAAATTAGGAACTC CTAGTGTTTCAAATATGTCGGA	Somers et al. 2004
Xgwm46	7B	GCACGTGAATGGATTGGAC TGACCCAATAGTGGTGGTCA	Somers et al. 2004
Xwmc364	7B	ATCACAATGCTGGCCCTAAAAC CAGTGCCAAAATGTCGAAAGTC	Somers et al. 2004
Xgwm333	7B	GCCCGGTCATGTAAAACG TTTCAGTTTGCGTTAAGCTTTG	Roder et al. 1998; Somers et al. 2004
Barc278	7B	GCATGCACTACGCTCAGAATAAAC TAAAAGGCCCGTCAACATACAAGTA	Song et al. 2005

2.1.8. QTL analysis

For each trait, each marker was analyzed for significance in single factor analysis of variance by using PROC GLM of SAS 9.3, with trait as the dependent variable and marker as the independent variable. A probability level of 0.05 was used as the significance threshold for marker-trait association (MTA), which interpreted as indication of the presence of a putative QTL.

Percent phenotypic variance explained by MTA *(%R^2)* was obtained by multiplying the R^2 values (coefficient of determination) provided in the SAS PROC GLM results by 100.

The average additive effect was the difference of mean trait values for two allelic classes, which was calculated as (A – B)/2 by using Excel 2010, where, A and B were mean trait values of the Platte and CO940610 allelic classes, respectively.

2.2. RESULTS

2.2.1. Trait distribution and means

Population distribution

Ten of 20 trait-environment combinations conformed to expectations for a normally distributed population according to the Shapiro-Wilk test at $P>0.05$ (Figure 13). For the trait-environments that deviated, the distributions were approximately normal based on a visual evaluation and examination of residuals. Transgressive segregation was apparent in all cases, indicating the presence of favorable alleles in both parents, CO940610 and Platte (Figure 13).

Trait means

Grain yield was significantly (P<0.01) lower in the rainfed condition in Akron compared to the fully irrigated condition in Greeley. Grain yield in Akron was 3245.3 kg ha^{-1}, a reduction of about 28% compared to Gy in Greeley (4506.1 kg ha^{-1}) (Table 5). Spike length and Ph in Akron also were significantly (P<0.01) less than those in Greeley.

Test weight, Kw, and Kd in Akron were significantly (P<0.01) lower than those in Greeley. Test weight was slightly lower, about 3 kg hL^{-1} (about 4%), while Kw and Kd were reduced about 30% and 14%, respectively (Table 5).

In contrast, Gpc and Kh in Akron were significantly (P<0.01) higher compared to Greeley. An increase of about 8 g kg^{-1} and 2 SKCS hardness units were recorded for Gpc and Kh, respectively (Table 5). There was no significant difference detected for Gac between Akron and Greeley.

Table 5. Means, standard errors (SE) and ranges for nine traits of the CO940610/Platte population (n=186) in Akron and Greeley in the 2009/10 growing season.

Traits	Akron (rainfed)		Greeley (fully irrigated)		*P*†
	Mean ± SE	Min. - Max.	Mean ± SE	Min. - Max.	
Grain yield, kg ha^{-1}	3245 ±16	1938 - 3868	4506 ± 24	2973 - 5175	<0.001
Spike length, cm	8.2 ± 0.02	7.5 - 9.2	9.15 ± 0.03	8.1 - 10.2	<0.001
Plant height, cm	65.3 ± 0.18	59.9 - 71.6	75.89 ± 0.24	68.4 - 84.6	<0.001
Grain protein conc., g kg^{-1}	146.7 ± 0.30	133.6 - 164.4	139.17 ± 0.39	127.1 - 159.3	<0.001
Grain ash conc., g kg^{-1}	14.3 ± 0.04	13.0 - 17.0	14.45 ± 0.04	12.9 - 17.2	0.028
Test weight, kg hL^{-1}	74.92 ± 0.07	71.3 – 77.2	77.82 ± 0.05	73.8 - 79.4	<0.001
Kernel weight, mg	26.3 ± 0.12	22.8 - 31.9	37.55 ± 0.21	30.6 - 47.5	<0.001
Kernel diameter, mm	2.6 ± 0.00	2.4 - 2.8	2.95 ± 0.01	2.7 - 3.3	<0.001
Kernel hardness (0 – 100)	72.3 ± 0.27	58.1 - 80.8	70.87 ± 0.31	53.8 - 83.0	<0.001

† Significance of difference between the means

Parent performance

CO940610 had significantly higher Gy, Ph and Kh than Platte in both environments. Meanwhile, Platte had significantly longer Sl and higher Gpc than CO940610 in both environments (Table 6).

Table 6. Means for nine traits of the two parents, CO940610 and Platte, at Akron and Greeley, CO in the 2009/10 growing season.

	Akron (rainfed)			Greeley (fully irrigated)		
Trait	**CO940610**	**Platte**	***P*†**	**CO940610**	**Platte**	***P***
Grain yield, kg ha^{-1}	3538	2739	<0.001	4827	4685	<0.001
Spike length, cm	7.57	8.72	0.019	8.50	9.70	<0.001
Plant height, cm	66.39	60.31	<0.001	78.04	72.33	0.009
Grain protein conc., g kg^{-1}	141.01	156.90	<0.001	124.61	157.26	<0.001
Grain ash conc., g kg^{-1}	14.18	14.89	0.072	13.85	14.65	0.125
Test weight, kg hL^{-1}	75.47	74.62	<0.001	74.43	78.06	<0.001
Kernel weight, mg	29.12	25.57	<0.001	38.23	38.79	0.775
Kernel diameter, mm	2.64	2.57	0.045	2.90	3.12	0.002
Kernel hardness (0 - 100)	73.05	67.72	0.057	75.31	59.34	<0.001

† Significance of difference between the parents

2.2.2. Correlation among traits

The correlations among traits in the two environments are shown in Table 7.

Grain yield was positively correlated with Ph (r=0.36, P<0.001 in Akron and r=0.23, P<0.01 in Greeley) and Tw (r=0.27, P<0.001 in Akron and r=0.38, P<0.001 in Greeley). In contrast, Gy was significantly and negatively correlated with Gpc (r=-0.36, P<0.001 in Akron and r=-0.34, P<0.001 in Greeley) and Gac (r=-0.29, P<0.001 in Akron and r=-0.38, P<0.001 in Greeley). This confirmed that Gy has an inverse association with Gpc and Gac.

Ph was also positively correlated with Sl (r=0.17, P<0.05 in Akron and r=0.29, P<0.001 in Greeley), Tw (r=0.34, P<0.001 in Akron and r=0.43, P<0.001 in Greeley), Kw (r=0.25, P<0.001 in Akron and r=0.43, P<0.001 in Greeley) and Kd (r=0.25, P<0.001 in Akron and r=0.37, P<0.001 in Greeley).

Grain protein concentration and Gac were also negatively correlated with Tw. Correlation coefficients between Gpc and Tw were -0.28 (P<0.001) and -0.14 (P<0.05) in Akron and Greeley, respectively; and between Gac and Tw were -0.53 (P<0.001) and -0.36 (P<0.001) in Akron and Greeley, respectively. This indicated that an increase in Tw was associated with a decrease in Gpc and Gac.

Test weight, Kw and Kd were highly significantly (P<0.001) correlated with each other, indicating that there are proportional relationships among these traits. Of these, Kw and Kd were very highly correlated (r=0.94, P<0.001 in both environments).

Kernel diameter and Kh were negatively correlated (P<0.001). Correlation coefficients between them were -0.24 (P<0.001) and -0.33 (P<0.001) in Akron and Greeley, respectively. This suggests that smaller kernel is associated with a softer kernel.

Table 7. Pearson correlation coefficients among traits of the CO940610/Platte population (n=186) at Akron and Greely in the 2009/10 growing season. Correlations for Akron are below the diagonal and those for Greeley are above the diagonal.

Trait†	Gy	Sl	Ph	Gpc	Gac	Tw	Kw	Kd	Kh
Gy		-0.02	0.23**	-0.34***	-0.38***	0.38***	-0.13	-0.16*	0.04
Sl	-0.12		0.29***	-0.03	0.07	0.24**	0.04	-0.01	0.08
Ph	0.36***	0.17*		-0.08	0.09	0.43***	0.43***	0.37***	-0.03
Gpc	-0.36***	-0.01	-0.39***		0.46***	-0.14*	0.29***	0.36***	-0.27***
Gac	-0.29***	0.10	-0.04	0.45***		-0.36***	0.27***	0.25***	-0.07
Tw	0.27***	0.01	0.34***	-0.28***	-0.53***		0.34***	0.31***	-0.02
Kw	-0.04	-0.03	0.25***	0.06	0.08	0.47***		0.94***	-0.41***
Kd	-0.01	-0.08	0.25***	0.09	0.07	0.45***	0.94***		-0.33***
Kh	-0.03	-0.03	0.01	0.07	0.04	-0.07	-0.24***	-0.14	

*, **, *** indicates significance at the 0.05, 0.01, and 0.001 levels of probability, respectively.

† Gy, grain yield; Sl, spike length; Ph, plant height; Gpc, grain protein concentration at 12% moisture level; Gac, grain ash concentration; Tw, test weight; Kw, kernel weight; Kd, kernel diameter; Kh, kernel hardness.

2.2.3. Heritability

The estimated heritability values for each measured phenotypic trait in each environment and combined across environments are presented in Table 8.

The heritability estimates of the traits were low (<0.50), moderate (0.50 to 0.75), and high (0.75). Grain yield showed low heritability estimates in each environment and combined across environments. This indicated that Gy is largely affected by environments, making it challenging to improve in breeding programs.

Grain protein concentration had low heritability in Akron, but moderate heritability in Greeley, suggesting that environment considerably influenced the heritability estimates. In contrast, Sl, Gac, Tw, Kw, Kd and Kh had moderate to high heritability estimates in each environment and combined across environments, indicating that for these traits a large part of the observed variation was due to genetics.

Table 8. Heritability estimates (H^2) and 90% confidence intervals for nine traits of CO940610/Platte RIL population in Akron and Greeley in the 2009/10 growing season.

Trait	Akron (rainfed)			Greeley (fully irrigated)			Two locations		
	H^2	90% confidence†		H^2	90% confidence		H^2	90% confidence	
		Lower limit	Upper limit		Lower limit	Upper limit		Lower limit	Upper limit
Grain yield	0.15	-0.08	0.33	0.14	-0.09	0.33	0.25		
Spike length	0.53	0.40	0.63	0.65	0.56	0.73	0.74	0.68	0.79
Plant height	0.37	0.20	0.50	0.60	0.49	0.69	0.63	0.55	0.70
Grain protein concentration	0.20	-0.02	0.37	0.55	0.43	0.65	0.42	0.28	0.53
Grain ash concentration	0.68	0.59	0.75	0.75	0.68	0.80	0.76	0.70	0.81
Test weight	0.69	0.60	0.76	0.70	0.62	0.77	0.69	0.60	0.76
Kernel weight	0.76	0.69	0.81	0.78	0.72	0.83	0.79	0.73	0.83
Kernel diameter	0.73	0.66	0.79	0.79	0.73	0.84	0.80	0.75	0.84
Kernel hardness	0.79	0.73	0.83	0.74	0.67	0.79	0.82	0.77	0.86

† Based on Knapp et al. (1985)

2.2.4. Marker analysis

Eighteen markers were genotyped, which included one sequence-tagged site (STS) and 17 SSR (Table 9). These markers were screened for polymorphisms between CO940610 and Platte before genotyping the population of 186 RIL. Four markers on chromosome 6B, *Barc136, Barc198, Xwmc182a,* and *Xwmc397,* had an excess of lines with the CO940610 allele. Segregation of the markers on 6B was significantly ($P<0.05$) distorted compared to the expected segregation ratio (Table 9). Three of thirteen markers on 7B (*Barc278, Xgwm46,* and *Xwmc364*) had an excess of lines with the Platte allele, with distorted segregation at $P<0.01$. The other markers conformed to the Mendelian independent segregation ratio.

Table 9. The goodness of fit of observed marker data for the CO940610/Platte RIL population based on deviation from expected segregation for the $F_{5:6}$ generation.

Marker	Type of marker	Chr.	Number of RIL								Chi-square
			Platte allele		CO940610 allele		Heterozygote		Genotyped		
			E‡	O	E	O	E	O	U	**M**	
Bx7-MAR	STS†	1B	84.5	85	84.5	84	12	17	186	0	0.00
Barc136	SSR	6B	91	111	91	71	11	0	182	4	8.36**
Barc198	SSR	6B	90	118	90	62	12	5	185	1	16.81**
Xwmc182a	SSR	6B	92	120	92	64	12	2	186	0	16.44**
Xwmc397	SSR	6B	90	119	90	61	11	1	181	5	18.05**
Xwmc182b	SSR	7B	93	85	93	101	12	0	186	0	1.21
Xgwm606	SSR	7B	91.5	93	91.5	90	11	0	183	3	0.02
Xgwm569	SSR	7B	92.5	92	92.5	93	12	0	185	1	0.00
Barc278	SSR	7B	89	69	89	109	12	8	186	0	8.54**
Xgwm333	SSR	7B	91	82	91	100	12	2	184	2	1.59
Xgwm46	SSR	7B	92.5	73	92.5	112	12	1	186	0	7.81**
Xgwm573	SSR	7B	91	81	91	101	12	3	185	1	1.98
Xgwm76	SSR	7B	92	80	92	104	12	0	184	2	2.88
Xgwm611	SSR	7B	92	87	92	97	12	0	184	2	0.44
Xwmc17a	SSR	7B	88	78	88	98	11	2	178	8	2.05
Xwmc17b	SSR	7B	90.5	78	90.5	103	11	0	181	5	3.18
Xwmc364	SSR	7B	92.5	73	92.5	112	12	1	186	0	7.81**
Xwmc426	SSR	7B	91	82	91	100	12	4	186	0	1.59

† STS, sequence-tagged site; SSR, simple sequence repeat.

‡ E or O are the number of expected or observed marker, respectively; U or M stands for number of markers that were usable or missing out of marker genotyped, respectively.

*, ** indicates segregation distortion is significant at the 0.05 and 0.01 level of probability, respectively.

2.2.5. Construction of linkage map

The linkage groups were created by JoinMap (Van Ooijen 2006) with minimum LOD score of 8.0. Of the 17 SSR markers screened, four markers (*Barc136, Barc198, Xwmc182a,* and *Xwmc397*) mapped to chromosome 6B, while 11 markers (*Xwmc182b, Barc278, Xgwm46, Xgwm76, Xwmc364, Xgwm606, Xgwm569, Xgwm333, Xgwm573, Xwmc17a,* and *Xwmc426*) mapped to chromosome 7B (Figure 7).

The four markers on 6B spanned 9.5 cM, with the average interval between markers of 3.2 cM. Meanwhile, 11 markers on 7B spanned 71.4 cm, with the average interval between markers of 7.1 cM. The *Glu-B1* marker has reliably mapped to chromosome 1B, so no additional markers on this chromosome were evaluated.

Marker (*Xgwm611* and *Xwmc17b*) did not group with any of the other markers. According to maps on the Grain Genes database (http://wheat.pw.usda.gov/cgi-bin/GG3/report.cgi?class=locus;name=Xgwm611), Xgwm611 is located on the long arm of 7B, distant from all the other markers in our study.

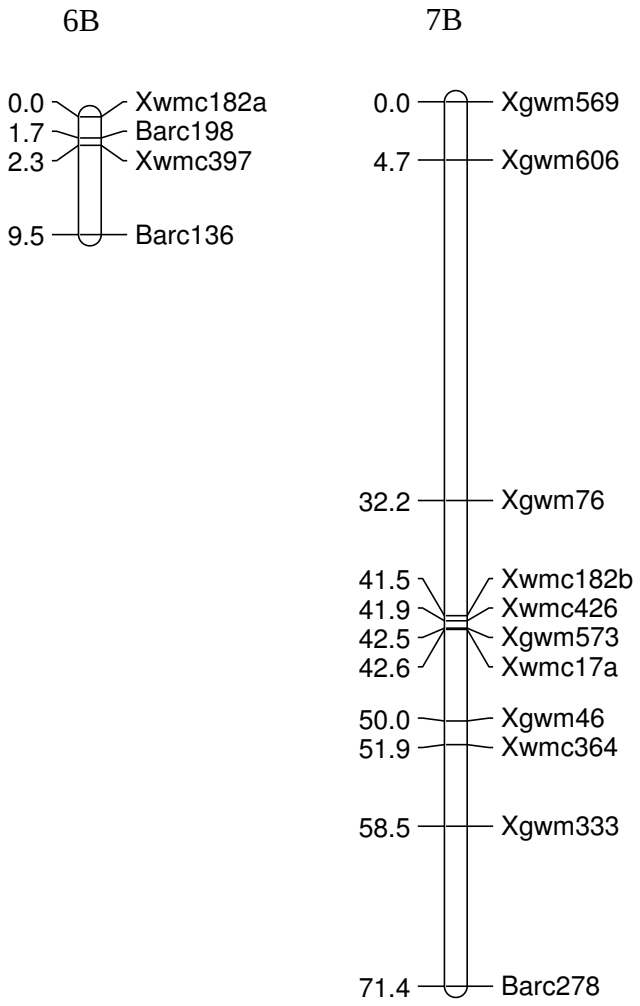


Figure 7. Linkage map constructed in JoinMap for the CO940610/Platte RIL population. Cumulative distances between markers are given in cM, calculated from recombination frequencies according to the Haldane mapping function.

2.2.6. Marker-trait associations

Single-factor analysis of variance conducted with PROC GLM was used to find marker-trait associations (MTA). There were 64 significant MTA among 9 traits in two environments (Table 10). They are also visualized in Figure 8.

Grain yield

Four markers on two chromosomes 1B and 7B were associated with grain yield. Markers *Glu-B1* on chromosome 1B, and *Xgwm611* and *Xwmc17b* on chromosome 7B were each associated with grain yield in one environment. This indicated that these MTA were not robust across environments.

On chromosome 7B, the markers *Xgwm76* and *Xwmc17b* were detected with a positive additive effect, suggesting that Platte contributed the higher grain yield allele while the other (*Xgwm611*) was detected with a negative additive effect, indicating that CO940610 contributed the higher grain yield allele. This suggests that these markers are linked to distinct QTL on the same chromosome.

Spike length

Nine markers on three chromosomes (1B, 6B, and 7B) were associated with Sl, with 13 MTA in both environments. Some of these associations were robust across environments, while others were not. Markers *Glu-B1, Barc136, Xgwm219,* and *Xgwm569* were associated with Sl only in Akron, while marker *Xgwm33* was detected only in Greeley. Meanwhile, markers *Xwmc397, Xwmc182a, Xgwm219, Xgwm569,* and *Xwmc17b* were all associated with Sl in both environments. CO940610 and Platte both contributed longer spike length alleles depending on the loci.

Table 10. Markers associated with traits of the CO940610/Platte RIL population in Akron and Greeley, CO in the 2009/10 growing season.

			Akron Rainfed			Greeley Irrigated		
Trait	**Marker**	**Chromosome**	***P*-value**	**R^2 (%)‡**	**a†**	***P*-value**	**R^2 (%)**	**a**
Grain yield	*GluB1*	1B	0.038	2.6	36.543			
(kg ha^{-1})	*Xgwm76*	7B	0.036	2.4	34.633	0.016	3.1	54.032
	Xgwm611	7B				0.038	2.4	-46.516
	Xwmc17b	7B	0.018	3.1	39.866			
Spike length	*GluB1*	1B	0.037	2.6	-0.046			
(cm)	*Barc136*	6B	0.022	2.9	0.049			
	Xwmc397	6B	0.001	6.0	0.073	0.029	2.7	0.067
	Xwmc182a	6B	0.002	5.2	0.067	0.021	2.9	0.071
	Barc198	6B	0.002	5.5	0.068	0.022	2.9	0.071
	Xgwm569	7B	0.004	4.4	-0.059			
	Xgwm333	7B				0.050	2.1	0.058
	Xwmc17b	7B	0.003	4.7	0.062	<.0001	13.6	0.149
Plant height	*GluB1*	1B				0.047	2.3	-0.501
(cm)	*Xgwm569*	7B	0.048	2.1	-0.348			
	Xwmc17b	7B				0.056	2.0	0.464
Grain protein conc.	*Xgwm569*	7B	0.075	1.7	0.537			
(g kg^{-1})								

Table 10. (Continued)

			Akron Rainfed			Greeley Irrigated		
Trait	**Marker**	**Chromosome**	***P*-value**	***R^2* (%)‡**	**a†**	***P*-value**	***R^2* (%)**	**a**
Grain ash conc.	*GluB1*	1B	0.000	7.3	-0.154	0.035	2.6	-0.094
(g kg^{-1})	*Xgwm606*	7B	0.002	5.4	0.134	0.024	2.8	0.095
	Xgwm569	7B	0.035	2.4	0.090			
	Xgwm573	7B	0.049	2.1	0.085	0.026	2.7	0.097
	Xgwm46	7B	0.009	3.7	0.114	0.017	3.1	0.104
	Xgwm333	7B				0.011	3.6	0.111
	Barc278	7B				0.048	2.2	0.089
	Xwmc17a	7B	0.004	4.7	0.120	0.002	5.3	0.127
	Xwmc364	7B	0.010	3.5	0.112	0.004	4.5	0.125
	Xwmc426	7B	0.051	2.1	0.085	0.042	2.3	0.089
Test weight	*GluB1*	1B	0.042	2.5	0.111			
(kg hL^{-1})	*Barc136*	6B				0.033	2.5	-0.091
	Xwmc397	6B	0.028	2.7	-0.124			
	Xwmc182a	6B	0.045	2.2	-0.110			
	Barc198	6B	0.035	2.5	-0.117			
	Xgwm569	7B				0.022	2.8	-0.094
	Xgwm76	7B				0.041	2.3	0.085
	Xgwm333	7B	0.047	2.2	0.105			
	Xwmc17b	7B	0.010	3.6	0.134	0.000	7.9	0.155
Kernel weight	*GluB1*	1B	0.053	2.2	-0.252	<.0001	9.0	-0.862
(mg)	*Xwmc397*	6B	0.056	2.0	-0.250			
	Xwmc182a	6B	0.024	2.8	-0.288			
	Barc198	6B	0.038	2.4	-0.268			

Table 10. (Continued)

			Akron Rainfed			Greeley Irrigated		
Trait	**Marker**	**Chromosome**	***P*-value**	**R^2 (%)†**	**a‡**	***P*-value**	**R^2 (%)**	**a**
Kernel diameter	*GluB1*	1B				0.028	2.9	-0.017
(mm)	*Barc136*	6B	0.030	2.6	-0.010	0.039	2.4	-0.015
	Xwmc397	6B	0.024	2.8	-0.011	0.054	2.1	-0.015
	Xwmc182a	6B	0.007	3.9	-0.013			
	Barc198	6B	0.007	4.0	-0.013	0.030	2.6	-0.017
	Xgwm46	7B				0.055	2.0	-0.014
Kernel hardness	*Xwmc397*	6B	0.002	5.5	-0.929	0.027	2.7	-0.750
(0-100)	*Xwmc182a*	6B	0.020	2.9	-0.666			
	Barc198	6B	0.004	4.5	-0.826	0.047	2.2	-0.670
	Xgwm569	7B	0.048	2.1	-0.539			
	Xgwm333	7B	0.006	4.2	-0.761			
	Xwmc364	7B	0.022	2.8	-0.637			

† percentage of phenotypic variation explained by the marker.

‡ average additive effect, with a positive value indicating an increasing effect of the Platte allele and a negative value indicating an increasing effect of the CO940610 allele.

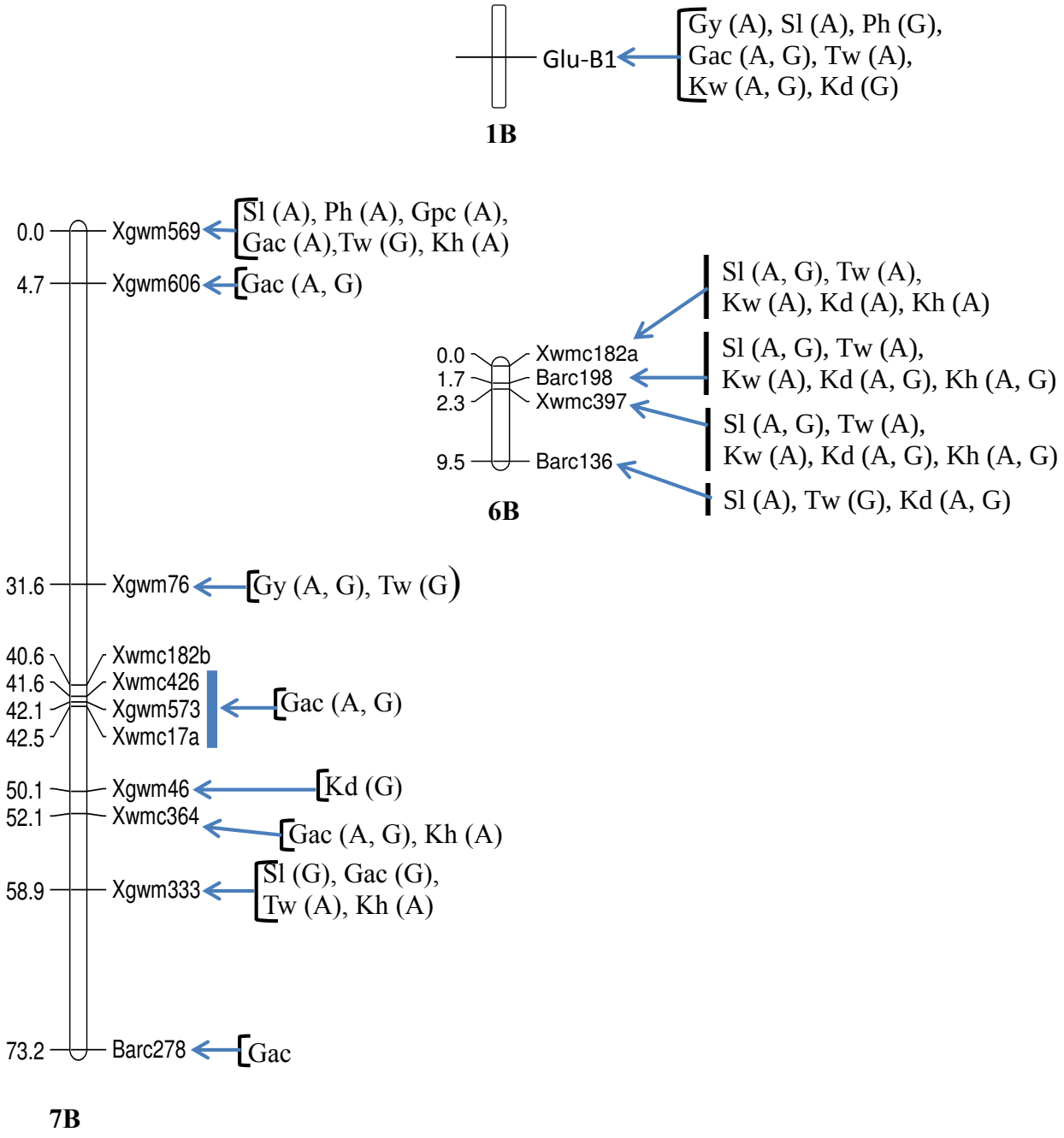


Figure 8. QTL maps of the CO940610/ Platte/ RIL population with location of marker-trait associations indicated. Cumulative distances between markers are given in cM. A = detected in Akron, G = detected in Greeley. Gy, grain yield; Sl, plant height; Sl, spike length; Gpc, grain protein concentration; Gac, grain ash concentration; Tw, test weight; Kw, kernel weight; Kd, kernel diameter; Kh, kernel hardness.

Markers *Glu-B1, Xgwm219,* and *Xgwm569* had a negative additive effect, indicating that CO940610 contributed the allele for longer Sl. Meanwhile, markers *Barc136, Xwmc397, Xwmc182a, Xgwm333,* and *Xwmc17b* detected a positive additive effect, indicating that Platte contributed the longer Sl allele.

It appears that markers are linked to distinct QTL on the chromosome 7B. On chromosome 7B, markers *Xgwm333* and *Xwmc17b* were detected with a positive additive effect for Sl, but marker *Xgwm569* located at a distance of 40 cM from the other markers was detected with a negative effect.

Plant height

Three markers were significantly associated with Ph, but each marker was detected only in a single environment. Markers *GluB1* on 1B and *Xwmc17b* on 7B were associated with Ph only in Greeley, while marker *Xgwm569* was detected only in Akron. This indicated that the MTA for Ph were not robust across environments.

CO940610 and Platte both contributed alleles for higher Ph depending on the loci. Markers *GluB1* on 1B and *Xgwm569* on 7B had a negative additive effect, while marker *Xwmc17b* had a positive additive effect for Ph. This also suggested that markers *Xgwm569* and *Xwmc17b* on chromosome 7B are linked to distinct QTL on the same chromosome.

Grain protein concentration

Grain protein concentration is one of the traits evaluated in this study that did not have any significant associations with the markers at $P < 0.05$. However, marker *Xgwm569* on 7B was associated with Gpc at $P = 0.075$. At this locus, Platte contributed the higher Gpc allele.

Grain ash concentration

One marker on chromosome 1B and nine markers on chromosome 7B were significantly associated with Gac, with a total of 17 associations in both environments. Seven of these markers showed stability across environments. Marker *GluB1* on 1B, and markers *Xgwm606*, *Xgwm573*, *Xgwm46*, *Xwmc17a*, *Xwmc364*, and *Xwmc426* on 7B were associated with Gac in both environments.

However, marker *Xgwm569* on 7B was significantly associated with this trait only in Akron, and the markers *Xgwm333* and *Barc278* on 7B were associated with Gac only in Greeley. Thus these markers were less robust across environments than markers in the central part of chromosome 7B.

Both parents contributed alleles for higher Gac depending on the loci. The marker *GluB1* was detected with a negative additive effect, suggesting that CO940610 contributed the higher Gac allele. Meanwhile, all other markers on chromosome 7B were detected with a positive additive effect, indicating that Platte contributed the higher Gac allele.

Test weight

Nine markers on three chromosomes (1B, 6B, and 7B) were significantly associated with Tw with 10 associations. Most of these markers showed instability across environments. Markers *GluB1* on 1B, *Xwmc397*, *Xwmc182a*, and *Barc198* on 6B, and *Xgwm333* on 7B were significantly associated with Tw only in Akron. Meanwhile, markers *Barc136* on 6B, and *Xgwm569* and *Xgwm76* on 7B were significantly associated with the trait only in Greeley. Only marker *Xwmc17b* on 7B showed significant associations in both environments.

Both parents contributed alleles for greater Tw. Markers *Barc136, Xwmc397, Xwmc182a,* and *Barc198* on 6B, and *Xgwm569* on 7B were detected with a negative additive effect, indicating CO940610 contributed greater Tw. Meanwhile, markers *GluB1* on 1B, *Xgwm76, Xgwm333,* and *Xwmc17b* on 7B had a positive effect, suggesting Platte contributed the alleles for greater Tw.

On chromosome 7B, marker *Xgwm569* had a negative additive effect while markers *Xgwm76, Xgwm333,* and *Xwmc17b* were detected with positive effects, demonstrating that these markers are linked to distinct QTL on this chromosome.

Kernel weight

Four markers on chromosomes 1B and 6B were significantly associated with Kw, with five associations. All of these associations were detected with a negative additive effect, indicating that CO940610 contributed the greater Kw allele.

The MTA for Kw on 6B were not robust across environments. Markers *Xwmc397, Xwmc182a,* and *Barc198* on 6B were significantly associated with Kw only in Akron. However, marker *GluB1* on 1B showed robust and significant associations across both environments.

Kernel diameter

Six markers on all three chromosomes were significantly associated with Kd for a total of nine associations. Some of these associations were not robust across environments. Markers *GluB1* on 1B and *Xgwm46* were significantly associated with Kw only in Greeley, while marker *Xwmc182a* on 6B had a significant association with the trait only in Akron.

However, the other associations showed stability across environments. Markers *Barc136, Xwmc397,* and *Barc198* on 6B were associated with Kd in both environments.

All the significant markers associated with this trait had a negative additive effect, suggesting that CO940610 contributed the allele for increased Kd.

Kernel hardness

Six markers on chromosomes 6B and 7B were significantly associated with Kh, with eight associations in both environments. Some of the associations were not robust across environments. Marker *Xwmc182a* on 6B, and markers *Xgwm569, Xgwm333,* and *Xwmc364* on 7B were significantly associated with Kh only in Akron. Meanwhile, the other associations were stably detected over environments. Markers *Xwmc397* and *Barc198* on 6B had significant associations with Kh in both environments.

All of these associations had a negative additive effect, indicating that CO940610 was the source of alleles for harder kernels.

Stability of QTL across environments

Nine traits evaluated in this study were significantly associated with 16 markers. Some MTA were robust across environments, while some other associations were not. A total of 45 significant MTA was detected. Nineteen of the total associations (42%) appeared in both environments, where 17 associations (38%) occurred only in Akron, and 20% were present only in Greeley (Table 11).

Chromosome 6B had a higher number of robust QTL, while chromosome 1B and 7B had more unstable QTL (Table 11). On chromosome 6B, there were 21 significant MTA. Eight of these associations were detected in both environments, while 13 others appear in either Akron or Greeley. In contrast, only two of seven significant MTA on chromosome 1B appeared in both

environments. On chromosome 7B, nine of 22 significant MTA occurred in both environments, while 13 of these associations were present in either Akron or Greeley.

Table 11. Number of significant marker-trait associations by chromosome. Stable associations were those detected in both environments and unstable ones were detected only in one environment.

	Significant marker-trait associations			Total	Percentage
	1B	6B	7B		
Stable	2	8	9	19	42
Unstable	5	13	13		
- Akron	*3*	*12*	*5*	17	38
- Greeley	*2*	*1*	*8*	9	20
Total	7	21	22	45	

Comparison of QTL in RIL and DH populations

Consistency of MTA or QTL in the DH and RIL population are summarized in Table 12. At *Glu-B1*, CO940610 contributed the higher value allele for Sl, Ph, Gac, Kw and Kd, while Platte was the source of the higher value allele for Gy and Tw. Some MTA for evaluated traits in the RIL corresponded to QTL in the DH at three chromosomal regions 1B, 6B and 7B.

Table 12. Environments[†] in which QTL were detected in the doubled haploid (DH) (El-Feki 2010) and recombinant inbred line (RIL) populations derived from the cross CO940610/Platte.

		DH				RIL	
Trait	**Chromosome**	**08FD**	**08FW**	**09GD**	**09GW**	**10AD**	**10GW**
Grain yield	1B					P	
	7B			CO		P	P, CO
Spike length	1B					CO	
	6B		P	P		P, CO	P
	7B			P		P, CO	P
Plant height	1B						CO
	6B						CO
	7B					CO	P
Grain protein conc.	6B		CO				
	7B	P	P	P		P	
Grain ash conc.	1B	CO	CO	CO	CO	CO	CO
	6B					P	P
	7B			P		P	P
Test weight	1B				P	P	
	6B		CO			CO	CO
	7B					P	P, CO
Kernel weight	1B	CO		CO		CO	CO
	6B					CO	
Kernel diameter	1B						CO
	6B					CO	CO
	7B		P				CO
Kernel hardness	6B	CO	CO	CO		CO	CO
	7B					CO	

[†] 08FD, Fort Collins Dry 2007/08; 08FW, Fort Collins Wet 2007/08; 09GD, Greeley Dry 2008/09; 09GW, Greeley Wet 2008/09; 10AD, Akron Rainfed 2009/10; and 10GW, Greeley fully irrigated 2009/10.

[‡] CO and P indicate CO940610 and Platte were the source of the higher value allele, respectively.

2.3. DISCUSSION

2.3.1. Trait means

Grain yield was reduced about 30% in Akron rainfed environment, compared to the fully irrigated Greeley environment. Reduced water is probably the major factor, although other environmental differences could have played a role. El-Feki (2010) reported that Gy was reduced about 19 to 21% under water deficit treatments in the CO940610/Platte DH population. Edae et al. (2014) reported 50% Gy reduction at Greeley rainfed compared to Greeley fully irrigated conditions. Grain yield reduction of 57% was due to drought stress in the study by Balla et al. (2011). Water stress during grain filling causes less sucrose and starch accumulation in grain (Ahmadi & Baker 2001), resulting in Gy reduction.

Although El-Feki (2010) found that Sl was significantly shorter in the fully irrigated treatment compared to the limited irrigation treatment in the DH population, this study found that Sl was reduced under drought stress. Amiri et al. (2013) also reported terminal drought stress reduced Sl by 2.58%. Mirbahar et al. (2009) also found Sl was significantly shorter in terminal drought. The reason for the difference between our results and El-Feki's results is not immediately apparent.

Plant height was shorter under drought stress in agreement with El-Feki (2010), Amiri et al. (2013), and Mirbahar et al. (2009). Plant height is associated with the movement of carbohydrates in the plant, especially under water stress condition (Blum et al. 1989).

Grain protein concentration increased under drought conditions, in agreement with previous reports. For example, El-Feki et al. (2013) found that Gpc was significantly higher under the limited irrigation compared to the full irrigation in the DH population. Water stress and available nitrogen are two major factors influencing Gpc (Daniel & Triboı 2002; Zheng et al. 2009; Prasad

et al. 2003). Drought stress during grain filling reduced sucrose and starch in grain (Ahmadi & Baker 2001; Foulkes et al. 2002), resulting in higher final Gpc (Weightman et al. 2008).

Test weight has been reported to be both higher and lower under full irrigation. The current study showed that Tw in Akron (74.92 kg hL^{-1}) was significantly (P<0.001) lower than in Greeley (77.82 kg hL^{-1}) in the CO940610/Platte RIL population. In the hard winter wheat association mapping panel (HWWAMP), Tw in the water deficit treatment was significantly lower than in the fully irrigated treatment in ARDEC 2012/13 (Chapter 4 of this dissertation). Conversely, some other reports showed Tw is significantly higher in the water deficit condition compared to the fully irrigated condition. In the same HWWAMP, Tw in water deficit treatment was significantly higher than in the fully irrigated treatment in LIRF 2011/12 (Chapter 4 of this dissertation). This also agreed with the CO940610/Platte RIL DH population (El-Feki 2010) and the wheat association mapping II (WAMII) (Edae et al. 2014; Edae 2013). Test weight is genetically controlled, and is influenced by many other factors e.g., grain moisture, dryer, drying methods, measurements, and kernel shapes. Experimental management also influenced Tw. These may cause the contradictory results.

Kernel weight was reduced in the rainfed environment. About 30% reduction of Kw was recorded in Akron compared to Greeley (Table 5). Drought stress during grain filling reduces sucrose and starch (Ahmadi & Baker 2001), consequently decreasing Kw.

2.3.2. Correlation among traits

Grain yield was negatively correlated with Gpc and Gac, a finding that agrees with previous reports (El-Feki et al. 2013; Bilgin et al. 2010; Hrušková & Švec 2009). This suggests that simultaneously breeding for Gy and Gpc is a challenging goal. However, Gy had positive

correlation with Tw (Bilgin et al. 2010; El-Feki et al. 2013) and Ph (Edae et al. 2014). This indicated that selection of greater Tw may lead to higher Gy.

Test weight, Kw and Kd were positively correlated with each other (P<0.001). This agreed with the previous reports (El-Feki et al. 2013). Of these relationships, Kw and Kd were tightly correlated with each other. The correlation coefficient between them was 0.94 (P<0.001) in both environments. El-Feki et al. (2013) reported correlation coefficient between Kw and Kd ranged from 0.89 to 0.91 (P<0.01) in the DH under four environments. This suggested that improving one of these traits would likely improve the other.

2.3.3. Heritability estimates

Heritability estimates of Gpc were 0.20 and 0.55 in Akron and Greeley, respectively (Table 8). Grain protein concentration was complex character, controlled by many genes and with large influence by environmental conditions. Clarke et al. (2009) reported that heritability estimate for Gpc ranged from 0.43 to 0.84, and suggested that heritability estimates are influenced by the number of replications, years and locations tested, and by the genetics of the populations. Barnard et al. (2002) estimated broad-sense heritability of Gpc at 0.57. In the DH population, heritability estimates for Gpc in four environments ranged from 0.52 to 0.61 (El-Feki 2010).

Spike length, Gac, Tw, Kw, Kd and Kd had moderate (0.50-0.75) to high (>0.75) heritability and stability across environments, indicating that the majority of their phenotypic variation is due to genetics. Therefore, it can be easier to make progress from selection for these traits in breeding programs. El-Feki (2010) reported the range of their heritability estimates for the DH population in four environments as follows: Sl (0.54 – 0.78), Gac (0.73 – 0.77), Tw (0.73 – 0.77), Kw (0.78 – 0.87), Kd (0.76 – 0.86), and Kh (0.79 – 0.88). Meanwhile, Gy showed low heritability

estimates in each environment and combined across environments. This indicated that Gy is highly affected by environments, making it challenging to improve in breeding programs.

Heritability estimates based on combined data for the two environments increased in comparison with individual environment estimates (Table 8). By combining data over multiple environments, the heritability estimates are expected to rise with an increasing number of environments when the genotype by environment interaction is large (Hill et al. 2012). A higher number of samples and replications increases the accuracy, reducing effects of environmental factors.

2.3.4. Marker analysis and genetic map construction

The 18 markers genotyped and mapped consisted of one STS (*Glu-B1*) and 17 SSR. Actually, 15 of these 17 SSR were mapped into two linkage groups 6B (4) and 7B (11).

Four markers *Xwmc182a, Barc198, Xwmc397* and *Barc136* on 6B were mapped in a linkage group of 9.5 cM in the RIL (Figure 2.2.5.1). El-Feki et al. (2013) also mapped these four markers in linkage group 6B.1 within a distance of 12.2 cM in the DH population, but in a slightly different order. This indicated that the two studies had nearly the same mapping result, with four tightly linked markers. Similarly, the other 11 SSR were mapped on 7B, spanning 71.4 cM. Seven of these 11 markers were mapped in one linkage group spanning 85.9 cM in the DH population (El-Feki et al. 2013), three (*Xgwm46, Xwmc17a* and *Xwmc364*) were newly added in the RIL population. The two studies strongly agreed, though not exactly, in marker position and order.

One of 17 SSR, *Xwmc182b*, was mapped on the 7B linkage group in the RIL population. This agreed with El-Feki (2010). However, this marker was not included in the 7B map of the DH population that appeared in El-Feki et al. (2013). Marker *Xgwm611* was not linked to this 7B

linkage group in the RIL population, although it was mapped on 7B in the DH population, but distant (52.4 cM) from the dense (central) markers of the linkage group (El-Feki et al. 2013).

2.3.5. Marker-trait associations

El-Feki (2010) and El-Feki et al. (2013) reported many QTL for agronomic and grain quality traits in the CO940610/Platte DH population in four environments. The current study used the same parents (CO940610 and Platte) to develop a RIL population for the purpose of validating selected QTL from El-Feki's studies.

In this study, a total 64 MTA for nine traits were detected in three chromosomal regions (1B, 6B and 7B) in two environments. Of these, the numbers of MTA recorded were as follows: Gy (5), Sl (12), Ph (3), Gpc (1), Gac (17), Tw (10), Kw (5), Kd (9), and Kh (8). Some MTA coincided with QTL found in the DH population, while other MTA were located close to, far apart from, or contradictory to QTL detected in the DH population.

Grain yield

One MTA on 1B and four MTA on 7B were detected for Gy. The grain yield-*Xgwm76* association on 7B coincided with QTL for Gy detected in the DH population (El-Feki 2010). However, the current study found that Platte contributed the higher Gy allele, while El-Feki (2010) found CO940610 conferred the higher Gy allele at this locus. This contradictory finding should be further investigated. In addition, no QTL for Gy was detected on 1B in the DH population.

Spike length

One MTA on 1B, seven MTA on 6B and four MTA on 7B were detected for Sl. Of these, the seven MTA for Sl on 6B coincided with QTL for Sl found in the DH population (El-Feki 2010).

The Sl-*Xgwm333* association on 7B was 9.1 cM apart from the QTL interval (flanking markers) for Sl detected in the DH population. This indicated that some of the results conformed to the previous study.

Plant height

One MTA on 1B and two MTA on 7B were detected for Ph in the current study. However, none of these MTA were found in the DH population (El-Feki 2010).

Grain protein concentration

There was no MTA detected for Gpc in the current study, with the exception of marker *Xgwm569* on 7B, with $P = 0.075$. Platte conferred the higher Gpc allele (Table 10). El-Feki et al. (2013) detected QTL for Gpc on 7B across three environments, 08FD, 08FW and 09GD in the DH population. Platte contributed the higher Gpc allele at this QTL. The QTL interval was between *Xgwm569* and *Xbarc278*, spanning 32.7 cM. SSR markers within this QTL interval included *Xgwm569, Xgwm76, Xgwm573, Xgwm46, Xgwm333* and *Xbarc278*. This suggested that Gpc-*Xgwm569* association in the current study conformed to the results of El-Feki et al. (2013). This MTA seems to be detected more often in drought environments (Table 12).

Grain ash concentration

One MTA on 1B and 16 MTA on 7B were detected for Gac. The Platte alleles at markers on 7B conferred higher Gac, while the CO940610 at *Glu-B1* conferred higher Gac (Table 4). El-Feki et al. (2013) also detected QTL for Gac in these two chromosomal regions. On chromosome 1B, *Glu-B1* was significantly associated with Gac in both Akron and Greeley in the RIL population, while QTL for Gac were detected in all four environments in the DH population (El-Feki et al. 2013). El-Feki et al. (2013) also reported that the QTL on 1B had the higher contribution from

the CO940610 allele. This suggested that the Gac-*Glu-B1* association agreed partially with findings from El-Feki et al. (2013), although *Glu-B1* was not included in the QTL interval detected in the DH population. However, *Glu-B1* was located on the same chromosomal region 1B.1 detected by El-Feki et al. (2013). In addition, the MTA or QTL for Gac on 1B was robust across environments.

Of nine markers significantly associated with Gac on 7B in the RIL population, five markers were located in the same chromosomal region, spanning 22.5 cM, that El-Feki et al. (2013) detected. These MTA showed more stability across environments than QTL detected in the DH population, where a QTL was found only in 09GD. Platte also contributed the higher value Gac allele at QTL detected in 09GD. Overall, the MTA or QTL for Gac on 1B and 7B in both populations concurred.

Test weight

One MTA on 1B, four MTA on 6B and five MTA on 7B were detected for Tw in the RIL population. Meanwhile, El-Feki et al. (2013) detected QTL for this trait on 1B and 6B, in 09GW and 08FW, respectively.

Glu-B1 was significantly associated with Tw in Akron, and Platte contributed the higher Tw allele. The QTL interval for Tw in the DH population was exactly the same as for Gac. The QTL for Tw in 09GW was also found to have the heavier Tw allele from Platte (El-Feki et al. 2013). This suggested that MTA or QTL for Tw on 1B from both studies agreed.

All four MTA on 6B were significantly associated with Tw and were proximate (2.9 cM apart) to the peak position of the QTL for Tw detected in the DH population. CO940610 contributed the

higher value Tw allele at these MTA or QTL. This indicated the MTA for Tw on 6B in the current study were similar to the findings of El-Feki et al. (2013).

Kernel weight

Two MTA on 1B and three MTA on 6B were detected for Kw in the RIL population. The MTA for Kw on 1B overlapped QTL for Kw on 1B detected in the DH population. Marker *Glu-B1* fell into the QTL interval for Kw detected in the DH population. Both studies showed CO940610 contributed the heavier Kw allele. This indicates the repeatability of the Kw-*Glu-B1* association, which explained 2.2 – 9.0% of Kw variation (Table 10). However, El-Feki et al. (2013) did not detect any QTL for Kw on 6B, which was contradictory to these results.

Kernel diameter

One MTA on 1B, seven MTA on 6B, and one MTA on 7B were detected for Kd in the RIL population. El-Feki et al. (2013) detected one QTL on 7B in two environments, but 75.4 cM apart from MTA for Kd found in the current study. This suggested that MTA for Kd did not coincide with the results of El-Feki et al. (2013).

Kernel hardness

Five MTA on 6B and three MTA on 7B were detected for Kh. Previously, El-Feki et al. (2013) found QTL for Kh on 6B stable across three environments (08FW, 08 FD and 09GD). Five MTA for Kh on 6B in the RIL population corresponded to El-Feki's linkage group 6B.1, while QTL for Kh found in the DH population was in linkage group 6B.2. This suggested that no MTA detected for Kh in the current study was the same as QTL for Kh in the DH population. However, both studies showed that CO940610 contributed the allele for increased Kh at MTA and QTL on chromosome 6B.

Linkage group 6B

Four markers (*Xwmc182a, Barc198, Xwmc397* and *Barc136*) were mapped in a linkage group on 6B, spanning a 9.5 cM interval. There were 26 MTA detected for five traits on this linkage group including Sl (7), Tw (4), Kw (3), Kd (7), Ha (5). Platte contributed the higher value Sl allele, explaining 2.9 - 6.0% of the phenotypic variation, while CO940610 contributed the higher value alleles for Tw, Kw, Kd and Ha, with 2.0 - 5.5% of the phenotypic variation explained. These MTA also showed stability across environments.

El-Feki et al. (2013) mapped exactly the same markers in the same position on 6B, spanning a 12.2 interval. The authors detected QTL for mixograph peak time that coincided within this linkage group and QTL for Tw in proximity to this group of markers, just 2.9 cM apart. By using meta-QTL analysis, Tyagi et al. (2014) detected QTL for Fusarium head blight (FSHB) on 6B, which coincided exactly with *Xwmc397*. Gupta et al. (2011) also used meta-analysis and detected QTL for FSHB, Gpc and grain weight on 6B, although it is difficult to determine which region of the chromosome is involved.

This suggests that the linkage group on 6B is of value for wheat improvement. It should be further investigated, and may prove to be useful in breeding programs to improve grain quality and yield of wheat under both rainfed and irrigated conditions.

Linkage group 7B

Four markers (*Xwmc182b, Xwmc426, Xgwm573,* and *Xwmc17a*) were mapped in a linkage group on 7B, spanning a 1.9 cM interval. Marker *Xwmc182a* was associated with Gpc in the DH population (El-Feki 2010) and in the BC_3 backcross population (Chapter 3 of this dissertation). Markers *Xwmc426, Xgwm573,* and *Xwmc17a* were associated with Gac across two environments

(Akron and Greeley). Grain ash concentration has positive correlation with Gpc. This suggests that this linkage group around *Xwmc182b* may be useful for breeding program.

Conclusions

- Linkage maps constructed in the CO940610/Platte RIL and DH populations were mostly consistent.
- Similar patterns of association were found for most grain quality traits investigated with the three chromosomal regions on 1B, 6B, and 7B, confirming previous findings in the CO940610/Platte DH population.
- Tightly linked markers for grain quality traits located around *Xwmc182a* on 6B and *Xwmc182b* on 7B may be useful for selection.

CHAPTER 3: VALIDATION OF QUANTITATIVE TRAIT LOCI FOR GRAIN QUALITY TRAITS IN WINTER WHEAT USING A CO940610/PLATTE BACKCROSS POPULATION

SUMMARY

Grain quality traits affect the end uses of wheat. Important traits include grain protein concentration (Gpc), grain ash concentration (Gac), test weight (Tw), kernel weight (Kw), kernel diameter (Kd), and kernel hardness (Kh). A previous study reported three chromosomal regions (1B, 6B, and 7B) associated with QTL co-located for grain quality traits in a doubled haploid (DH) population derived from the cross CO940610/Platte. That study found that the higher value Gpc allele was contributed by the parent CO940610 at *Glu-B1* on 1B and *Xwmc182a* on 6B, and by the parent Platte at *Xwmc182b* on 7B.

To validate the previous results, the objectives of this study were to (1) transfer the alleles at three loci (*Glu-B1*, *Xwmc182a*, and *Xwmc182b*) from CO940610 into the genetic background of Platte through marker-assisted backcrossing (MABC); (2) evaluate the population of 35 selected BC_3F_2 lines (MABC lines), grown in Fort Collins and Greeley firrigated, and Greeley water deficit environments in the 2012/13 growing season, for Gpc, Gac, Tw, Kw, Kd, and Kh; and (3) conduct marker-trait association (MTA) analysis for the traits evaluated in the MABC lines.

Near-infrared spectroscopy (NIRS) was used to determine Gpc, Gac and Tw. The Single Kernel Characterization System (SKCS) was employed to measure Kw, Kd, and Kh. Correlation coefficients and heritability estimates were calculated for each trait. Marker-trait associations were detected by single-factor analysis of variance (ANOVA).

Test weight, Kw, and Kd were positively correlated with each other. In particular, Kw and Kd were highly correlated (r=0.85 to 0.93, P<0.001). Kernel weight and Kh had high (>0.75) heritability estimates, while the other evaluated traits had low (<0.50) to moderate (0.50-0.75) heritability estimates.

There were 13 significant MTA detected for Ph, Gpc, Gac, Kw, and Kd in either one or two environments. Of these, the MTA for Gpc and Gac were more robust across environments than the others. CO940610 and Platte were observed to contribute the higher value Gpc allele at *Xwmc182a* and *Xwmc182b*, respectively. In Fort Collins fully irrigated conditions, the Gpc of the MABC line with allelic combination CO-CO-PL at *Glu-B1-Xwmc182a-Xwmc182b* was significantly higher than the line carrying PL-PL-CO, confirming the hypothesis that the former and the latter combinations have the highest and lowest Gpc, respectively. The findings were consistent, in general, with results for the DH population. However, the MTA for Gpc detected at *Glu-B1* was contrary to the previous findings.

3.0. INTRODUCTION

Wheat grain can be used for various types of human food or as a livestock feed. It is used primarily to produce flour for bread, but is used widely in the production of many other baked goods, cooked products, and alcoholic beverages. These end-use products are determined by several wheat grain characteristics. Wheat breeding programs put considerable effort into improving grain quality in addition to increasing yield. Grain quality is characterized by several quality parameters such as grain protein concentration (Gpc), grain protein composition, grain ash concentration (Gac), flour color, test weight (Tw), polyphenol oxidase activity, dough-mixing properties, and single kernel characteristics.

Grain protein concentration is one of the grain quality traits in wheat that can have a significant impact on the end-use quality of the product (Khan et al. 2000). Seed storage proteins in wheat consist of two main groups, gliadins and glutenins, which are the major components of wheat gluten structure. Gluten proteins, which comprise up to 85% of endosperm proteins, play a key role in the determination of dough and bread-making quality. Gliadins are important in dough viscosity and extensibility, while glutenins affect dough strength. Gliadins are separated into four groups, α, ß, γ, and ω. Glutenins are classified into two unequal groups, the low molecular weight subunits (LMW-GS) and the high molecular weight subunits (HMW-GS) (Payne 1987). The HMW-GS are controlled by orthologous genes *Glu-1* on the long arm of chromosomes 1A, 1B, and 1D. The LMW-GS are governed by orthologous genes *Glu-3*, and γ- and ω-gliadins encoded by *Gli-1* and *Gli-3* genes on the short arm of the same set of chromosomes. Meanwhile, α- and ß-gliadins are encoded by *Gli-2* genes on the short arm of chromosomes 6A, 6B, and 6D (Payne 1987).

A previous QTL mapping study conducted in four environments found clusters of QTL for grain quality characteristics in regions of chromosomes 6B and 7B in the population of 185 DH lines (El-Feki 2010; El-Feki et al. 2013). A cluster of seven QTL on 6B included two for mixograph peak time, three for Kh, one for Gpc, and one for Tw. Another cluster of 18 QTL on 7B consisted of three for mixograph peak time, one for mixograph peak height, three for mixograph peak width, two for mixograph right slope, one for Kw, three for Gpc, one for Gac, and four for flour yellowness. Markers *Glu-B1*, *Xwmc182a* and *Xwmc182b* are representative of these chromosomal regions of interest on 1B, 6B, and 7B, respectively. They were observed to be associated with Gpc in one to three environments in previous studies using CO940610/Platte mapping populations. CO940610 contributed the allele for higher Gpc at *Glu-B1* and *Xwmc182a*,

while the Platte allele conferred the higher trait value at *Xwmc182b* as seen in Table 13 (El-Feki 2010).

Marker-assisted selection (MAS) has been successfully applied to practical breeding in wheat. Vishwakarma et al. (2014) used marker-assisted backcrossing (MABC) to improve Gpc in the wheat cultivar HUW468 by selecting for the appropriate allele at marker *Xucw108*, which is linked to the high-value allele of the *Gpc-B1* gene on chromosome 6B. The result showed significantly higher Gpc in improved lines by MABC, and no yield penalty. Kumar et al. (2011) improved Gpc of seven MAS-derived progenies by 14.8 to 17.9% compared to their recipient parents using MABC. The authors suggested that combining MAS and phenotypic selection is useful to improve wheat genotypes for high Gpc without yield loss.

Before the chromosome regions detected by (El-Feki 2010) are employed in breeding programs, they should be validated by analysis of a RIL population (Chapter 2 of this dissertation) or by MAS. The hypothesis of this study was that the Gpc of the MABC line with allelic combination PL-PL-CO was the lowest and CO-CO-PL was the highest of the eight possible allelic combinations.

Therefore, to validate the previous results, the objectives of this study were to (1) transfer the alleles at three loci (*Glu-B1*, *Xwmc182a*, and *Xwmc182b*) from CO940610 into the genetic background of Platte through marker-assisted backcrossing (MABC); (2) evaluate the population of 35 selected BC_3F_2 lines (MABC lines), grown in Fort Collins and Greeley irrigated, and Greeley water deficit environments in the 2012/13 growing season, for Gpc, Gac, Tw, Kw, Kd, and Kh; and (3) conduct marker-trait association analysis for the traits evaluated of the MABC lines.

3.1. MATERIALS AND METHODS

3.1.1. Mapping population

Three markers (*Glu-B1, Xwmc182a,* and *Xwmc182b*) were used to select five doubled haploid (DH) lines from the population of 185 DH lines (El-Feki et al. 2013), which was produced from the cross of CO940610, as female, and Platte as male (El-Feki et al. 2015). CO940610 is a hard white winter experimental wheat line developed by the Colorado State University Wheat Breeding Program from the cross KS87H22/MW09. Platte is a hard white winter cultivar developed by HybriTech Seed International, Inc. (a former division of Monsanto Co., St. Louis, MO), with pedigree Tesia 79/Chat'S'//'Abilene'.

The three markers *Glu-B1*, *Xwmc182a,* and *Xwmc182b* are located on chromosomes 1B, 6B, and 7B, respectively. Based on a previous study, these markers were associated with grain protein concentration (Gpc) in one to three environments (El-Feki 2010). CO940610 contributed the higher value allele for Gpc at *Glu-B1* and *Xwmc182a,* while the Platte allele conferred the higher trait value at *Xwmc182b* as seen in Table 13 (El-Feki 2010).

Table 13. Marker associations with grain protein concentration in the CO940610/Platte DH population in three environments (El-Feki 2010).

		08FW[†]			08FD			09GD		
Loci	Chr.	a[‡]	LOD[††]	R^2 (%)[¶]	a	LOD	R^2 (%)	a	LOD	R^2 (%)
Glu-B1	1B	-1.3	3.03	4.8						
Xwmc182a	6B	-1.6	4.18	7.0				-1.4	2.73	4.5
Xwmc182b	7B	1.6	3.95	6.7	2.2	4.75	9.2	1.6	3.52	5.8

[†] 08FD, Fort Collins Dry 2007/08; 08FW, Fort Collins Dry 2007/08; 09GD, Greeley Dry 2008/09.

[‡] average additive effect (g kg^{-1}), with a positive value indicating an increasing effect of the Platte allele and a negative value indicating an increasing effect of the CO940610 allele.

[††] logarithm (base 10) of odds

[¶] percentage of phenotypic variation explained by the marker.

Five selected DH lines were PCODH19, PCODH41, PCODH52, PCODH82 and PCODH91. The first four of these lines had genotype CO-CO-CO at the loci *Glu-B1*, *Xwmc182a*, and *Xwmc182b*, respectively, while PCODH91 had the alleles PL-CO-CO (Table 14).

Table 14. Allelic constitution of five selected CO940610/Platte DH lines used for backcrossing to Platte as the recurrent parent.

	Selected CO940610/Platte DH lines				
Loci	PCODH19	PCODH41	PCODH52	PCODH82	PCODH91
Glu-B1	CO†	CO	CO	CO	PL
Xwmc182a	CO	CO	CO	CO	CO
Xwmc182b	CO	CO	CO	CO	CO

† CO and Platte indicate selected allele.

Crossing the selected DH lines and Platte produced F_1 plants. Subsequently, heterozygous plants at three loci of interest selected using MAS were backcrossed three times to Platte, obtaining the BC_3F_1 plants (Figure 9). Platte was used as the female in the first crosses (Platte x DHx) and the last backcrosses (Platte x BC_2F_1). An exception was that cross PCODH91 and Platte always produced homozygous genotype (PL-PL) at *Glu-B1* locus.

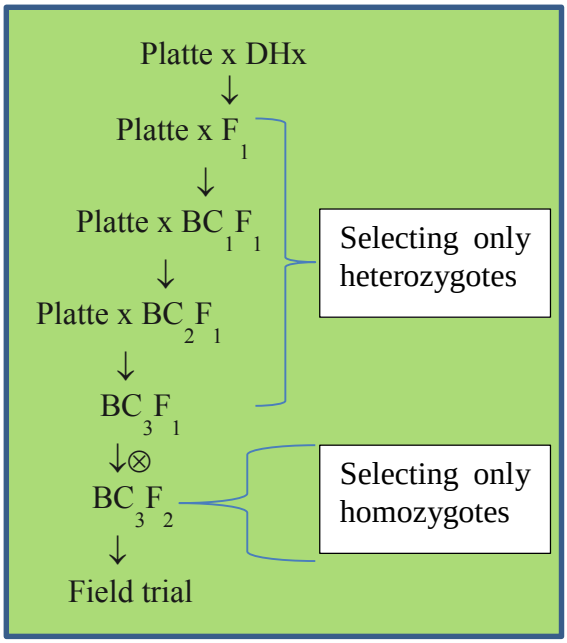


Figure 9. Scheme of selection and developing backcross populations.

The BC_3F_1 plants, which are expected to have an average of 6.25% of their genomic composition from CO940610 and 93.75% from Platte, were self-pollinated to produce the BC_3F_2 plants. Only homozygous BC_3F_2 plants at the loci of interest were selected using MAS. The selected

homozygous BC_3F_2 plants at three loci of interest could be classified into one of eight possible allelic combinations (Table 15).

Table 15. Eight combinations of alleles at the loci *Glu-B1, Xwmc182a, and Xwmc182b* on chromosomes 1B, 6B, and 7B, respectively.

		Selected alleles		
Hypothesis	Genotype class	*Glu-B1*	*Xwmc182a*	*Xwmc182b*
	PL-PL-PL	Platte	Platte	Platte
Lowest Gpc	PL-PL-CO	Platte	Platte	CO940610
	PL-CO-PL	Platte	CO940610	Platte
	PL-CO-CO	Platte	CO940610	CO940610
	CO-PL-PL	CO940610	Platte	Platte
	CO-PL-CO	CO940610	Platte	CO940610
Highest Gpc	CO-CO-PL	CO940610	CO940610	Platte
	CO-CO-CO	CO940610	CO940610	CO940610

The five selected DH lines mentioned above generated five families of BC_3F_2, each with eight possible homozygous allelic combinations (Table 15). Therefore, the five families could result in maximum of 40 different genotype combinations. Five of these 40 genotype combinations, however, did not survive or were missed during the selection process. Plants with the other genotype combinations were selected, grown in the CSU greenhouse for seed increase, and the resulting seeds were used for field trials. The 35 selected plants and their genotypes are listed in Table 16.

The whole process of genotyping, selection, and seed increase was conducted in the CSU greenhouse and in the Byrne plant genetics laboratory.

Table 16. The 35 selected plants and their genotypes of the CO940610/Platte DH (BC_3) backcross population.

Entry no.	Developed from COP DH line	Selected allele			
		Glu-B1	***Xwmc182a***	***Xwmc182b***	**Genotype class**
1	19	Platte	Platte	Platte	PL-PL-PL
2	19	Platte	Platte	CO940610	PL-PL-CO
3	19	Platte	CO940610	Platte	PL-CO-PL
4	19	Platte	CO940610	CO940610	PL-CO-CO
5	19	CO940610	Platte	Platte	CO-PL-PL
6	19	CO940610	Platte	CO940610	CO-PL-CO
7	19	CO940610	CO940610	Platte	CO-CO-PL
8	19	CO940610	CO940610	CO940610	CO-CO-CO
9	41	Platte	Platte	Platte	PL-PL-PL
10	41	Platte	Platte	CO940610	PL-PL-CO
11	41	Platte	CO940610	Platte	PL-CO-PL
12	41	Platte	CO940610	CO940610	PL-CO-CO
13	41	CO940610	Platte	Platte	CO-PL-PL
14	41	CO940610	Platte	CO940610	CO-PL-CO
15	41	CO940610	CO940610	Platte	CO-CO-PL
16	41	CO940610	CO940610	CO940610	CO-CO-CO
19	52	CO940610	Platte	Platte	CO-PL-PL
20	52	CO940610	Platte	CO940610	CO-PL-CO
21	52	CO940610	CO940610	Platte	CO-CO-PL
22	52	CO940610	CO940610	CO940610	CO-CO-CO
23	82	Platte	Platte	Platte	PL-PL-PL
24	82	Platte	Platte	CO940610	PL-PL-CO
25	82	Platte	CO940610	CO940610	PL-CO-CO
26	82	CO940610	Platte	Platte	CO-PL-PL
27	82	CO940610	Platte	CO940610	CO-PL-CO
28	82	CO940610	CO940610	Platte	CO-CO-PL
29	82	CO940610	CO940610	CO940610	CO-CO-CO
30	91	Platte	Platte	Platte	PL-PL-PL
31	91	Platte	Platte	CO940610	PL-PL-CO
32	91	Platte	CO940610	Platte	PL-CO-PL
33	91	Platte	CO940610	CO940610	PL-CO-CO
34	91	CO940610	Platte	Platte	CO-PL-PL
35	91	CO940610	Platte	CO940610	CO-PL-CO
36	91	CO940610	CO940610	Platte	CO-CO-PL
37	91	CO940610	CO940610	CO940610	CO-CO-CO

3.1.2. Genotyping

Germination and vernalization: Each wheat seed was sown separately in ProMix Biofungicide soil media, manufactured by Premier Horticultural Inc. (Quakertown, PA), in a cell of a tray in the CSU greenhouse. After seven days, as the seed germinated, the seedlings were moved into a cold room (about 4^{o}C). About 6-8 weeks later, the seedlings were transplanted into a 15-cm diameter pot filled with the ProMix Biofungicide soil in the greenhouse.

Leaf sampling: About two weeks after transplanting, an approximately 8-cm leaf segment was sampled for DNA extraction. Leaf samples were collected into separate labeled tubes, which were placed on ice, then transferred to a -80^{o}C freezer prior to DNA extraction.

DNA extraction: A detailed procedure was described in section 2.1.5 (Chapter 2 of this dissertation). It is summarized in the following steps.

(1) The leaf samples were freeze-dried for about 48 hours in a freeze drier (Freeze Dry System, Labconco ® 7522900).

(2) The freeze-dried leaf samples were ground for 30 to 90 seconds using a shaker machine (Mini-BeadBeater^{TM}, BioSpec Products, Inc.).

(3) DNA was extracted, purified, and diluted using a wheat extraction protocol (Riede and Anderson, 1996).

Marker-assisted selection: Three markers (*Bx7-MAR, Xwmc182a,* and *Xwmc182b*) representing the loci of interest (*Glu-B1*, *Xwmc182a,* and *Xwmc182b*, respectively) were used to select the desired plants. Because there were many samples to screen in each generation, the screening was conducted in a step-wise fashion as described below.

First, all DNA samples from each generation were diluted, then amplified by using marker Bx-7 in a polymerase chain reaction (PCR) to obtain genotypes at *Glu-B1*. Amplification conditions for the PCR reaction were an initial cycle at 95°C for 5 minutes, followed by 37 cycles of 95°C for 30 seconds, 58°C for 30 seconds and 72°C for 1 minute, followed by a final extension at 72°C for 5.15 minutes. Primers used were forward 5'-CCTCAGCATGCAAACATGCAGC-3' and reverse 5'-CTGAAACCTTTGGCCAGTCATGTC-3' (Butow et al. 2004). A 1.4% agarose gel was used to separate the bands, which were stained with ethidium bromide and visualized under ultra-violet (UV) light. Only plants that were heterozygous at *Glu-B1* were selected in this step.

Second, PCR was conducted with *Xwmc182* primers for samples that were heterozygous at *Glu-B1*. Thermal cycling conditions were as follows: 95°C, 5 minutes followed by 4 cycles of 94°C, 30 seconds; annealing temperature step-downs every 1 cycle of 1°C (from 56°C to 51°C), 40 seconds; 72°C, 1 minute. The annealing temperature for the final 34 cycles was 51°C with denaturation and extension times as above. Primers used were forward 5'-GTATCTCACGAGCATAACACAA-3' and reverse 5'- GAAAGTGTATGGATCATTAGGC-3' (Somers et al., 2004). With these primers, two loci were amplified, designated *Xwmc182a* and *Xwmc182b*, each of which produced a DNA band for each of the parents. Polyacrylamide gels (PAGE) were used to separate the bands, which were then silver-stained as described in El-Feki et al. (2013). Only plants that were heterozygous at both *Xwmc182a* and *Xwmc182a* were selected at the second step.

3.1.2. Experimental design and trial management

The population consisted of 35 selected BC_3F_2 lines, two parents, and three checks. The two parents were Platte and CO940610 (El-Feki et al. 2015). The check cultivars were Longhorn

(Reg. No. GSTR 9100198) (AgriPro Biosciences Inc. 1998), Hatcher (PI 638512) (Haley et al. 2005) and Ripper (PI 644222) (Haley et al. 2007).

The field trials for the population were conducted in the 2012/13 growing seasons at two Colorado locations, CSU's Agricultural Research, Development, and Education Center (ARDEC) in Fort Collins and the USDA-Agricultural Research Service Limited Irrigation Research Farm (LIRF) in Greeley.

The experiments were designed as randomized complete blocks with two replicates in each treatment at LIRF and four replicates at ARDEC. The entries in each block were randomized using Excel with the function "=RAND()". Experimental units were two-row plots 1.52 m long and 0.51 m wide. Space between rows was 23 cm, and space between plots was 28 cm. Seed density was approximately 1,728,000 seeds ha^{-1}. Fertilization and weed control were typical of practices for winter wheat in the area, and no disease or insect pest control was needed.

ARDEC

The trial was located at 40.652 N, 104.996 W, and elevation of 1558 m near Fort Collins, Colorado. The soil type is Fort Collins loam, 0 to 1 percent slopes (http://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm).

The trial was conducted only in fully irrigated condition (ARDEC Wet). Water was regularly applied using a linear overhead sprinkler system until the wheat ripened, for a total of 216 mm of irrigation water (ARDEC record) plus 124 mm of precipitation (http://www.coagmet.colostate.edu/) over the growing season (Table 36 and 37).

LIRF

Side-by-side experiments were located at 40.449 N, 104.638 W, and elevation of 1427 m in Greeley, Colorado. The soil type was Nunn clay loam, 0 to 1 percent slopes (http://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm).

Water was applied using a drip irrigation system. Irrigation was supplied to both treatments beginning during heading and grain filling (Zadoks et al. 1974). The wet treatment (LIRF Wet received 114 mm of irrigation water and 105 mm of precipitation (http://www.coagmet.colostate.edu/), for a total of 219 mm of water. The dry treatment (LIRF Dry) received 38 mm of irrigation water and 105 mm of precipitation (http://www.coagmet.colostate.edu/), for a total of 143 mm of water (Table 36 and 37).

3.1.3. Phenotypic evaluation

Plant height

Height of five plants from the soil to the tip of the spike excluding awns was manually measured in each plot, approximately two weeks before harvest.

Grain protein concentration, grain ash concentration, and test weight

These traits were measured by the Near Infrared Reflectance System (NIRs) at the CSU Wheat Quality Laboratory. A Foss-Tecator NIR systems Model 6500 instrument (Foss North America, Eden Prairie, MN) was used to measure these traits in reflectance mode with a wavelength of 450–2498 nm (El-Feki et al. 2013).

Wheat kernels were cleaned before measurement. A sample of approximately 100 g was added into a transport module in a rectangular quartz cup (NIRSystems, part 0IH–0379). The estimates of the traits were obtained with calibrations from the manufacturer for Gpc and Gac, and a

calibration developed in the CSU Wheat Quality Laboratory for Tw. The output measured by the instrument was recorded on the computer. The units of traits were converted to g kg^{-1} for Gpc and Gac, and kg hL^{-1} for Tw.

Single kernel characteristics

Three single kernel traits evaluated in this study were kernel weight (Kw), kernel diameter (Kd) and kernel hardness (Kh). These traits were measured by the Single Kernel Characterization System (SKCS, model 4100, Perten Instruments, Springfield, IL) in the CSU Wheat Quality Laboratory.

Wheat kernels were carefully cleaned. Broken kernels, weed seeds, metal staples, and any other foreign materials were manually removed. The cleaned wheat kernels were placed in the access hopper of the SKCS instrument, then the hopper was placed in the SKCS instrument for kernel measurement. The instrument analyzed 100 kernels individually, measuring Kw, Kd and Kh. Kernel weight was measured by a load cell, and reported in mg. Kernel diameter was measured by electrical current, and expressed in mm. Kernel hardness was analyzed by pressure force, and expressed as an index scale of 0 (extremely soft) to 100 (extremely hard). These results were recorded on the computer.

3.1.4. Statistical analysis

The raw (original) phenotypic data were plotted in heat maps using Excel 2010 and in frequency distribution curves using JMP Pro 11 to identify the outliers and remove obvious mistakes. Additional inspection of the data was also implemented by visualizing plots of residual distributions in SAS 9.3 (SAS Institute Inc., Cary, NC, 2010).

After removing the outliers, the original phenotypic data were analyzed with six different spatial models to adjust for spatial variations using PROC MIXED in SAS 9.3. The model with the lowest Akaike's information criterion (AIC) value was used for estimating best linear unbiased predictions (BLUPs). These data were used for the following analyses: plotting normal distributions, calculating trait means using Excel 2010 and trait correlations (using PROC CORR of SAS 9.3), determining the significance of marker-trait associations (PROC GLM), and calculating means of genotype classes (PROC GLM, Lsmeans option). The original phenotypic data, however, were used for calculating broad-sense heritability estimates with PROC GLM and Excel. Detailed descriptions of statistical analyses are provided in Chapter 2.

3.2. RESULTS

3.2.1. Trait distribution and means

Distribution

Sixteen of 18 trait-environment combinations conformed to expectations for a normally distributed population according to the Shapiro-Wilk test at $P>0.05$ (Figure 14). For two trait-environments that deviated, the distributions were approximately normal based on a visual evaluation and examination of residuals.

Trait means

Plant height in ARDEC Wet was about 9 cm higher than in LIRF Wet and 11 cm higher than in LIRF Dry (Table 17). This likely reflects the difference in available water and temperature at these three environments. Total available water in the ARDEC Wet treatment was 55% higher than in the LIRF Wet treatment and 138% higher than in the LIRF Dry treatment. There was,

however, only a slight difference in Ph between LIRF wet and LIRF Dry, indicating that the irrigation differential (Wet 53% higher than Dry), which began at heading, did not have much impact on plant growth in that location.

Grain protein concentration in ARDEC Wet, 165.23 g kg^{-1}, and in LIRF Wet, 160.81 g kg^{-1}, were significantly (P<0.0001) lower than that in LIRF Dry, 172.50 g kg^{-1}. The difference of Gpc between the LIRF Dry and Wet treatments indicated that reduced soil moisture was associated with increased Gpc.

Test weight (73.97 kg hL^{-1}), Kw (24.36 mg) and Kh (65.32 SKCS hardness units) in LIRF Dry were significantly (P<0.0001) lower than those in either LIRF Wet or ARDEC Wet. This indicates that greater drought stress was associated with reduced grain filling and consequently smaller, lighter, and softer kernels.

Table 17. Means, standard errors, and ranges for traits of the CO940610/Platte backcross population (n=35) for the ARDEC wet treatment and the LIRF wet and dry treatments in the 2012/13 growing season.

Env.	ARDEC Wet [1]		LIRF Wet [2]		LIRF Dry [3]		*P*-value‡		
Traits†	Means ± SE	Range	Means ± SE	Range	Means ± SE	Range	[1]&[2]	[1]&[3]	[2]&[3]
Ph (cm)	59.67 ± 0.13	58.22 - 61.71	50.47 ± 0.18	48.54 - 52.97	48.42 ± 0.34	43.61 - 51.84	<0.001	<0.001	<0.001
Gpc (g kg^{-1})	165.23 ± 0.23	162.65 - 169.71	160.81 ± 0.12	159.31 - 162.74	172.50 ± 0.50	166.24 - 176.69	<0.001	<0.001	<0.001
Gac (g kg^{-1})	16.30 ± 0.05	15.84 - 16.93	15.97 ± 0.04	15.54 - 16.40	16.35 ± 0.06	15.62 - 17.37	<0.001	0.546	<0.001
Tw (kg hL^{-1})	74.86 ± 0.08	73.95 - 75.77	75.48 ± 0.06	74.56 - 76.09	73.97 ± 0.10	72.40 - 74.93	<0.001	<0.001	<0.001
Kw (mg)	26.41 ± 0.19	24.30 - 28.68	30.50 ± 0.15	29.10 - 32.53	24.36 ± 0.08	23.44 - 25.69	<0.001	<0.001	<0.001
Kd (mm)	2.54 ± 0.01	2.46 - 2.61	2.74 ± 0.00	2.72 - 2.76	2.53 ± 0.00	2.47 - 2.61	<0.001	0.059	<0.001
Kh (0-100)	70.55 ± 0.47	64.09 - 77.79	70.33 ± 0.44	65.71 - 75.07	65.32 ± 0.15	63.54 - 67.65	0.732	<0.001	<0.001

† Ph, plant height; Gpc, grain protein concentration at 12% moisture; Gac, grain ash concentration; Tw, test weight; Kw, kernel weight; Kd, kernel diameter; Kh, kernel hardness.

‡ Probability level for difference between pairwise combinations of environments.

3.2.2. Correlation among traits

Pearson correlation coefficients among seven traits in three environments are shown in Table 18 and 19. Plant height was negatively correlated with Gpc in two of the three environments. The correlation coefficients were r=-0.35 (P<0.05) and -0.37 (P<0.05) at LIRF Dry and ARDEC Wet, respectively. The correlation between Ph and Gpc was not significant in LIRF Wet. Plant height, however, was positively correlated with Kw. The correlation coefficients were 0.47 (P<0.01), 0.44 (P<0.01) and 0.62 (P<0.001) at LIRF Wet, LIRF Dry and ARDEC Wet, respectively.

Test weight, Kw, and Kd were significantly correlated with each other. The correlation coefficients between Tw and Kw were 0.65 (P<0.001), 0.47 (P<0.01), and 0.41 (P<0.05) in LIRF Wet, LIRF Dry and ARDEC Wet, respectively. The correlations between Tw and Kd were 0.56 (P<0.001) and 0.47 (P<0.01) in LIRF Wet and LIRF Dry, respectively. In particular, Kw and Kd were hightly correlated. Their correlation coefficients ranged from 0.85 to 0.93 (P<0.001) in three environments.

Table 18. Pearson correlation coefficients among traits of the CO940610/Platte backcross population (n=35) at LIRF, Greeley, CO in the 2012/13 growing season. Correlations for the wet treatment are above the diagonal and those for the dry treatment are below the diagonal.

Traits†	Ph	Gpc	Gac	Tw	Kw	Kd	Kh
Ph		-0.07	-0.11	0.01	0.47**	0.36*	-0.33
Gpc	-0.35*		0.32	-0.14	-0.08	-0.16	-0.14
Gac	-0.22	0.31		0.19	0.22	0.02	0.00
Tw	0.07	-0.33	-0.30		0.65***	0.56***	-0.09
Kw	0.44**	-0.48**	-0.10	0.47**		0.85***	-0.23
Kd	0.31	-0.50**	-0.12	0.47**	0.91***		-0.02
Kh	-0.14	-0.31	-0.12	0.18	-0.04	-0.06	

*, **, *** indicates significance at the 0.05, 0.01, and 0.001 levels of probability, respectively.

† Ph, plant height; Gpc, grain protein concentration at 12% moisture level; Gac, grain ash concentration; Tw, test weight; Kw, kernel weight; Kd, kernel diameter; Kh, kernel hardness.

Table 19. Pearson correlation coefficients among traits of the CO940610/Platte backcross population (n=35) in the ARDEC wet treatment, Fort Collins, CO in the 2012/13 growing season.

Traits†	Gpc	Gac	Tw	Kw	Kd	Kh
PH	-0.37*	-0.01	0.15	0.62***	0.59***	-0.12
Gpc		0.32	0.04	-0.16	-0.17	-0.25
Gac			-0.09	0.24	0.13	-0.26
Tw				0.41*	0.30	-0.21
Kw					0.93***	-0.44**
Kd						-0.29

*, **, *** indicates significance at the 0.05, 0.01, and 0.001 levels of probability, respectively.

† Ph, plant height; Gpc, grain protein concentration at 12% moisture level; Gac, grain ash concentration; Tw, test weight; Kw, kernel weight; Kd, kernel diameter; Kh, kernel hardness.

3.2.3. Heritability

The estimated heritability values for each measured phenotypic trait in each environment and combined across environments are presented in Table 20. Grain protein concentration had low heritability estimates. Its heritability estimates were 0.19, 0.33 and 0.47 in LIRF Wet, ARDEC Wet and LIRF Dry, respectively.

The heritability estimates of Gac, Tw, Kw, Kd and Kh varied largely from environment to environment. For example, the heritability estimate of Kh was 0.80 and 0.81 in ARDEC Wet and LIRF Wet, respectively, while it was 0.40 in LIRF Dry.

Across three environments, heritability estimates were low (<0.50) for Gpc, moderate (0.50-0.75) for Ph, Gac, Tw and Kd, but high (>0.75) for Kw and Kh.

Table 20. Broad-sense heritability estimates (H^2) and 90% confidence intervals for seven traits of the CO940610/Platte BC_3F_2 backcross population in the ARDEC wet, LIRF wet, and LIRF dry environments in the 2012/13 growing season.

	ARDEC Wet		LIRF Wet		LIRF Dry		Three environments	
Trait†	H^2	90% confidence interval‡	H^2	90% confidence interval	H^2	90% confidence interval	H^2	90% confidence interval
Ph	0.40	0.01 - 0.62	0.28	-0.28 - 0.59	0.42	-0.02 - 0.67	0.59	0.34 - 0.73
Gpc	0.33	-0.11 - 0.57	0.19	-0.44 - 0.54	0.47	0.06 - 0.70	0.41	0.01 - 0.63
Gac	0.57	0.29 - 0.73	0.26	-0.31 - 0.58	0.56	0.21 - 0.75	0.67	0.46 - 0.78
Tw	0.66	0.44 - 0.78	0.62	0.33 - 0.79	0.39	-0.08 - 0.66	0.58	0.29 - 0.74
Kw	0.77	0.62 - 0.85	0.38	-0.11 - 0.65	0.25	-0.32 - 0.58	0.76	0.60 - 0.84
Kd	0.65	0.42 - 0.78	0.20	-0.43 - 0.55	0.22	-0.39 - 0.56	0.67	0.47 - 0.78
Kh	0.80	0.66 - 0.87	0.81	0.65 - 0.89	0.40	-0.07 - 0.66	0.85	0.75 - 0.90

† Ph, plant height; Gpc, grain protein concentration at 12% moisture level; Gac, grain ash concentration; Tw, test weight; Kw, kernel weight; Kd, kernel diameter; Kh, kernel hardness.

‡ Based on Knapp et al. (1985).

3.2.4. Detection of significant marker-trait associations

The significant marker-trait associations detected are shown in Table 21. Plant height was associated with marker *Xwmc182b* in both ARDEC wet and LIRF wet treatments with an additive effect of -0.40 and -0.37 cm, respectively, indicating that CO940610 contributed the taller plant height allele.

All three markers of interest were associated with Gpc. *Glu-B1* and *Xwmc182b* were associated with Gpc in the LIRF dry and ARDEC wet treatments with positive additive effects of 1.07 and 0.60 g kg^{-1}, respectively, indicating that Platte contributed alleles for higher Gpc at these loci. *Xwmc182a* was associated with Gpc in the ARDEC wet and LIRF wet treatments, with negative additive effects of -0.57 and -0.28 g kg^{-1}, indicating that CO940610 contributed alleles for higher Gpc at those loci.

Grain ash concentration was also associated with all three tested markers. The negative additive effect of *Glu-B1* for Gac was -0.11 g kg^{-1} in the ARDEC wet treatment. *Xwmc182a* had additive effects of -0.08 and -0.14 g kg^{-1} in the LIRF wet and dry treatments, respectively. This indicates that the CO940610 allele contributed higher Gac at both of these loci. On the other hand, a positive additive effect (0.12) was detected for *Xwmc182b*, indicating that Platte contributed higher Gac at this locus. In addition, there were significant epistatic interactions of *Glu-B1* with *Xwmc182a* and *Xwmc182b* in the ARDEC wet treatment.

Marker *Glu-B1* was significantly associated with kernel weight, with a negative effect of -0.29 and -0.21 mg in the LIRF wet and dry treatments, respectively. This marker was also associated with kernel diameter with a negative effect of -0.01 mm in the LIRF dry treatment. This evidence indicates that CO940610 contributed greater Kw and Kd at this locus. In addition, there was

evidence of an epistatic interaction of *Glu-B1* with *Xwmc182a* for kernel diameter in the LIRF dry environment.

Table 21. Significance of loci detected with analysis of variance for single factors or digenic epistatic interactions for seven traits in the CO940610/Platte BC_3F_2 population in the 2012/13 growing season.

Traits/Markers	Environments	a†	R^2(%)‡	*P*-value
Plant height (cm)				
Xwmc182b	ARDEC wet	-0.40	28.6	0.0009
Xwmc182b	LIRF wet	-0.37	12.8	0.0352
Grain protein concentration (g kg^{-1})				
Glu-B1	LIRF dry	1.07	13.2	0.0317
Xwmc182a	ARDEC wet	-0.57	17.7	0.0117
Xwmc182a	LIRF wet	-0.28	14.7	0.0229
Xwmc182b	ARDEC wet	0.60	19.2	0.0085
Grain ash concentration (g kg^{-1})				
Glu-B1	ARDEC wet	-0.11	16.0	0.0174
Xwmc182a	LIRF wet	-0.08	11.9	0.0427
Xwmc182a	LIRF dry	-0.14	15.9	0.0177
Xwmc182b	ARDEC wet	0.12	21.2	0.0053
*Glu-B1*Xwmc182b*	ARDEC wet			0.0369
Kernel weight (mg)				
Glu-B1	LIRF wet	-0.29	11.8	0.0432
Glu-B1	LIRF dry	-0.21	19.0	0.0088
Kernel diameter (mm)				
Glu-B1	LIRF dry	-0.01	15.8	0.0180
*Glu-B1*Xwmc182a*	LIRF dry			0.0376

† Average additive effect: positive values indicate an increasing effect of Platte alleles and negative values indicate an increasing effect of CO940610 alleles.

‡ Percent of phenotypic variation explained by the marker.

3.2.5. Epistatic interactions

When pairwise interactions were examined, significant ($P<0.05$) epistasis was detected between markers *Glu-B1* and *Xwmc182a* (Table 22) and between *Glu-B1* and *Xwmc182b* (Table 23). The interactions were between *Glu-B1* and *Xwmc182a* loci for kernel diameter in the LIRF dry

treatment and between *Glu-B1* and *Xwmc182b* loci for grain ash concentration in the ARDEC wet treatment. The Platte allele at *Xwmc182a* provides an increase in kernel diameter only when the CO940610 allele is present at the *Glu-B1* locus (Table 22). In contrast, when the Platte allele at *Glu-B1* and CO940610 allele at *Xwmc182b* were present, they caused a decrease in grain ash concentration (Table 23).

Table 22. Genotype class means for the epistatic interaction of loci *Glu-B1* and *Xwmc182a* in the CO940610/Platte BC_3F_2 population.

Trait	Genotype class			
(Environment)	*Glu-B1*	*Xwmc182a*	Mean ± SE	
Kernel diameter (mm)	CO940610	CO940610	2.52 ± 0.01	b†
(LIRF dry)	CO940610	Platte	2.55 ± 0.01	a
	Platte	CO940610	2.52 ± 0.01	b
	Platte	Platte	2.51 ± 0.01	b

† Means followed by the same letter do not differ significantly at the 0.05 probability level.

Table 23. Genotype class means for the epistatic interaction of loci *Glu-B1* and *Xwmc182b* in the CO940610/Platte BC_3F_2 population.

Trait	Genotype class			
(Environment)	*Glu-B1*	*Xwmc182b*	Mean ± SE	
Grain ash concentration (g kg-1)	CO940610	CO940610	16.35 ± 0.07	a†
(ARDEC wet)	CO940610	Platte	16.45 ± 0.07	a
	Platte	CO940610	16.01 ± 0.07	b
	Platte	Platte	16.41 ± 0.08	a

† Means followed by the same letter do not differ significantly at the 0.05 probability level.

3.2.6. Trait mean comparisons among eight genotype classes

Means for seven traits are shown in Tables 24, 25, and 26. For the most part, plant heights among the eight genotype classes were not significantly different in any of the three

environments. The only exception was that plant height of genotype class ‘PL-CO-CO’ (60.30 cm) in the ARDEC wet environment was significantly higher than that of ‘CO-PL-PL’ (59.27 cm) and ‘CO-CO-PL’ (59.09 cm).

In ARDEC Wet, Gpc of genotype class ‘CO-CO-PL’ (166.53 g kg^{-1}) was significantly higher than that of ‘PL-PL-CO’ (164.69 g kg^{-1}). This confirmed the hypothesis that Gpc is the highest for the combination ‘CO-CO-PL’ and lowest for ‘PL-PL-CO’. This expected result, however, was not repeated in LIRF Wet and LIRF Dry, where Gpc of these genotype classes were not significantly different.

Test weight showed mostly non-significant differences among genotype classes in all three environments, with the exception that test weight of genotype class ‘PL-CO-CO’ (75.28 kg hL^{-1}) was higher than that of ‘CO-PL-CO’ (74.60 kg hL^{-1}) in ARDEC Wet.

Kernel weight, Kd, and Kh showed mostly non-significant differences among genotype classes in all three environments. Each of these traits had one exception in LIRF Dry. Kernel weights of genotype classes ‘CO-PL-PL’ (24.80 mg) and ‘CO-PL-CO’ (24.73 mg) in LIRF Dry were higher than that of ‘PL-PL-PL’ (23.92 mg), ‘PL-PL-CO’ (24.16 mg), and ‘PL-CO-PL’ (24.00 mg). Kernel diameter followed the same pattern as Kw. Kernel hardness of genotype class ‘CO-PL-PL’ (66.47) was higher than that of ‘PL-PL-PL’ (65.03), ‘PL-PL-CO’ (65.24), ‘PL-CO-CO’ (64.57), and ‘CO-PL-CO’ (65.11).

Table 24. Least squares means of seven traits in each genotype class of the CO940610/Platte BC_3F_2 population in the ARDEC wet treatment in the 2012/13 growing season. Genotype classes are defined in Table 15.

Genotype classes	Plant height (cm)		Grain protein concentration (g kg^{-1})		Grain ash concentration (g kg^{-1})		Test weight (kg hL^{-1})		Kernel weight (mg)		Kernel diameter (mm)		Kernel hardness (0-100)	
PL-PL-PL	59.34	abc†	165.43	abc	16.23	abc	74.89	ab	26.01	a	2.53	a	68.96	a
PL-PL-CO	59.84	abc	164.69	a	15.97	a	74.91	ab	26.28	a	2.54	a	71.29	a
PL-CO-PL	59.32	abc	166.93	c	16.66	d	74.93	ab	26.25	a	2.54	a	69.55	a
PL-CO-CO	60.30	c	165.19	abc	16.05	ab	75.28	b	26.33	a	2.54	a	71.74	a
CO-PL-PL	59.27	ab	164.72	ab	16.46	cd	74.79	ab	26.68	a	2.57	a	72.53	a
CO-PL-CO	60.13	bc	164.03	a	16.43	cd	74.60	a	26.69	a	2.54	a	70.55	a
CO-CO-PL	59.09	a	166.53	c	16.44	cd	74.82	ab	25.78	a	2.52	a	69.52	a
CO-CO-CO	59.94	abc	164.94	ab	16.27	bc	74.78	ab	27.13	a	2.56	a	70.19	a

† Means followed by the same letter are not significantly different at $P<0.05$.

Table 25. Least squares means of seven traits in each genotype class of the CO940610/Platte BC_3F_2 population in the LIRF wet treatment in the 2013 growing season. Genotype classes are defined in Table 15.

Genotype classes	Plant height (cm)		Grain protein concentration (g kg^{-1})		Grain ash concentration (g kg^{-1})		Test weight (kg hL^{-1})		Kernel weight (mg)		Kernel diameter (mm)		Kernel hardness (0-100)	
PL-PL-PL	50.19	a†	161.06	bc	15.95	ab	75.24	a	29.89	a	2.73	a	69.70	a
PL-PL-CO	50.62	a	160.79	abc	15.91	ab	75.50	a	30.39	a	2.73	a	70.02	a
PL-CO-PL	49.91	a	160.77	abc	16.02	ab	75.63	a	30.06	a	2.73	a	71.80	a
PL-CO-CO	50.47	a	161.20	c	15.95	ab	75.28	a	30.29	a	2.74	a	70.32	a
CO-PL-PL	50.12	a	160.19	ab	15.74	a	75.66	a	30.65	a	2.74	a	70.87	a
CO-PL-CO	51.05	a	160.15	a	15.94	ab	75.54	a	30.90	a	2.74	a	71.19	a
CO-CO-PL	50.00	a	160.78	abc	16.12	b	75.56	a	30.43	a	2.73	a	68.71	a
CO-CO-CO	51.02	a	161.53	c	16.07	b	75.44	a	31.08	a	2.74	a	70.58	a

† Means followed by the same letter are not significantly different at $P<0.05$.

Table 26. Least squares means of seven traits in each genotype class of the CO940610/Platte BC_3F_2 population in the LIRF dry treatment in the 2013 growing season. Genotype classes are defined in Table 15.

Genotype classes	Plant height (cm)		Grain protein concentration (g kg^{-1})		Grain ash concentration (g kg^{-1})		Test weight (kg hL^{-1})		Kernel weight (mg)		Kernel diameter (mm)		Kernel hardness (0-100)	
PL-PL-PL	48.34	a†	173.31	b	16.18	ab	73.99	a	23.92	a	2.51	a	65.03	a
PL-PL-CO	49.20	a	173.50	b	16.31	ab	73.71	a	24.16	a	2.51	a	65.24	a
PL-CO-PL	47.01	a	173.98	b	16.56	ab	74.18	a	24.00	a	2.51	a	65.42	ab
PL-CO-CO	48.16	a	173.97	b	16.38	ab	74.05	a	24.42	ab	2.53	ab	64.57	a
CO-PL-PL	48.76	a	168.92	a	16.03	a	74.35	a	24.80	b	2.55	b	66.47	b
CO-PL-CO	49.28	a	172.07	ab	16.29	ab	73.77	a	24.73	b	2.55	b	65.11	a
CO-CO-PL	47.09	a	171.60	ab	16.49	ab	73.92	a	24.25	ab	2.52	ab	65.47	ab
CO-CO-CO	48.96	a	172.97	b	16.53	b	73.95	a	24.49	ab	2.53	ab	65.34	ab

† Means followed by the same letter are not significantly different at $P<0.05$.

3.3. DISCUSSION

3.3.1. Trait means, correlation and heritability

Trait means

About 1% and 19% reduction of Ph occurred in LIRF Dry compared to LIRF Wet and ARDEC Wet, respectively. El-Feki (2010) found Ph was reduced by 14.5 - 18.2% under drought conditions in the DH population. Amiri et al. (2013) reported that terminal drought stress reduced Ph by 1.23%. Mirbahar et al. (2009) reported that water stress significantly reduced Ph. Plant height is associated with the movement of carbohydrates in the plant, especially under water stress conditions (Blum et al. 1989). Drought stress restricts dry mater accumulation and nutrient access in plants (Ahmadi & Baker 2001).

Grain protein concentration and Gac in the water deficit treatment were significantly (P<0.001) higher than those traits in the watered treatments, confirming that Gpc and Gac increase under drought conditions. Water stress during grain filling is known to decrease sucrose and starch content of wheat grain relative to protein (Ahmadi & Baker 2001; Jenner & Rathjen 1975), thus, resulting in an increase in final Gpc and Gac (Weightman et al. 2008). Drought and heat stresses are two events that often occur simultaneously in the field (Aprile et al. 2013). Triboï et al. (2003) reported that the rate of N (indication of protein) accumulation per day increased with an increase in post-anthesis temperature.

Kernel weight under water deficit was significantly (P<0.001) reduced compared to full irrigation. Dry mater and grain moisture are major factors influencing kernel weight. Under terminal drought stress, grain accumulates less starch and sucrose, thus, causing lighter grain.

Correlation among traits

Kernel weight and Kd were positively correlated with each other. This agreed with El-Feki et al. (2013). As a grain accumulates more dry matter and/or has higher grain moisture content, it is heavier and larger. In contrast, Gpc was negatively correlated with Kw and Kd. This conformed to previous reports (Al-Saleh & Brennan 2012; Fjell et al. 1985). Grain dry matter (mostly starch) determines grain weight and size. Higher amounts of starch and/or water in grain dilute Gpc. This suggests that simultaneous improvement of Gpc and Kw would be challenging.

Heritability

Grain protein concentration had low heritability estimates, ranging from 0.19 to 0.47. Grain protein concentration is a complex character, controlled by many genes and significantly influenced by environmental conditions. Clarke et al. (2009) reported that heritability estimates for Gpc ranged from 0.43 to 0.84, and suggested that heritability estimates are influenced by the number of replications, years and locations tested, and by the genetics of the populations. Barnard et al. (2002) estimated broad-sense heritability of Gpc at 0.57. In the DH, heritability estimates for Gpc in four environments ranged from 0.52 to 0.61 (El-Feki 2010).

Kernel weight and Kh had high (>0.75) heritability, indicating that their variation in this population is largely controlled by genetics. Therefore, it can be easier to make progress in selection for these traits in breeding programs. El-Feki (2010) reported a similar range of heritability estimates for Kw (0.78 – 0.87) and Kh (0.79 – 0.88) in the DH in four environments.

3.3.2. Marker-trait-associations

Only three markers *Glu-B1*, *Xwmc182a*, and *Xwmc182b*, were investigated in the current study, so few significant MTA were detected for the seven traits observed (Table 21).

Marker *Glu-B1* was significantly associated with Gpc, Gac, Kw, and Kd. This indicates either that *Glu-B1* governs multiple traits (pleiotropy) or that there are multiple linked genes in that region that influence those traits. The Kw-*Glu-B1* association was detected in the LIRF Wet and LIRF Dry, indicating this MTA was robust across these two environments. El-Feki et al. (2013) detected QTL for Gac, Tw and Kw in chromosomal region 1B.1, where *Glu-B1* was located.

Markers *Xwmc182a* and *Xwmc182b* also showed evidence either pleiotropy or linkage of multiple genes. *Xwmc182a* was associated with Ph, Gpc, and Gac. The Gpc-*Xwmc182a* association was observed to be robust across environments ARDEC Wet and LIRF Wet. The Gac-*Xwmc182a* association was robust across environments LIRF Wet and LIRF Dry. El-Feki et al. (2013) found QTL for Gpc and Tw on 6B.1, but did not map this marker on it. *Xwmc182b* was also detected in association with Ph, Gpc, and Gac, but only in a single environment. El-Feki et al. (2013) also detected QTL for Gpc and Gac on 7B, where *Xwmc182b* was located.

3.3.3. Combined genotype trait means

As described in section 3.1, CO940610 contributed the higher Gpc allele at *Glu-B1* and *Xwmc182a*, while the Platte allele conferred the higher value at *Xwmc182b* in the study by El-Feki (2010) (Table 13). With that assumption, in the order of *Glu-B1*, *Xwmc182a*, and *Xwmc182b*, a combined genotype CO-CO-PL should have the highest Gpc, while PL-PL-CO should have the lowest Gpc. In ARDEC Wet, the mean Gpc of CO-CO-PL lines was 16.44 g kg^{-1}, which was significantly higher than that of PL-PL-CO lines, 15.97 g kg^{-1} (Table 24). This indicates that the selection for the alleles of interest was successful in obtaining the predicted results. However, this result was not repeated in the other environments. Because the allelic effects for Gpc were small (El-Feki 2010), the environmental conditions differed, and MTA for

Gpc were not very robust across environments, repeatability of the results became difficult to demonstrate.

Vishwakarma et al. (2014) successfully used marker-assisted backcrossing (MABC) to improve Gpc in wheat cultivar HUW468 (recipient parent) using SSR marker *Xucw108* linked to *Gpc-B1* and a donor parent Glu269 (DBW16/GluPro//2*DBW16). The improved lines had significantly higher Gpc. Tonk et al. (2010) reported a highly significant increase in Gpc when *T. aestivum* and *T. dicoccoides* were crossed, suggesting that the increase in Gpc is most likely due to the transferring of the high protein genes from the *T. dicoccoides* to the hexaploid cultivars. Yildirim et al. (2013) transferred the *Gli-B1* locus encoding γ-gliadin 45 and *Glu-B3* locus encoding low molecular weight (LMW)-2 glutenin of a high-quality Canadian durum wheat cultivar, Kyle, to the wheat variety Sacricanak-98 using MABC (backcrossed 4 times to the recurrent parent). The result was a considerable increase in Gpc and gluten quality. Chee et al. (2001) reported the successful transfer of a QTL for Gpc in wheat into a different genetic background. In these studies, the allelic effects were larger and the consistency of the MTA was greater than in our study.

At ARDEC Wet, the allelic effects of Gpc at *Xwmc182a* and *Xwmc182b* were -0.57 and 0.60 g kg^{-1}, respectively, while they were -0.16 and 0.16 g kg^{-1} in the DH (El-Feki 2010). This indicates that the directions of allelic effects were similar at individual loci in both studies, but the size effects were much higher at both loci in the current study. The difference of Gpc between CO-CO-PL and PL-PL-CO was 0.47 g kg^{-1}, about 60% as low as the sum of allelic effects at these loci (1.17 g kg^{-1}).

Conclusions

- MTA for Gpc detected at *Xwmc182a* and *Xwmc182b* were consistent with findings in the CO940610/Platte DH population. Also, a MTA for Gpc at *Glu-B1* was found.
- MTA for Gpc and Gac were found at all three loci investigated. These MTA at locus *Xwmc182a* were stable across two of three environments.
- The hypothesized results of combining alleles in a uniform background were confirmed in one environment, but not in the other two.
- The MABC lines with allelic combinations need further investigation.

CHAPTER 4: GENOME-WIDE ASSOCIATION STUDY FOR GRAIN QUALITY TRAITS OF A WINTER WHEAT ASSOCIATION MAPPING PANEL UNDER TWO WATER REGIMES

SUMMARY

With the development of high-throughput sequencing and statistical methodology, the genome-wide association study (GWAS) has become an effective approach for identifying quantitative trait loci (QTL) in crops including bread wheat (*Triticum aestivum* L.). GWAS is able to increase mapping resolution, reduce research time, and evaluate a broader range of alleles compared to traditional bi-parental mapping. Grain quality traits of wheat are complex traits that influence end-use quality. These quality traits are also correlated to varying degrees with grain yield (Gy). Drought conditions strongly affect both grain yield and quality traits. Under water deficit, Gy is reduced, but grain protein concentration (Gpc) generally increases.

The specific objectives of this study were to (1) evaluate the Hard Winter Wheat Association Mapping Panel (HWWAMP) for Gy, Gpc, grain protein deviation (Gpd), grain ash concentration (Gac), and test weight (Tw) in multiple fully irrigated and moisture-stressed environments; (2) examine correlations among the evaluated traits; and (3) perform a GWAS to identify single nucleotide polymorphisms (SNP) associated with the evaluated traits.

The study was conducted under fully irrigated (Wet) and water deficit (Dry) treatments at Greeley, Colorado in 2011/12 and Fort Collins, Colorado in 2012/13, for a total of four environments. The HWWAMP includes 299 entries (cultivars and advanced breeding lines), representative of the U.S. Great Plains region.

Both locations experienced about 50% grain yield reduction in the water deficit treatments compared to the fully irrigated treatments, while Gpc and Gac were higher in the drought treatments. Grain yield and test weight (Tw) were negatively correlated with Gpc and Gac, but Gy was positively correlated with Tw. Grain protein concentration, Gpd, and Gac correlated positively with each other, and in particular Gpc was highly correlated with Gpd. Broad-sense heritability estimates for all evaluated traits varied from low (<0.05) to high (>0.75) across the four environments. Of these, the heritability for Gac and Tw showed more stability over environments than the other traits.

The Genomic Association and Prediction Integrated Tool (GAPIT) package in R was used to determine the association between markers, using a set of 16,052 filtered SNP and data of the five phenotypic traits. Principal components (PC), used for controlling population structure, and a kinship matrix were incorporated in a mixed linear model (MLM). Different approaches for analysis in GAPIT changed the number of significant marker-trait associations (MTA) enormously. Employing a significance threshold of $P<0.001$, there were 307 significant MTA detected in the four environments for the five traits. The highest number of MTA was recorded for Tw (94) followed by Gac (75), Gpc (69), and Gy (43), with the least for Gpd (26). Considering a more stringent threshold of $P < 1.67 \times 10^{-4}$, there were only 40 significant MTA detected in the four environments for the five traits. The highest number of MTA was recorded for Tw (17), followed by Gac (12) and Gpc (5), with the least for Gy (3) and Gpd (3). The allelic effect of each MTA was small, and the phenotypic variation explained was approximately the same among MTA, about 5%.

4.0. INTRODUCTION

Wheat (*Triticum aestivum* L.) is one of the most highly produced cereal crops, after maize (*Zea mays* L.) and rice (*Oryza sativa* L.), and is important in meeting global food demands. The production in million tons was 1,018 for maize, 741 for rice, and 716 for wheat in 2013 (FAO 2013). Wheat supplies about 20% of the total food calories consumed worldwide, and is the major source of dietary proteins (Wen et al. 2012). The global demand for wheat is predicted to increase as the world's population increases to an estimated 9.6 billion by 2050 (UN 2013).

Wheat productivity is mainly limited by biotic and abiotic stresses. Abiotic factors are considered the main source (70%) of yield reductions (Boyer 1982). Among the abiotic stresses, drought is the most limiting factor reducing productivity of wheat and other crops in many parts of the world (Rampino et al. 2006; Mohammadi et al. 2015; Ahuja et al. 2010). Drought stress may occur from early in the growing season to terminally at grain filling and maturation, particularly in arid and semi-arid regions of the world. The climates of these regions usually have inadequate precipitation, a shortage of irrigation water, and high crop evapotranspiration demand.

Drought conditions also influence grain quality traits of wheat. Under drought stress grain protein concentration (Gpc) usually increases, and is negatively correlated with grain yield (Gy) (El-Feki et al. 2013; Daniel & Triboı 2002; Guttieri et al. 2000; X. Sun et al. 2010; Weightman et al. 2008). Water stress also changes grain protein composition in winter wheat. In the study by Saint Pierre et al. (2008), under water stress, the percentage of soluble protein decreased, but that of polymeric protein was not affected. Ahmed & Fayyaz-ul-Hassan (2015) found that under drought conditions, proline content increased tenfold and had a synergistic relationship with grain protein and grain ash. Grain ash concentration (Gac) was significantly higher when plants

were irrigated fewer times (Seleiman et al. 2011). Grain protein deviation (Gpd), a deviation from regression between Gpc and Gy, is an index for selection of both Gpc and Gy in breeding programs (Bogard et al. 2008; Guttieri et al. 2015; Monaghan et al. 2001; Mosleth et al. 2015). It is possible to use Gpd as a measure of the ability of a cultivar to gain a higher Gpc than expected from Gy (Monaghan et al. 2001). Test weight (Tw), a measure of the bulk density of grain, has a positive correlation with Gy and has direct effect on Gy (Mohammadi et al. 2012; Doğan 2009).

Association mapping (AM) or linkage disequilibrium (LD) mapping, an alternative approach to traditional bi-parental quantitative trait loci (QTL) mapping, is a tool to understand the genetic architecture of complex traits down to the sequence level. Bi-parental QTL mapping can evaluate only two alleles per locus at a time, has low mapping resolution, detects QTL specific to a particular population, is costly in time and materials, and detects chromosome regions but not genes. Meanwhile, AM overcomes many of these disadvantages due to its increased mapping resolution, reduced research time, and greater allele number (Zhu et al. 2008). Association mapping utilizes diverse germplasm instead of developing bi-parental mapping populations (Buntjer et al. 2005), thereby saving much more time than traditional QTL mapping. However, limitations of AM are the high risk of type I error (false positives) and the high sampling variance of rare alleles. The false discoveries are a major concern (Abiola et al. 2003) and can be partially attributed to spurious associations caused by population structure and unequal relatedness among individuals (Zhang et al. 2010).

Two AM approaches are in general use: (1) candidate-gene association mapping, and (2) whole genome scan or genome-wide association study (GWAS). Candidate-gene association mapping relates polymorphism in selected candidate genes controlling phenotypic variation for specific

traits, while GWAS surveys genetic variation in the whole genome to identify signals of association for various complex traits (Risch & Merikangas 1996).

Advances in genomic technology, interest in identifying novel alleles, and robust statistical methods have stimulated interest in AM for genetic research (Zhu et al. 2008). Association mapping was first applied for plant research in 2001, in a study on maize flowering time (Thornsberry et al. 2001). Association mapping has been used successfully to detect QTL in wheat for end-use quality traits (Breseghello & Sorrells 2006; Plessis et al. 2013; Zheng et al. 2009), grain yield and yield components (Edae et al. 2014; Maccaferri et al. 2011; Neumann et al. 2011; Dodig et al. 2012; Sukumaran et al. 2015), disease resistance (Adhikari et al. 2011; Ghavami et al. 2011; Maccaferri et al. 2010; Crossa et al. 2007; Yu et al. 2011; Peng et al. 2009; Maccaferri et al. 2015), and root traits (Canè et al. 2014).

Previous AM studies have identified QTL for yield, yield components, and drought related traits using the Hard Winter Wheat Association Mapping Panel (HWWAMP) in the Great Plains region (Awad 2015; Grogan 2015). However, those studies did not include grain quality traits in this population. In addition, there were QTL detected for wheat grain quality traits in the CO940610/Platte doubled haploid (DH) population (El-Feki et al. 2013) and in the CO940610/Platte recombinant inbred line (RIL) population (Chapter 2 of this dissertation).

The overall goal of this project was to use phenotypic and genotypic data from the HWWAMP to find single nucleotide polymorphism (SNP)-trait associations for four winter wheat grain quality traits (grain ash concentration, grain protein concentration, grain protein deviation, and test weight) and grain yield. The specific objectives of this study were to

(1) evaluate the HWWAMP for yield and grain quality traits in multiple fully-irrigated and moisture-stressed environments;

(2) examine correlations among the evaluated traits; and

(3) perform a GWAS for phenotypic and genotypic data to identify SNP associated with the evaluated traits.

4.1. MATERIALS AND METHODS

4.1.1. Association mapping panel

The germplasm used in this study was a large, diverse panel of 299 genotypes representative of the U.S. Great Plains region. The panel was developed by the Triticeae Coordinated Agricultural Project (TCAP, http://www.triticeaecap.org) and is designated as the Hard Winter Wheat Association Mapping Panel (HWWAMP). These genotypes consist of 258 hard red winter wheats and 41 hard white winter wheats, originating from both public and private wheat breeding programs in Colorado, Kansas, Montana, Nebraska, North Dakota, Oklahoma, South Dakota, and Texas. The genotypes include 185 recent cultivars, 106 experimental lines that were in advanced stages of testing in 2008, and 8 historic varieties developed before 1960. The historic varieties were the landrace Turkey; the two ancestral cultivars, Cheyenne and Kharkof; and five cultivars (Comanche, Wichita, Kiowa, Bison, and Tascosa). The panel was developed by breeders in the Great Plains region based on significance in production or contribution to the pedigrees of contemporary germplasm. Information about the genotypes including name, seed color, head type, chaff color, release year, origin, accession, and pedigrees is contained in Supplemental Table S.4.1 described by Guttieri et al. (2015).

4.1.2. Experimental design and management

The field trials for the HWWAMP were conducted in four environments in Colorado in two growing seasons, 2011/12 and 2012/13. In both years, the dry and wet treatments were planted side-by-side, separated by a buffer strip. Within each environment, experimental units (plots) were arranged in an unreplicated row-column design with two check varieties. The entry 'Wichita' (CI 11952) was included in the panel twice. The two check varieties 'Hatcher' (Haley et al. 2005) and 'Settler CL' (Baenziger et al. 2012) were replicated 15 times each per treatment, and were also included as experimental entries. Seed density was approximately 1,700,000 seeds ha^{-1}. Fungicide was applied at LIRF as a preventative for stripe rust, and there was an untreated outbreak of Russian wheat aphids in the ARDEC plots. Weed control was typical of practices for winter wheat in the area.

LIRF, Greeley, 2011/12

In the 2011/12 growing season, the panel and two checks, Hatcher and Settler CL, were grown in the USDA-Agricultural Research Service Limited Irrigation Research Farm (LIRF) in Greeley, Colorado (latitude 40º 4484' N; longitude 104º 636' W; elevation 1427 m). The soil type was Nunn clay loam, 0 to 1 percent slopes (http://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm). Soil pH ranges from 7.4 to 8.4. Planting date was on October 30, 2011, and harvest dates were July 3 and 13, 2012 for dry and wet treatments, respectively. The experimental unit was a six-row plot 1.83 m long and 1.52 m wide. Space between rows was 23 cm and between plots was 28 cm.

Water was applied using a drip irrigation system. Irrigation was supplied to both treatments beginning approximately at the booting stage (Zadoks et al. 1974). Amount of water applied was 335 mm (10 times) and 102 mm (five times) in wet (LIRF Wet) and dry (LIRF Dry) treatments,

respectively, between green-up in early spring and physiological maturity. Both wet and dry treatments received 82 mm precipitation between 1 Jan and harvest (Table 36 and 37).

ARDEC, Fort Collins, 2012/13

In the 2012/13 growing season, the panel and checks, Hatcher and Settler CL, were grown in CSU's Agricultural Research, Development, and Education Center (ARDEC) in Fort Collins, Colorado (latitude 40º 652' N; longitude 104º 996' W; elevation 1558 m). The soil type is Fort Collins loam, 0 to 1 percent slopes (Web Soil Survey).

A linear overhead sprinkler irrigation system was used for the wet treatment (ARDEC Wet) with a total of 222 mm applied during the growing season. No irrigation was used on the dry treatment (ARDEC Dry). Both wet and dry treatments received 124 mm precipitation between January 1 and July 17 (harvest) and 38 mm pre-planting irrigation in September 2012 (Table 36 and 37).

The trial was planted on October 2, 2012, and harvested on July 16 – 17, 2013. A total of 90 kg N/ha was applied during the growing season. Of this amount, 56 kg N/ha was applied in fall and 34 kg N/ha in spring before jointing. The experimental unit was the same as in growing season 2011/12.

4.1.3. Phenotypic evaluation

4.1.3.1. Grain yield

Grain in each experimental unit (six-row plot) was harvested using a Hege plot combine (Wintersteiger, Salt Lake City, Utah) after plants reached maturity. The harvested grain samples were dried for about three days at approximately 40^{o}C, weighted, and Gy was converted to kg ha^{-1}.

4.1.3.2. Grain ash concentration and grain protein concentration

Grain ash concentration (g kg^{-1}) and Gpc (g kg^{-1}) were measured by the Near Infrared Reflectance System (NIRs). A sample of approximately 100 g of cleaned wheat kernels was added to a transport module in a rectangular quartz cup (NIRSystems, part 01IH-0379). A Foss-Tecator NIR systems Model 6500 instrument (Foss North America, Eden Prairie, MN) was used to determine these two traits in the reflectance mode with a wavelength of 450-2498 nm. The estimates of the traits were obtained with calibrations from the manufacturer for both traits. This work was done in the Wheat Quality Laboratory, Department of Soil and Crop Science, Colorado State University.

4.1.3.3. Grain protein deviation

Grain protein deviation was obtained by conducting linear regression between Gpc and Gy using PROC REG of SAS 9.3 (SAS Institute Inc., Cary, NC, 2010). The BLUPs of Gpc and Gy were used for running the linear regression. The Gpc and Gy were the dependent and independent variables, respectively. The residuals produced by PROC REG were used as Gpd.

4.1.3.4. Test weight

A cleaned kernel sample was poured into a closed hopper centered over a metal cylinder 5.5 cm in diameter and 4.5 cm in tallness. The hopper valve was quickly opened to allow the grain to fill the cylinder. A standard striker held in both hands with the flat sides in a vertical position was used to remove the excess grain from the top of the cylinder with three full-length, zigzag motions. The cylinder was carefully placed on a scale platform, and the weight was read by an electronic scale. The unit for test weight in this study was kg hL^{-1}.

4.1.4. Phenotypic data analysis

Outliers

The original phenotypic data for each trait were inspected with a "heat map", which was created in Excel to determine whether any data points were unreasonably large or small and to visualize spatial trends. In addition, the original phenotypic data were also evaluated for normality of frequency distribution and potential outliers were detected by using JMP. Based on these inspections, several data points were removed. In the ARDEC Dry treatment, the removed data were for three traits (Gac, Gpc and Gpd) of plot 295 (entry 179, TX01M5009-28). In addition, Tw in the ARDEC Dry treatment had 19 missing data points; most of them were in row 1 of this treatment. This row was located on the very western edge of the TCAP HWWAMP field, adjacent to a field of sorghum and separated by a single 1.5 m-wide strip of 'Prairie Red' wheat. The entire row had very poor growth, probably due to severe drought (it was only irrigated by one nozzle from the irrigation system). In the LIRF Dry, the removed data were for three traits (Gac, Gpc and Gpd) of plot 222 (entry 136, DUKE) and plot 314 (entry 301, HATCHER, one of the occurrences of the repeated checks). As these suspected outliers were deleted, the normality and outliers were evaluated again to see if the distribution is normal.

BLUPs and spatial adjustment

After the suspected outliers in the original phenotypic data were removed, the remaining entries were used to estimate BLUPs within each treatment, with entries considered a random effect. However, the plots in row 1 of ARDEC Dry were given a border effect value of one, while all other plots in this treatment received values of zero. The border effect was considered as a covariate. The estimation was performed with PROC MIXED of SAS 9.3 (SAS Institute Inc., Cary, NC, 2010), including a spatial adjustment. Six different models (SAS statements) were

applied to adjust the experimental field spatial variation. These included non-adjusted, spherical, exponential, regular power, anisotropic power, and Matérn spatial models. The model with lowest Akaike's information criterion (AIC) was considered as the most appropriate one for a given trait and environment.

Broad-sense heritability estimation

Each treatment had two checks (Hatcher and Settler CL), which were replicated 15 times. Since these plots had no genetic variance, the average phenotypic variance of the two checks was considered to be the environmental variance. Variance of the entries in the panel, excluding the two checks was considered the population's phenotypic variance. Genetic variance was estimated by subtracted the environmental variance from the phenotypic variance. The estimate of broad-sense heritability (H^2) was calculated as in the following equation:

$$H^2 = \frac{V_G}{V_G + V_e} = \frac{\text{Phenotypic variance} - \text{Environmental variance}}{\text{Phenotypic variance}}$$

Where, V_G and V_e are genotypic and error variances, respectively.

4.1.5. Genome-wide association study

4.1.5.1. SNP markers

SNP marker genotyping was done at the USDA-ARS Small Grains Genotyping Lab in Fargo, North Dakota, using a high-density 90,000 gene-associated single nucleotide polymorphism (SNP) Illumina iSelect genotyping platform as described in Wang et al. (2014). SNP genotype calling was performed with the Genome Studio Polyploid Clustering Module v1.0 software developed in collaboration with Illumina, Inc. (Wang et al. 2014). The data were filtered to remove monomorphic SNP, multi-allelic SNP, SNP with low call rate, SNP that produced clusters, and SNP with small cluster distances to produce a set of 26,553 SNP. Manual curation

returned 21,555 SNP calls on 299 genotypes, of which 21,564 SNP were missing <10% of calls. The minor allele frequency (MAF) of 5,497 SNP was <5%. SNP with MAF <5% and SNP with ≥10% of calls missing were removed to produce a set of 16,052 filtered SNP (Mary Guttieri, University of Nebraska-Lincoln, personal communication).

The software package GAPIT (Genomic Association and Prediction Integrated Tool) in R (Lipka et al. 2012) was used to determine the association between markers (using the 16,052 filtered SNP) and phenotypic traits. A genetic map of the SNP was obtained from http://wheatgenomics.plantpath.ksu.edu/ (download date: June 15, 2015). The 16,052 filtered SNP were comprised of 14,829 SNP uniquely mapped on the 21 wheat chromosomes and 1,223 unmapped SNP, which were assigned chromosome number 22. Of the 14,829 SNP, 5,971 were distributed in the A-genome, 7,244 in the B-genome, and 1,614 in the D-genome (Table 27).

Table 27. Distribution of markers across genomes and chromosome in the hard winter wheat association-mapping panel (HWWAMP), provided by Mary Guittieri (University of Nebraska-Lincoln, personal communication).

Chromosome	Genome			Total SNP
	A	B	D	
1	998	1,188	420	2,606
2	915	1,101	520	2,536
3	747	1,065	230	2,042
4	713	425	52	1,190
5	754	1,475	153	2,382
6	920	1,182	117	2,219
7	924	808	122	1,854
Mapped SNP	5,971	7,244	1,614	14,829
Unmapped SNP				1,223
Grand total SNP				16,052

4.1.5.2. Model selection

Many SNP-trait association models were analyzed in GAPIT (Table 28). Each of the models is a different combination of the 'population parameters previously determined' (P3D) (Zhang et al.

2010), kinship matrix, principal components, compression, trait, and environment. The model could either include or exclude P3D. If P3D is excluded from the model, it takes much more time (about one hour for each model) compared to the model with P3D.

Table 28. Positive R^2 values obtained with different model combinations used for marker-trait association (MTA) identification.

Kinship matrix	Types	Positive R^2		
		P3D with compression	P3D without compression	No P3D and no compression
IBS	External	all	all	n/a
rrBLUP	External	all	all	n/a
Loiselle	Internal	all	all	n/a
VanRaden	Internal	most	all	n/a
EMMA	Internal	most	most	most

Kinship matrices were calculated externally and imported or internally calculated by GAPIT. The internal kinship matrices calculated by GAPIT consisted of Loiselle, VanRaden, and EMMA. These internal kinship matrices were not used for further analyses because VanRaden and EMMA produced negative R^2 (G. Sun et al. 2010), while Loiselle provided less significant MTA than either IBS or rrBLUP, e.g., for Gpc in the LIRF Dry (Figure 10). The external kinship matrices included Identity-by-State (IBS), which was calculated by Scott Reid (Department of Soil and Crop Sciences, Colorado State University) using the TASSEL software package (Bradbury et al. 2007), and rrBLUP (Endelman 2011), which was computed by Mary Guttieri (University of Nebraska-Lincoln, personal communication) in the rrBLUP package for R (Endelman 2011). These two external kinship matrices showed similarity in terms of the matrix generated and the MTA detected as well (Figure 10).

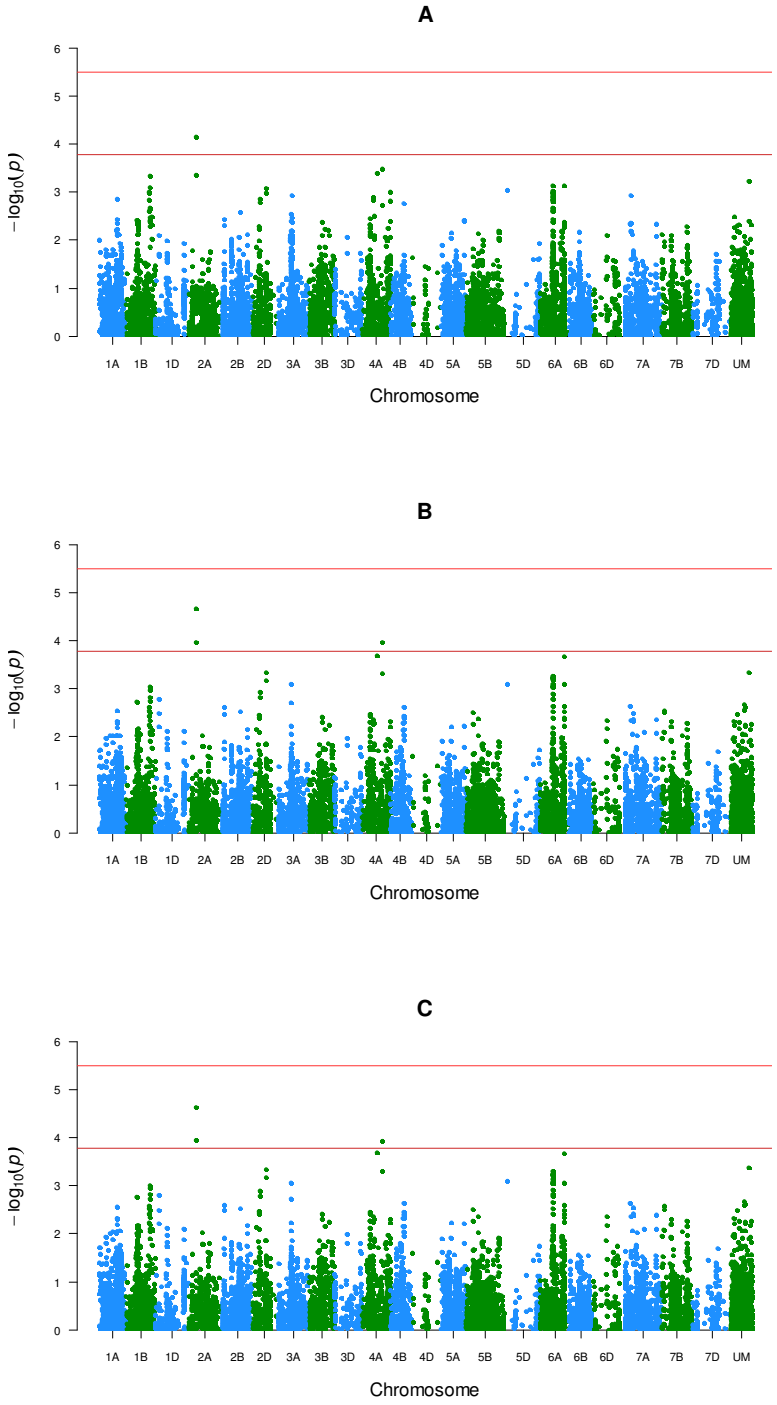


Figure 10. Manhattan plots for Gpc in the LIRF Dry with the different kinship matrices used for analysis. The X-axis is the genomic position of each SNP; the Y-axis is the negative logarithm of the P-value obtained from the GWAS model. The lines identify the two threshold lines of significance. A, Loiselle kinship matrix used; B, IBS kinship matrix used; C, rrBLUP kinship matrix used. UM, unmapped SNP.

Principal components (PC) (Price et al. 2006) in mixed linear model (MLM) (Yu et al. 2006; Zhang et al. 2010) were used for controlling population structure in this set of germplasm. The number of PC varied depending on MTA model. The optimum number of PC for a model was selected based on the Bayesian information criterion (BIC) (Schwarz 1978), the larger the BIC, the better the number of PC. For MTA models with P3D, the number of PC up to five was manually provided, then GAPIT itself ran all six options from zero to five and selected the optimum number of PC for that MTA model.

An MTA model can be run with or without compression (Lipka et al. 2012). Based on the above preliminary analysis of different models, it was decided to use P3D, rrBLUP kinship matrix, optimum PC selection, and non-compression for reporting results.

Multiple comparison adjustment

GAPIT provides an unadjusted *P*-value for each MTA. Because the analysis includes thousands of SNP loci and hundreds of wheat lines, there is a multiple testing effect, causing many false positive results. GAPIT also provided false discovery rate-adjusted *P*-value (FDR) (Benjamini & Hochberg 1995).

The approach suggested by Bonferroni (1935 & 1936) was included on the Manhattan plots. The upper line in Figure 10 shows the Bonferroni significance threshold. To avoid the high Bonferroni penalty, Gao et al. (2008) suggested using the adjusted point-wise significance level (α_P) in GWAS, calculated as in the following equation.

$$\alpha_P = \frac{\text{Experimental wide error rate } (\alpha_e)}{\text{Effective number of independent test } (M_{eff})}$$

In the current study, α_e was kept at nominal level of 0.05, while M_{eff} was 276, obtained from an algorithm in R (Gao et al. 2008). As a result, α_P equaled 1.6×10^{-4}, which is the lower line in

Figure 4.1.5.1 indicating Gao's significance threshold. Gao et al. (2008) reported that this adjusted threshold is closer to the permutation-based correction and is more accurate than the multiple testing adjusted approaches suggested by Bonferroni (1935 & 1936) or Li & Ji (2005). Therefore, this study used α_P as a significance threshold for strong MTA.

Phenotypic variation explained by SNP

The percentage of variation explained by the marker (R^2) was calculated by subtracting R^2 for the model without a given SNP from R^2 for the model including the SNP.

The GAPIT program by default assigns a symbol indicating the allelic effect estimate with respect to the nucleotide that is second in alphabetical order. For example, if the nucleotides at a SNP are "A" and "T", then a positive allelic effect indicates that "T" is favorable (GAPIT user manual Version 2.0).

4.2. RESULTS

4.2.1. Trait distribution and means

Distribution

For all evaluated traits, both the original data and BLUPs were normally distributed according to the Shapiro-Wilk test at $P>0.05$. Therefore, it was not necessary to transform the data in any case before further use. The frequency distributions of the lines for each of the five traits in each environment are shown in Figure 15.

Trait means

Grain yield in the dry treatments was significantly (P<0.0001) lower than that in the wet treatments. Grain yield in LIRF Dry was 2572.7 kg ha^{-1}, about a 47% reduction compared to Gy in LIRF Wet. Grain yield in ARDEC Dry was 2535.6 kg ha^{-1}, about a 49% reduction in comparison with Gy in ARDEC Wet (Table 29). This indicates that drought stress strongly decreased Gy.

Grain protein concentration and Gac in the dry treatments were significantly (P<0.0001) higher than those in the wet treatments. In the LIRF Dry and LIRF Wet, Gpc was 152.7 and 145.4 g kg^{-1}; and Gac was 15.2 and 14.9 g kg^{-1}, respectively. In ARDEC Dry and ARDEC wet, Gpc was 148.7 and 131.3 g kg^{-1}, and Gac was 15.1 and 14.7 g kg^{-1}, respectively (Table 29). This suggests that drought stress increased Gpc and Gac.

Test weight in LIRF Dry was significantly (P<0.0001) higher than that in LIRF Wet, while Tw in ARDEC Dry was significantly (P<0.0001) lower than that in ARDEC Wet (Table 29). Grain protein deviation did not differ significantly between dry and wet treatments in either LIRF or ARDEC (Table 29).

Table 29. Means, standard errors (SE) and ranges for five evaluated traits of the Hard Winter Wheat Association Mapping Panel in LIRF 2011/12 and ARDEC 2012/13 growing seasons.

		Dry					Wet					
Env.	**Traits¶**	**n**	**Mean**	**SE**	**Min.**	**Max.**	**n**	**Mean**	**SE**	**Min.**	**Max.**	***P-value*†**
LIRF	Gpc. (g kg^{-1})	298	152.7	0.46	132.6	176.5	298	145.4	0.55	123.7	175.6	<0.0001
	Gpd (g kg^{-1})	295	1.05	0.30	-12.11	15.45	298	0.58	0.38	-16.95	20.81	0.3291
	Gac. (g kg^{-1})	298	15.2	0.04	13.5	17.3	298	14.9	0.04	12.6	17.5	<0.0001
	Tw (kg hL^{-1})	299	80.0	0.09	73.0	83.9	299	75.0	0.21	61.5	81.0	<0.0001
	Gy (kg ha^{-1})	296	2572.7	23.4	1463.8	3945.2	299	4813.7	63.1	685.8	7700.4	<0.0001
ARDEC	Gpc. (g kg^{-1})	298	148.7	0.31	135.9	165.0	299	131.3	0.37	114.1	150.5	<0.0001
	Gpd (g kg^{-1})	298	0.10	0.10	-4.12	5.37	299	0.48	0.18	-7.89	11.57	0.0593
	Gac. (g kg^{-1})	298	15.1	0.04	13.4	18.3	299	14.7	0.05	12.7	18.3	<0.0001
	Tw (kg hL^{-1})	280	74.6	0.11	68.4	78.3	299	78.6	0.10	71.2	82.9	<0.0001
	Gy (kg ha^{-1})	299	2535.6	6.2	2248.0	3021.2	299	4967.6	48.5	2635.0	7282.2	<0.0001

¶ Gpc, grain protein concentration; Gpc, grain protein deviation; Gac, grain ash concentration; Tw, test weight; Gy, grain yield.

n, number of samples evaluated.

† Significance of difference between treatment means.

4.2.2. Correlation among traits

Grain yield was significantly (P<0.001) and negatively correlated with Gpc (r = -0.52 to – 0.67) and Gac (r = -0.24 to -0.45) (Tables 30 and 31). This confirms that Gy had an inverse relationship with Gpc and Gac, hindering improvement of both Gy and Gpc in breeding programs. In contrast, Gy correlated positively with Tw. Correlation coefficients ranged from 0.14 (P<0.05) in LIRF Dry to 0.65 (P<0.001) in LIRF Wet (Tables 30 and 31). This suggests that selection for higher Tw is able to increase Gy.

Grain protein concentration, Gpd and Gac correlated positively with each other. Grain protein concentration was highly (P<0.001) correlated with Gpd. Correlation coefficients ranged from 0.72 to 87 (Tables 30 and 31), indicating that selection for higher Gpd can improve Gpc. In contrast, Gpc correlated negatively with Tw in three environments, but not in ARDEC Wet. The coefficients between these traits were -0.15 (P<0.01) in LIRF Dry, -0.22 (P<0.001) in ARDEC Dry, and -0.51 (P<0.001) in LIRF Wet (Tables 30 and 31). With low correlation, it is possible to select both traits in the favorable direction although they had an inverse relationship.

Grain protein deviation significantly (P<0.001, r=0.20) correlated with Gy only in LIRF Wet, but not in the three other environments (Tables 30 and 31). Grain protein deviation had low and unstable correlation with Gy, but very high correlation with Gpc, indicating that Gpd is a useful selection index to improve both Gy and Gpd simultaneously in a breeding program.

Table 30. Pearson correlation coefficients (n = 295 to 299) among traits of the Hard Winter Wheat Association Mapping Panel at LIRF, Greeley, CO in the 2011/12 growing season. Correlations for the wet treatment are above the diagonal and those for the dry treatment are below the diagonal.

	Gy†	Gpc	Gpd	Gac	Tw
Gy		-0.67***	-0.07	-0.45***	0.65***
Gpc	-0.61***		0.74***	0.50***	-0.51***
Gpd	-0.20***	0.87***		0.26***	-0.06
Gac	-0.25***	0.34***	0.26***		-0.53***
Tw	0.14*	-0.15**	-0.10	-0.23***	

*, **, *** indicates significance at the 0.05, 0.01, and 0.001 levels of probability, respectively.

† Gy, grain yield; Gpc, grain protein concentration at 12% moisture level; Gpd, grain protein deviation; Gac, grain ash concentration; Tw, test weight.

Table 31. Pearson correlation coefficients among traits of the Hard Winter Wheat Association Mapping Panel at ARDEC, Fort Collins, CO in the 2012/13 growing season. Correlations for the wet treatment are above diagonal and those for the dry treatment are below the diagonal.

	Gy†	Gpc	Gpd	Gac	Tw
Gy		-0.52***	0.01	-0.34***	0.31***
		299	299	299	299
Gpc	-0.55***		0.72***	0.44***	-0.07
	298‡		299	299	299
Gpd	-0.04	0.80***		0.28***	0.19***
	299	298		299	299
Gac	-0.24***	0.38***	0.30***		-0.17**
	298	298	298		299
Tw	0.34***	-0.22***	-0.08	-0.19**	
	280	279	280	279	

*, **, *** indicates significance at the 0.05, 0.01, and 0.001 levels of probability, respectively.

† Gy, grain yield; Gpc, grain protein concentration at 12% moisture level; Gpd, grain protein concentration deviation; Gac, grain ash concentration; Tw, test weight.

‡ number of sample used.

4.2.3. Heritability of traits

Broad-sense heritability estimates for Gy, Gpc and Gpd in ARDEC Dry were close to zero or negative values. In ARDEC Dry, the heritability estimates were -0.31, 0.08 and -0.06 for Gy, Gpc and Gpd, respectively (Table 32). This indicated that experimental errors or the estimation method might have caused the problems.

The heritability estimates for all evaluated traits varied considerably over environments, from low (<0.50) to high (>0.75). The heritability estimates for Gy were 0.19 in LIRF Dry and 0.79 in LIRF Wet (Table 32). This suggests that experimental errors and/or environmental factors had a large effect on the results.

Generally, heritability estimates for Gpc, Gpd, Gac and Tw were moderate (>0.50) to high, with the exception of Gpc and Gpd in ARDEC Dry. The heritability estimates for Gpc were 0.66, 0.91 and 0.92 in ARDEC Wet, LIRF Wet and LIRF Dry, respectively. The heritability estimates for Gac and Tw showed more stability across environments, ranging from 0.69 to 0.81 for Gac and from 0.67 to 0.88 for Tw (Table 32). These suggested that a large part of the expression of these traits was genetically controlled, thus, it would be easier to obtain progress from selection in a breeding program.

Table 32. Broad-sense heritability estimates (H^2) for five evaluated traits of the Hard Winter Wheat Association Mapping Panel in LIRF 2011/12 and ARDEC 2012/13 growing seasons.

	Broad-sense heritability			
	LIRF Dry	LIRF Wet	ARDEC Dry	ARDEC Wet
Grain yield	0.19	0.79	-0.31	0.34
Grain protein concentration	0.91	0.92	0.08	0.66
Grain protein deviation	0.53	0.83	-0.06	0.41
Grain ash concentration	0.69	0.70	0.69	0.81
Test weight	0.73	0.88	0.67	0.86

4.2.4. Marker-trait associations

MTA with P <0.001

MTA detected at the unadjusted significance level of *P* <0.001 for all five evaluated traits in the four environments are shown in Table 33. Considering the criterion of *P*<0.001, there were 307 significant MTA detected in the four environments for the five traits. The highest number of MTA was recorded for Tw (94) followed by Gac (75), Gpc (69), and Gy (43), with the least for Gpd (26).

Table 33. Marker-trait associations detected in the Hard Winter Wheat Association Mapping Panel at the unadjusted *P*-value < 0.001 for five traits in four environments.

Trait	Env.	SNP	Chr.	Position	P	FDR	MAF†	Allelic effect	R^2
Gy	LIRF Dry	BS00021704_51	6A	140.9	5.51E-05	0.57096	0.20	130.75	0.052
		Excalibur_c4518_2931	6A	140.9	7.11E-05	0.57096	0.21	-125.91	0.050
		wsnp_Ex_rep_c76495_73453891	6A	140.7	0.00026	0.74858	0.21	115.30	0.042
		BobWhite_c26503_61	6A	99.0	0.00035	0.74858	0.43	92.22	0.040
		RAC875_rep_c114621_200	2D	85.2	0.00035	0.74858	0.16	125.50	0.040
		RAC875_rep_c114621_464	2D	82.6	0.00041	0.74858	0.16	-124.30	0.040
		RAC875_c30919_311	2D	82.5	0.00044	0.74858	0.17	-121.63	0.039
		Excalibur_c3969_1208	3A	145.6	0.00062	0.74858	0.13	129.01	0.037
		BS00097302_51	2D	82.8	0.00065	0.74858	0.13	-149.02	0.037
		Ku_c21235_676	5A	98.9	0.00069	0.74858	0.13	-126.60	0.036
		Tdurum_contig29607_294	6A	140.9	0.00082	0.74858	0.11	145.23	0.035
		wsnp_RFL_Contig2104_1368653	2D	82.5	0.00088	0.74858	0.16	-117.71	0.035
		IAAV2593	5B	122.6	0.00095	0.74858	0.17	-114.83	0.035
		RAC875_c29079_177	2D	80.5	0.00099	0.74858	0.09	148.84	0.034
Gy	LIRF Wet	Excalibur_c28017_641	2A	109.2	0.00059	0.48607	0.32	-245.63	0.029
		Ku_c13307_1116	3A	88.0	0.00068	0.48607	0.13	372.59	0.028
		Kukri_c9595_242	6A	74.2	0.00071	0.48607	0.07	427.96	0.028
		wsnp_Ku_c3286_6111360	3A	90.5	0.00078	0.48607	0.13	-371.44	0.028
		wsnp_Ex_c52577_56128947	6A	74.2	0.00081	0.48607	0.07	434.17	0.028
		Kukri_c7794_1247	6A	74.2	0.00082	0.48607	0.07	433.84	0.028
		RFL_Contig4441_627	6A	74.2	0.00082	0.48607	0.07	433.84	0.028
		Kukri_c16463_584	6A	74.2	0.00082	0.48607	0.07	-433.84	0.028
		wsnp_Ku_c5243_9344536	3A	88.0	0.00085	0.48607	0.13	-363.55	0.027
		wsnp_Ex_rep_c67878_66584488	6A	74.2	0.00099	0.48607	0.06	446.45	0.027

Table 33. (Continued)

Trait	Env.	SNP	Chr.	Position	P	FDR	MAF†	Allelic effect	R^2
Gy	ARDEC Dry	BobWhite_rep_c52911_146	3D	143.0	0.00047	0.92168	0.49	-25.99	0.041
		RAC875_c40654_206	4A	120.1	0.00062	0.92168	0.26	25.58	0.039
Gy	ARDEC Wet	wsnp_Ex_c29382_38422739	UM	118.8	2.50E-05	0.40116	0.33	-253.85	0.061
		wsnp_Ex_c42372_48966781	6B	64.1	0.00017	0.54281	0.30	-222.43	0.048
		Tdurum_contig64795_494	UM	115.1	0.00022	0.54281	0.22	-244.84	0.047
		BobWhite_c10852_309	6B	64.1	0.00025	0.54281	0.30	-216.93	0.046
		BobWhite_c47831_87	6B	64.1	0.00025	0.54281	0.30	-216.93	0.046
		RFL_Contig311_951	6B	64.1	0.00025	0.54281	0.30	-216.93	0.046
		wsnp_Ex_c46160_51746546	6B	64.1	0.00025	0.54281	0.30	-216.93	0.046
		RAC875_c13240_96	2B	115.8	0.00032	0.54281	0.22	-236.46	0.044
		RAC875_c24962_1326	6B	64.1	0.00034	0.54281	0.30	-211.56	0.044
		Excalibur_c5136_2314	6B	64.1	0.00034	0.54281	0.30	211.56	0.044
		Excalibur_c11968_204	4A	103.0	0.00038	0.56104	0.09	-320.30	0.043
		Excalibur_c79066_165	6B	64.1	0.00046	0.58204	0.23	231.94	0.042
		BS00063365_51	2B	114.8	0.00047	0.58204	0.34	204.94	0.042
		Excalibur_c146_92	5B	183.9	0.00065	0.71735	0.09	316.29	0.039
		wsnp_Ra_c4660_8405634	2B	115.8	0.00075	0.71735	0.41	191.62	0.039
		wsnp_BE499016B_Ta_2_1	3B	75.4	0.00076	0.71735	0.39	-186.41	0.039
		Tdurum_contig78534_314	2B	115.8	0.00076	0.71735	0.41	191.61	0.039
Gpc	LIRF Dry	BS00021706_51	2A	47.2	2.30E-05	0.36967	0.37	-2.17	0.048
		Kukri_c14598_614	2A	47.2	0.00011	0.58118	0.38	-2.00	0.040
		BS00064369_51	4A	118.7	0.00012	0.58118	0.28	2.11	0.040
		Tdurum_contig22511_355	4A	89.1	0.0002	0.58118	0.18	2.36	0.037
		Excalibur_c4518_2931	6A	140.9	0.00021	0.58118	0.21	2.35	0.037
		TA004558-1018	UM	102.4	0.00043	0.58118	0.07	-3.48	0.033

Table 33. (Continued)

Trait	Env.	SNP	Chr.	Position	P	FDR	MAF†	Allelic effect	R^2
Gpc	LIRF Dry	Kukri_c31121_1460	2D	80.1	0.00046	0.58118	0.19	-2.16	0.033
		wsnp_RFL_Contig3501_3652740	4A	118.7	0.00049	0.58118	0.31	1.88	0.032
		wsnp_Ex_c35465_43610634	6A	74.2	0.0005	0.58118	0.07	-3.50	0.032
		IACX6453	6A	74.2	0.00056	0.58118	0.07	-3.54	0.032
		Kukri_c50910_866	6A	74.2	0.0006	0.58118	0.07	3.56	0.031
		wsnp_Ex_rep_c102845_87922204	6A	74.2	0.0006	0.58118	0.07	3.56	0.031
		wsnp_Ku_c7794_13356946	6A	74.2	0.0006	0.58118	0.07	3.56	0.031
		BS00084846_51	6A	74.2	0.0006	0.58118	0.07	-3.56	0.031
		Kukri_c3075_391	6A	74.2	0.00063	0.58118	0.07	3.47	0.031
		Excalibur_c44325_339	2D	80.4	0.00067	0.58118	0.14	-2.33	0.031
		IAAV583	6A	74.2	0.00072	0.58118	0.06	3.67	0.030
		Kukri_c46526_103	5D	23.9	0.00079	0.58118	0.36	1.73	0.030
		RFL_Contig2605_672	6A	74.2	0.00081	0.58118	0.07	3.40	0.030
		BS00021704_51	6A	140.9	0.00085	0.58118	0.20	-2.15	0.030
		BS00072153_51	3A	88.0	0.00086	0.58118	0.07	3.28	0.030
		BobWhite_c20782_697	6A	77.1	0.00086	0.58118	0.07	-3.63	0.029
		Kukri_c16463_584	6A	74.2	0.0009	0.58118	0.07	3.42	0.029
		Kukri_c7794_1247	6A	74.2	0.0009	0.58118	0.07	-3.42	0.029
		RFL_Contig4441_627	6A	74.2	0.0009	0.58118	0.07	-3.42	0.029
		wsnp_Ex_c4436_7981037	1B	141.8	0.00097	0.60053	0.18	2.33	0.029
Gpc	LIRF Wet	D_GBUVHFX02JKG4A_54	2D	22.5	9.01E-05	0.44258	0.36	-2.33	0.040
		Excalibur_c44325_339	2D	80.4	0.00018	0.44258	0.14	-2.88	0.037
		wsnp_Ku_c3286_6111360	3A	90.5	0.00023	0.44258	0.13	3.61	0.035
		BS00023026_51	3A	90.5	0.00028	0.44258	0.12	-3.59	0.035
		wsnp_Ex_c1149_2206471	3A	88.0	0.00028	0.44258	0.12	-3.59	0.035
		Ex_c9685_1264	3A	90.5	0.00028	0.44258	0.12	3.59	0.035

Table 33. (Continued)

Trait	Env.	SNP	Chr.	Position	P	FDR	MAF†	Allelic effect	R^2
Gpc	LIRF Wet	GENE-3343_183	UM	49.9	0.00037	0.44258	0.12	-3.56	0.033
		BS00068083_51	1B	78.4	0.00038	0.44258	0.26	2.35	0.033
		BS00037537_51	3A	90.5	0.00039	0.44258	0.13	3.47	0.033
		CAP12_c590_307	1B	136.0	0.00039	0.44258	0.06	4.36	0.033
		IAAV6317	3A	88.0	0.00041	0.44258	0.12	3.50	0.033
		Tdurum_contig29087_757	1B	136.0	0.00047	0.44258	0.06	4.28	0.032
		wsnp_Ku_c32404_42016343	3A	88.0	0.00048	0.44258	0.12	-3.50	0.032
		wsnp_Ex_c12269_19597415	3A	90.5	0.00055	0.44258	0.13	-3.36	0.031
		BS00101401_51	3A	88.0	0.00059	0.44258	0.12	-3.44	0.031
		tplb0030j08_1960	2D	18.5	0.00064	0.44258	0.32	-2.02	0.030
		Ku_c14313_1194	3A	90.6	0.00065	0.44258	0.13	-3.26	0.030
		wsnp_Ra_c35889_44345459	3A	88.0	0.00067	0.44258	0.12	-3.43	0.030
		IAAV1334	3A	88.0	0.00073	0.44258	0.12	-3.37	0.030
		Ra_c38505_544	3A	90.5	0.00075	0.44258	0.13	3.27	0.030
		RAC875_c66369_296	6A	77.0	0.00082	0.44258	0.19	-2.30	0.029
		BS00070511_51	3A	88.0	0.00083	0.44258	0.12	-3.31	0.029
		IAAV5030	3A	88.0	0.00083	0.44258	0.12	-3.31	0.029
		Kukri_c31121_1460	2D	80.1	0.00084	0.44258	0.19	-2.31	0.029
		wsnp_CAP12_c2297_1121142	3B	119.4	0.00087	0.44258	0.28	-2.11	0.029
		Excalibur_c44325_638	UM	13.7	0.0009	0.44258	0.33	2.20	0.029
		Kukri_c6288_364	3A	90.6	0.00099	0.44258	0.18	2.97	0.028
Gpc	ARDEC Dry	wsnp_Ex_c19207_28125072	4A	69.6	1.32E-05	0.21186	0.17	1.98	0.062
		Kukri_c55081_219	3D	101.1	0.0003	0.95154	0.13	1.85	0.042
		Excalibur_c24511_1196	4A	74.5	0.00037	0.95154	0.21	1.47	0.041
		Excalibur_c46394_762	4A	74.5	0.00049	0.95154	0.22	1.40	0.039
		Excalibur_c11727_837	2B	157.2	0.00072	0.95154	0.41	1.24	0.037

Table 33. (Continued)

Trait	Env.	SNP	Chr.	Position	P	FDR	MAF†	Allelic effect	R^2
Gpc	ARDEC Dry	GENE-0223_239	UM	20.5	0.00072	0.95154	0.18	1.57	0.037
		BobWhite_c20322_183	4A	137.5	0.00073	0.95154	0.10	-1.98	0.037
		BobWhite_c20322_559	4A	137.9	0.00086	0.95154	0.10	-1.93	0.036
		wsnp_JD_c100_159424	1B	99.5	0.00087	0.95154	0.05	2.46	0.036
		RAC875_c291_647	1B	99.5	0.00087	0.95154	0.05	-2.46	0.036
		BS00072025_51	4A	66.3	0.00099	0.95154	0.22	-1.32	0.035
		BS00045546_51	4A	137.9	0.001	0.95154	0.10	1.82	0.035
Gpc	ARDEC Wet	wsnp_Ku_c2637_5009091	6D	19.0	0.0006	0.95069	0.26	1.52	0.037
		BS00031339_51	5B	183.9	0.00061	0.95069	0.06	2.85	0.037
		D_contig36160_443	3D	149.0	0.00072	0.95069	0.16	1.86	0.036
		wsnp_Ex_c20457_29526403	6A	123.5	0.00098	0.95069	0.07	2.72	0.034
Gpd	LIRF Dry	wsnp_CAP11_rep_c4066_1921894	7A	62.7	4.27E-05	0.68503	0.09	-3.01	0.053
		BS00064369_51	4A	118.7	9.65E-05	0.72781	0.28	1.80	0.048
		wsnp_Ku_c3237_6024936	4A	70.0	0.00018	0.72781	0.14	2.18	0.045
		wsnp_Ex_c24474_33721784	4A	70.0	0.00024	0.72781	0.15	2.04	0.043
		IAAV5704	1B	70.1	0.00034	0.72781	0.37	-1.95	0.041
		Tdurum_contig54860_1393	7A	51.0	0.00038	0.72781	0.37	-1.57	0.040
		Excalibur_c60612_236	UM	14.6	0.00067	0.72781	0.07	-2.64	0.037
		wsnp_CAP7_c940_480745	1B	70.1	0.00098	0.72781	0.38	-1.78	0.034
Gpd	LIRF Wet	BS00039187_51	2B	152.6	7.64E-05	0.90722	0.31	1.96	0.052
		Tdurum_contig51717_1463	6A	74.2	0.00034	0.90722	0.33	-1.83	0.042
		BS00010811_51	6A	74.2	0.00035	0.90722	0.33	1.83	0.042
		Tdurum_contig51717_1582	6A	74.2	0.00035	0.90722	0.33	1.83	0.042
		Ra_c28284_223	6A	74.2	0.00057	0.90722	0.34	-1.76	0.039
		RAC875_rep_c111781_179	3B	6.0	0.00072	0.90722	0.40	-1.53	0.038
		RAC875_c4602_886	2B	153.1	0.00089	0.90722	0.41	1.54	0.036

Table 33. (Continued)

Trait	Env.	SNP	Chr.	Position	P	FDR	MAF†	Allelic effect	R^2
Gpd	LIRF Wet	Ra_c58335_189	3B	12.6	0.00094	0.90722	0.10	-2.89	0.036
		wsnp_Ex_c24958_34212226	6A	74.2	0.00097	0.90722	0.31	-1.68	0.036
		wsnp_Ex_c43412_49738738	6A	74.2	0.00097	0.90722	0.31	-1.68	0.036
		wsnp_Ex_c7876_13389283	6A	74.2	0.00100	0.90722	0.32	1.68	0.035
Gpd	ARDEC Dry	wsnp_Ex_c41074_47987860	4A	80.1	0.00046	0.95111	0.29	1.12	0.041
		RAC875_c20429_903	4A	80.1	0.00061	0.95111	0.30	1.07	0.039
		wsnp_Ex_c19207_28125072	4A	69.6	0.00064	0.95111	0.17	1.29	0.039
		BS00072025_51	4A	66.3	0.00072	0.95111	0.22	-1.13	0.038
		BS00037442_51	4B	71.5	0.00085	0.95111	0.06	-2.09	0.037
		BS00056645_51	2B	161.4	0.001	0.95111	0.16	-1.34	0.036
Gpd	ARDEC Wet	D_contig36160_443	3D	149.0	0.00034	0.88634	0.16	1.70	0.041
Gac	LIRF Dry	wsnp_Ra_c15715_24192817	4B	59.9	4.45E-05	0.27892	0.13	-0.29	0.053
		BS00064935_51	4B	63.2	6.00E-05	0.27892	0.13	0.29	0.051
		Excalibur_c26226_458	4B	59.9	6.71E-05	0.27892	0.12	-0.29	0.050
		IAAV8654	4B	63.2	9.50E-05	0.27892	0.12	0.28	0.048
		BS00064884_51	4B	63.2	0.00011	0.27892	0.11	-0.28	0.047
		Ex_c7593_754	4B	63.2	0.00011	0.27892	0.12	-0.28	0.047
		BS00065555_51	4B	63.2	0.00012	0.27892	0.13	0.27	0.047
		BobWhite_c18721_190	6A	43.1	0.00018	0.36129	0.49	0.18	0.044
		BobWhite_c38340_243	4B	59.9	0.00021	0.36606	0.13	-0.26	0.043
		IAAV4238	1A	144.9	0.00026	0.37693	0.43	0.16	0.042
		BS00063804_51	4B	63.4	0.00031	0.37693	0.13	0.25	0.041
		Excalibur_c29141_864	4B	62.6	0.00043	0.37693	0.16	0.21	0.039
		Kukri_c3986_346	6A	43.1	0.00044	0.37693	0.47	-0.17	0.039
		Ra_c2253_1271	6A	37.1	0.00045	0.37693	0.48	-0.17	0.039
		BS00011607_51	6A	71.2	0.00046	0.37693	0.13	-0.26	0.038

Table 33. (Continued)

Trait	Env.	SNP	Chr.	Position	P	FDR	MAF†	Allelic effect	R^2
Gac	LIRF Dry	RAC875_c2253_238	6A	37.1	0.00047	0.37693	0.48	0.17	0.038
		tplb0057e06_1607	6A	37.1	0.00047	0.37693	0.48	0.17	0.038
		Tdurum_contig98255_84	4B	64.0	0.00048	0.37693	0.13	0.24	0.038
		BS00059475_51	3B	144.7	0.00057	0.37693	0.46	-0.17	0.037
		RAC875_c2253_1255	6A	37.1	0.00059	0.37693	0.48	-0.16	0.037
		Ra_c59822_1439	4B	63.2	0.0006	0.37693	0.13	-0.23	0.037
		wsnp_Ex_c9763_16125630	6A	37.1	0.00061	0.37693	0.47	-0.16	0.037
		BobWhite_c4810_190	4B	64.0	0.00066	0.37693	0.13	0.24	0.036
		Tdurum_contig22511_355	4A	89.1	0.00067	0.37693	0.18	0.19	0.036
		BS00077789_51	6A	37.0	0.00067	0.37693	0.49	-0.16	0.036
		Excalibur_c8131_385	4B	64.0	0.00072	0.37693	0.13	0.23	0.036
		wsnp_Ku_c10377_17181816	6A	35.7	0.00075	0.37693	0.48	-0.16	0.036
		RAC875_c99174_63	6A	43.1	0.00075	0.37693	0.47	0.16	0.036
		BS00067163_51	3D	130.1	0.00078	0.37693	0.43	0.14	0.035
		TA004394-0527	UM	-	0.00078	0.37693	0.12	0.24	0.035
		BobWhite_c1907_124	4B	62.6	0.00079	0.37693	0.16	-0.20	0.035
		BS00009915_51	4B	62.6	0.00079	0.37693	0.16	-0.20	0.035
		RAC875_c2253_1756	6A	37.1	0.00082	0.37693	0.47	0.16	0.035
		GENE-1785_118	UM	-	0.00082	0.37693	0.42	0.14	0.035
		tplb0060j17_879	6A	40.6	0.00084	0.37693	0.48	0.16	0.035
		RAC875_c41786_587	6A	43.1	0.00085	0.37693	0.47	0.16	0.035
		Kukri_c10377_88	6A	37.1	0.0009	0.37693	0.48	0.16	0.035
		GENE-2847_1060	UM	-	0.00095	0.37693	0.14	-0.23	0.034
		Kukri_c65146_460	4B	63.4	0.00097	0.37693	0.14	-0.22	0.034
		Kukri_c21270_1870	4B	59.9	0.00097	0.37693	0.14	0.22	0.034

Table 33. (Continued)

Trait	Env.	SNP	Chr.	Position	P	FDR	MAF†	Allelic effect	R^2
Gac	LIRF Wet	BobWhite_c19554_544	2B	107.5	4.84E-05	0.45243	0.05	0.42	0.052
		BS00088489_51	2B	108.0	0.00011	0.45243	0.05	0.39	0.046
		GENE-1125_32	UM	-	0.00011	0.45243	0.05	0.39	0.046
		BS00022970_51	7A	64.6	0.00018	0.45243	0.48	0.17	0.044
		D_GBUVHFX02JKG4A_54	2D	22.5	0.00018	0.45243	0.36	-0.18	0.043
		Tdurum_contig54925_202	2B	107.4	0.00019	0.45243	0.05	0.38	0.043
		Excalibur_c25471_225	7A	64.6	0.0002	0.45243	0.48	0.17	0.043
		RFL_Contig337_1432	2B	108.0	0.00042	0.84449	0.06	-0.35	0.039
		BS00023064_51	5B	122.0	0.00071	0.95409	0.21	-0.20	0.035
		BS00067163_51	3D	130.1	0.00074	0.95409	0.43	0.15	0.035
		Excalibur_c50584_358	5D	196.1	0.00084	0.95409	0.15	-0.22	0.035
		GENE-1785_626	UM	39.9	0.00087	0.95409	0.42	-0.15	0.034
		Tdurum_contig54925_415	2B	108.0	0.00092	0.95409	0.05	-0.33	0.034
		GENE-1785_118	UM	-	0.00097	0.95409	0.42	0.15	0.034
Gac	ARDEC Dry	RAC875_rep_c71149_148	2B	152.6	0.0002	0.9658	0.16	-0.23	0.044
		wsnp_CAP11_c1820_985143	2B	152.6	0.00032	0.9658	0.16	0.23	0.041
		Ku_c16249_315	2B	152.6	0.00048	0.9658	0.16	0.22	0.039
		Tdurum_contig42643_194	6A	140.9	0.00054	0.9658	0.11	0.25	0.038
		Excalibur_c5082_158	6A	140.9	0.00059	0.9658	0.11	0.25	0.038
Gac	ARDEC Wet	Kukri_c66671_183	6A	140.9	4.37E-05	0.70185	0.09	0.39	0.054
		GENE-2559_63	UM	-	0.00016	0.92887	0.19	-0.28	0.046
		Kukri_c40099_249	5B	144.3	0.00027	0.92887	0.20	-0.26	0.042
		RAC875_c45135_184	5B	144.3	0.00033	0.92887	0.20	0.25	0.041
		Excalibur_c36630_2194	4B	55.5	0.00034	0.92887	0.23	0.23	0.041
		Excalibur_c5082_158	6A	140.9	0.00035	0.92887	0.11	0.31	0.041
		BS00064884_51	4B	63.2	0.00042	0.92887	0.12	-0.32	0.040

Table 33. (Continued)

Trait	Env.	SNP	Chr.	Position	P	FDR	MAF†	Allelic effect	R^2
Gac	ARDEC Wet	Tdurum_contig42643_194	6A	140.9	0.00052	0.92887	0.11	0.30	0.038
		BS00065555_51	4B	63.2	0.00057	0.92887	0.13	0.30	0.038
		Excalibur_c26226_458	4B	59.9	0.00063	0.92887	0.13	-0.30	0.037
		wsnp_Ra_c15715_24192817	4B	59.9	0.00066	0.92887	0.13	-0.30	0.037
		D_F1BEJMU02GB94Z_188	2D	8.5	0.00087	0.92887	0.22	-0.23	0.035
		IAAV8654	4B	63.2	0.00088	0.92887	0.13	0.29	0.035
		TA004394-0527	UM	-	0.00088	0.92887	0.13	0.30	0.035
		Ex_c7593_754	4B	63.2	0.00089	0.92887	0.12	-0.30	0.035
		Ra_c59822_1439	4B	63.2	0.00096	0.92887	0.14	-0.28	0.035
Tw	LIRF Dry	BS00047114_51	3B	48.1	1.41E-05	0.09412	0.25	0.51	0.056
		BS00065934_51	3B	45.9	1.73E-05	0.09412	0.26	0.50	0.054
		RAC875_c5799_224	3B	45.9	1.76E-05	0.09412	0.22	0.53	0.054
		IAAV3851	3A	169.9	4.96E-05	0.19930	0.17	0.55	0.048
		RAC875_c43028_62	3B	47.1	0.00012	0.38586	0.13	-0.54	0.043
		Excalibur_c76665_98	2B	109.2	0.00015	0.39087	0.21	-0.52	0.042
		Excalibur_c56787_95	4B	58.1	0.00020	0.45162	0.43	0.43	0.041
		Excalibur_c24391_321	3B	45.9	0.00033	0.47825	0.18	0.48	0.038
		Tdurum_contig29467_99	6B	57.1	0.00034	0.47825	0.28	-0.47	0.037
		BS00010332_51	3B	46.8	0.00036	0.47825	0.13	-0.50	0.037
		CAP11_c3631_75	4B	59.9	0.00037	0.47825	0.36	-0.42	0.037
		wsnp_Ra_c25590_35156699	1B	75.1	0.00039	0.47825	0.16	0.52	0.037
		wsnp_Ex_c1495_2864718	1B	75.1	0.00039	0.47825	0.16	-0.52	0.037
		CAP7_c3847_204	1B	43.9	0.00047	0.52999	0.49	0.38	0.036
		CAP12_c590_307	1B	136.0	0.00053	0.52999	0.06	-0.77	0.035
		Excalibur_c38928_307	6B	57.1	0.00058	0.52999	0.36	-0.43	0.034
		CAP12_c424_402	1B	74.4	0.00062	0.52999	0.15	-0.54	0.034

Table 33. (Continued)

Trait	Env.	SNP	Chr.	Position	P	FDR	MAF†	Allelic effect	R^2
Tw	LIRF Dry	BS00078055_51	1B	75.1	0.00063	0.52999	0.17	-0.49	0.034
		wsnp_Ra_rep_c74497_72390803	2B	119.1	0.00063	0.52999	0.29	0.42	0.034
		Tdurum_contig11700_1247	6B	57.4	0.00075	0.60016	0.42	0.38	0.033
		BS00079611_51	2A	128.9	0.00086	0.61574	0.08	0.58	0.032
		BobWhite_c18540_351	2B	119.1	0.00098	0.61574	0.29	0.40	0.032
		Ku_c22990_969	7B	142.9	0.00099	0.61574	0.07	-0.69	0.032
		Ra_c7721_225	6B	57.1	0.00099	0.61574	0.36	0.41	0.032
		wsnp_Ra_rep_c108218_91556581	6B	57.1	0.00099	0.61574	0.36	-0.41	0.032
Tw	LIRF Wet	RFL_Contig1115_407	2B	161.4	7.58E-05	0.78829	0.20	-1.17	0.044
		BS00056645_51	2B	161.4	0.00011	0.78829	0.16	1.25	0.042
		RAC875_c5799_224	3B	45.9	0.00029	0.78829	0.22	0.94	0.036
		BS00012081_51	UM	3.8	0.00033	0.78829	0.34	0.92	0.036
		Excalibur_c55959_651	1D	37.3	0.00038	0.78829	0.21	0.98	0.035
		CAP7_c3847_204	1B	43.9	0.00041	0.78829	0.49	0.80	0.035
		D_GA8KES401DAEOJ_64	2D	34.1	0.00048	0.78829	0.07	1.39	0.034
		BS00059475_51	3B	144.7	0.00060	0.78829	0.47	0.84	0.033
		RAC875_c31358_214	UM	79.7	0.00061	0.78829	0.35	-0.87	0.033
		RAC875_c15844_348	UM	78.7	0.00072	0.78829	0.35	-0.85	0.032
		BS00047114_51	3B	48.1	0.00074	0.78829	0.25	0.84	0.032
		BobWhite_c14271_1379	1B	43.9	0.00074	0.78829	0.45	0.76	0.032
		BS00072840_51	2B	161.4	0.00074	0.78829	0.22	0.94	0.032
		BS00068182_51	1B	62.6	0.00083	0.78829	0.24	-0.89	0.031
		BS00072839_51	2B	161.4	0.00084	0.78829	0.22	0.93	0.031
		RAC875_c43028_62	3B	47.1	0.00085	0.78829	0.13	-1.00	0.031
		CAP11_c1820_244	2B	153.1	0.00092	0.78829	0.09	1.18	0.030
		RAC875_rep_c119322_54	2B	157.2	0.00092	0.78829	0.09	1.18	0.030
		Tdurum_contig91865_242	3A	66.5	0.00093	0.78829	0.46	0.77	0.030

Table 33. (Continued)

Trait	Env.	SNP	Chr.	Position	P	FDR	MAF†	Allelic effect	R^2
Tw	ARDEC Dry	BS00047114_51	3B	48.1	4.10E-05	0.29529	0.24	0.60	0.058
		Excalibur_c56787_95	4B	58.1	4.24E-05	0.29529	0.44	0.57	0.058
		BS00065934_51	3B	45.9	5.52E-05	0.29529	0.25	0.59	0.056
		CAP11_c3631_75	4B	59.9	7.46E-05	0.29957	0.36	-0.57	0.054
		RAC875_c48283_1574	4B	63.0	0.00016	0.48778	0.07	-0.91	0.049
		Kukri_c35140_75	4B	64.3	0.00020	0.48778	0.48	-0.51	0.047
		BS00021960_51	5B	129.8	0.00024	0.48778	0.20	0.59	0.046
		BS00075960_51	7B	143.2	0.00027	0.48778	0.20	0.62	0.045
		RAC875_c43028_62	3B	47.1	0.00027	0.48778	0.12	-0.67	0.045
		BS00087825_51	3B	53.0	0.00039	0.51269	0.19	0.54	0.043
		RAC875_c5799_224	3B	45.9	0.00040	0.51269	0.21	0.55	0.043
		Tdurum_contig81548_426	5B	129.8	0.00043	0.51269	0.20	-0.56	0.042
		Tdurum_contig83087_104	6A	140.7	0.00045	0.51269	0.28	0.51	0.042
		Kukri_c29414_105	4B	64.3	0.00046	0.51269	0.48	0.49	0.042
		Excalibur_c17607_542	4B	63.0	0.00048	0.51269	0.46	-0.47	0.042
		BS00010332_51	3B	46.8	0.00051	0.51269	0.12	-0.64	0.041
		RAC875_c27536_611	4B	54.6	0.00062	0.58143	0.08	0.75	0.040
		Tdurum_contig51087_573	7B	67.5	0.00077	0.64787	0.15	-0.61	0.039
		RAC875_c101563_102	4B	64.0	0.00080	0.64787	0.48	-0.47	0.038
		RAC875_c12495_1391	4B	64.0	0.00084	0.64787	0.50	0.47	0.038
		IAAV5175	4B	64.3	0.00085	0.64787	0.48	0.46	0.038
		wsnp_RFL_Contig4151_4728831	4B	64.0	0.00098	0.68325	0.49	0.46	0.037
		RAC875_rep_c70391_128	4B	64.0	0.00098	0.68325	0.49	-0.46	0.037
Tw	ARDEC Wet	BS00047114_51	3B	48.1	1.82E-05	0.15064	0.25	0.57	0.056
		BS00105846_51	1B	28.8	2.29E-05	0.15064	0.41	-0.52	0.055
		BS00065934_51	3B	45.9	2.81E-05	0.15064	0.26	0.55	0.053

Table 33. (Continued)

Trait	Env.	SNP	Chr.	Position	P	FDR	MAF†	Allelic effect	R^2
Tw	ARDEC Wet	Kukri_c16814_103	7B	133.6	7.95E-05	0.31918	0.22	-0.58	0.047
		Kukri_c11959_587	7B	133.6	0.00018	0.50572	0.23	-0.54	0.043
		Excalibur_c7338_563	7B	133.6	0.00021	0.50572	0.23	-0.53	0.042
		CAP7_c3847_204	1B	43.9	0.00029	0.50572	0.50	0.43	0.040
		GENE-0293_285	UM	21.6	0.00032	0.50572	0.42	0.40	0.039
		wsnp_Ex_c4923_8767234	3A	104.9	0.00033	0.50572	0.14	-0.61	0.039
		Excalibur_c56787_95	4B	58.1	0.00040	0.50572	0.44	0.45	0.038
		Excalibur_c1055_565	7B	133.6	0.00041	0.50572	0.25	-0.49	0.038
		BS00087825_51	3B	53.0	0.00044	0.50572	0.19	0.48	0.037
		D_GBUVHFX02FQKHM_48	3D	148.4	0.00060	0.50572	0.17	-0.54	0.036
		BS00012264_51	7B	133.6	0.00062	0.50572	0.24	0.47	0.035
		RAC875_c4693_554	7B	133.6	0.00062	0.50572	0.24	-0.47	0.035
		wsnp_Ex_c14027_21925404	3D	143.0	0.00064	0.50572	0.17	-0.55	0.035
		RAC875_c56101_368	UM	80.9	0.00066	0.50572	0.47	-0.43	0.035
		CAP11_c3631_75	4B	59.9	0.00069	0.50572	0.36	-0.44	0.035
		Kukri_c43464_89	3D	149.0	0.00071	0.50572	0.18	-0.53	0.035
		RAC875_c31922_138	3D	149.0	0.00075	0.50572	0.17	0.53	0.034
		RAC875_c61950_1644	3D	148.4	0.00075	0.50572	0.17	0.53	0.034
		Excalibur_c51312_218	3D	148.4	0.00075	0.50572	0.17	-0.53	0.034
		Kukri_c35140_75	4B	64.3	0.00082	0.50572	0.46	-0.42	0.034
		IACX11443	7B	133.6	0.00087	0.50572	0.25	-0.46	0.033
		BS00023023_51	7B	120.8	0.00087	0.50572	0.26	0.42	0.033
		IACX11047	7B	133.6	0.00090	0.50572	0.25	0.46	0.033
		Tdurum_contig19413_656	UM	112.8	0.00094	0.50572	0.47	-0.42	0.033

† SNP, single nucleotide polymorphism; FDR, false discovery rate; MAF, minor allele frequency; R^2, phenotypic variation explained by SNP.

Strong MTA with P <1.67x10^{-4}

Among the detected MTA, strong MTA with $P<1.67x10^{-4}$ are represented in Figure 11 (Manhattan plots) as SNP above the lower threshold line, alternatively in Figure 12 (QQ plot), and in Table 34. Among strong MTA, only two (5%), which were recorded for Tw, were observed in three environments and the remaining 38 MTA (95%) were detected in a single environment (Table 34). The strong MTA for Gac were detected on chromosomes 2B (2 in LIRF Wet), 4B (7 in LIRF Dry), 6A (1 in ARDEC Wet), and unmapped (1 in ARDEC Wet and 1 in LIRF Wet).

The strong MTA for Gpc were found on chromosomes 2A (2 in LIRF Dry), 2D (1 in LIRF Wet), and 4A (1 in ARDEC Dry and 1 in LIRF Dry), while the strong MTA for Gpd were observed on chromosomes 2B (1 in LIRF Wet), 4A (1 in LIRF Dry), and 7A (1 in LIRF Dry). SNP *BS00064369_51* on chromosome 4A was associated with both Gpc and Gpd in LIRF Dry.

The strong MTA for Gy were obtained on chromosomes 6A (2 in LIRF Dry) and unmapped (UM) (1 in ARDEC Wet). The strong MTA for Tw were detected on chromosomes 1B (1 in ARDEC Wet), 2B (1 in LIRF Dry and 2 in LIRF Wet), 3A (1 in LIRF Dry), 3B (2 in ARDEC Dry, 2 in ARDEC Wet, and 4 in LIRF Dry), 4B (3 in ARDEC Dry), and 7B (1 in ARDEC Wet). Two of the Tw SNP, *BS00047114_51* and *BS00065934_51* on chromosome 3B, were repeatedly observed in three environments, ARDEC Dry, ARDEC Wet, and LIRF Dry.

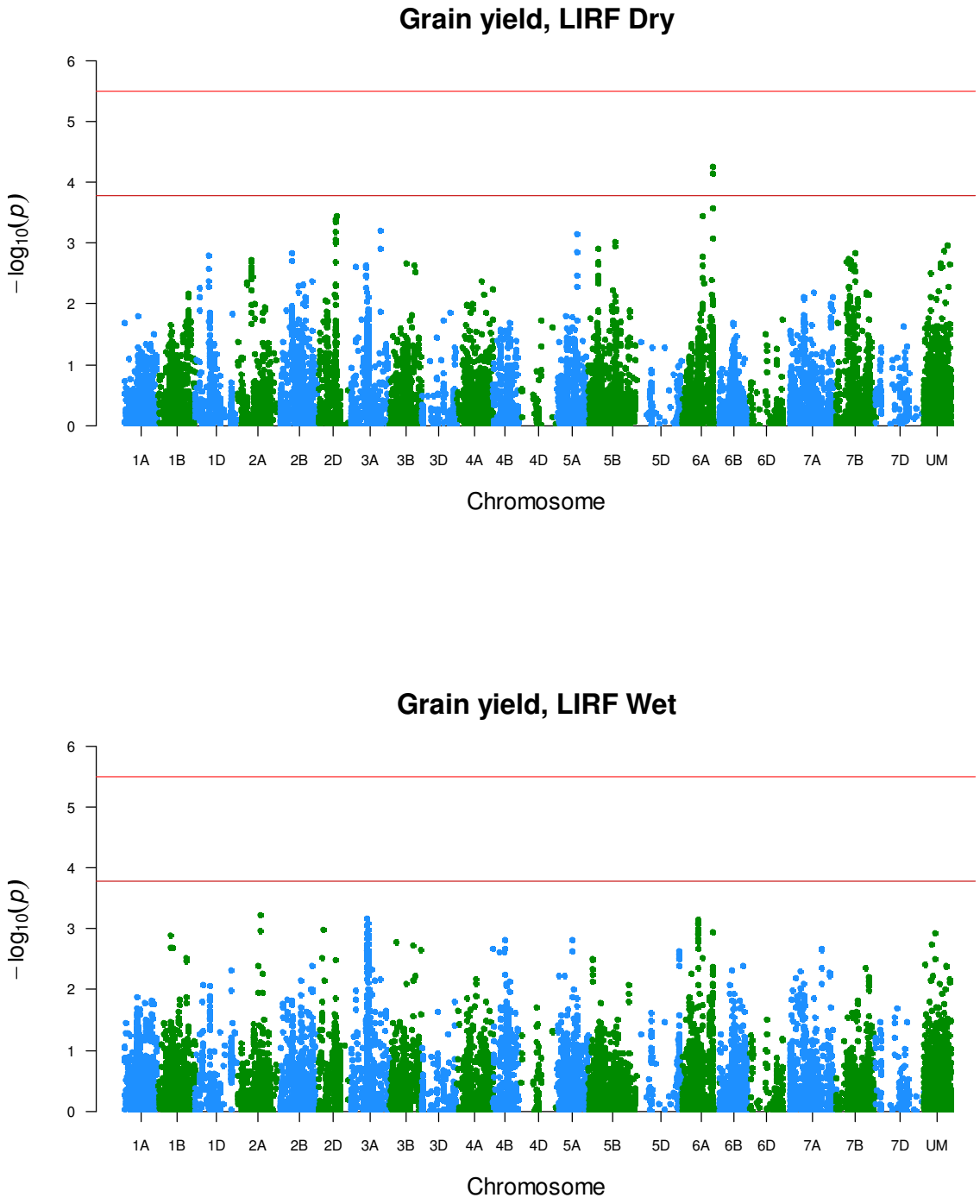


Figure 11. Manhattan plots for five traits of the Hard Winter Wheat Association Mapping Panel evaluated in four environments. The X-axis is the genomic position of each SNP; the Y-axis is the negative logarithm of the *P*-value obtained from the GWAS model. The lower line represents the significance threshold proposed by Gao et al. (2008) and the upper line is the Bonferroni significance threshold. UM, unmapped SNP.

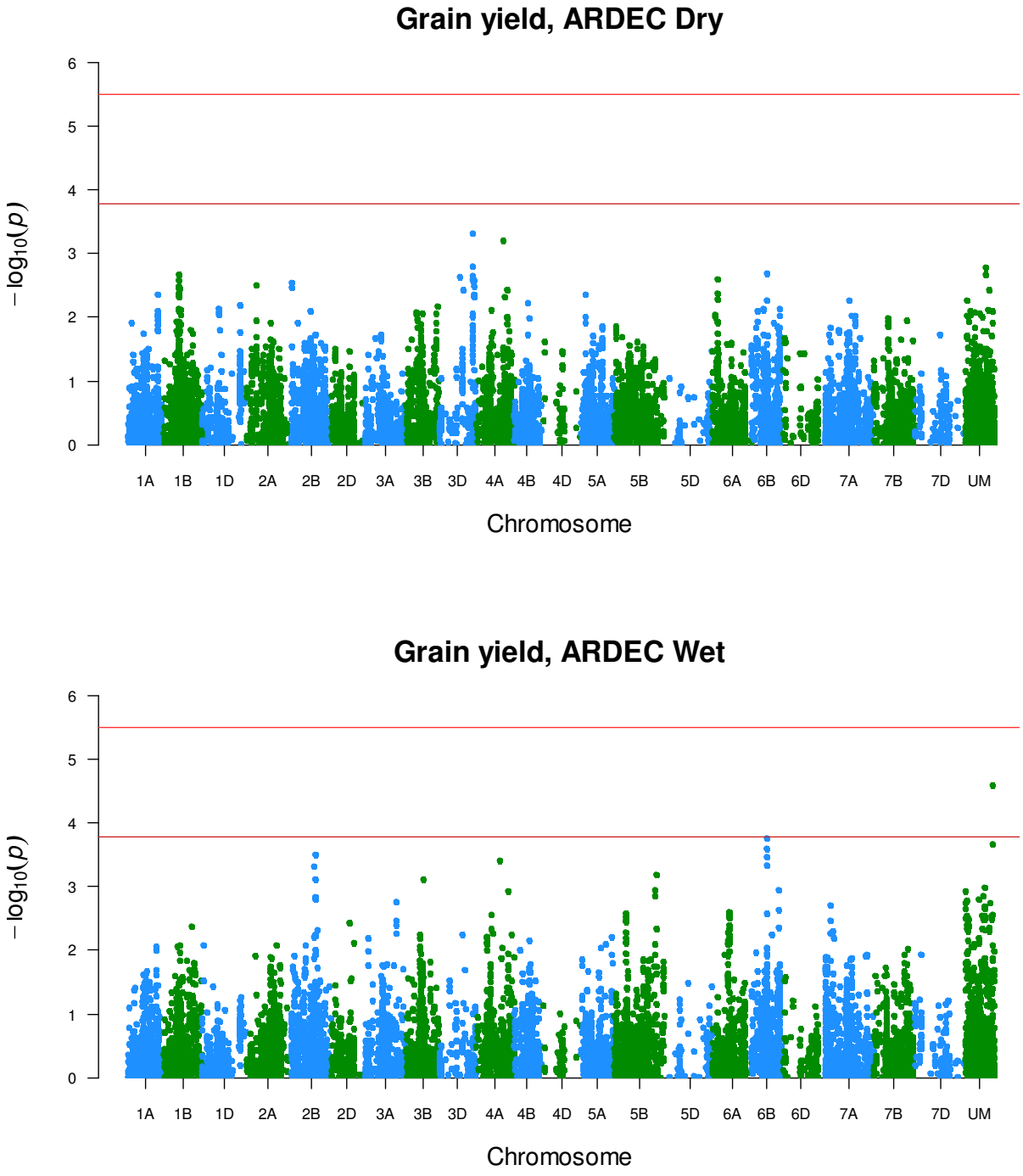


Figure 11. (Continued)

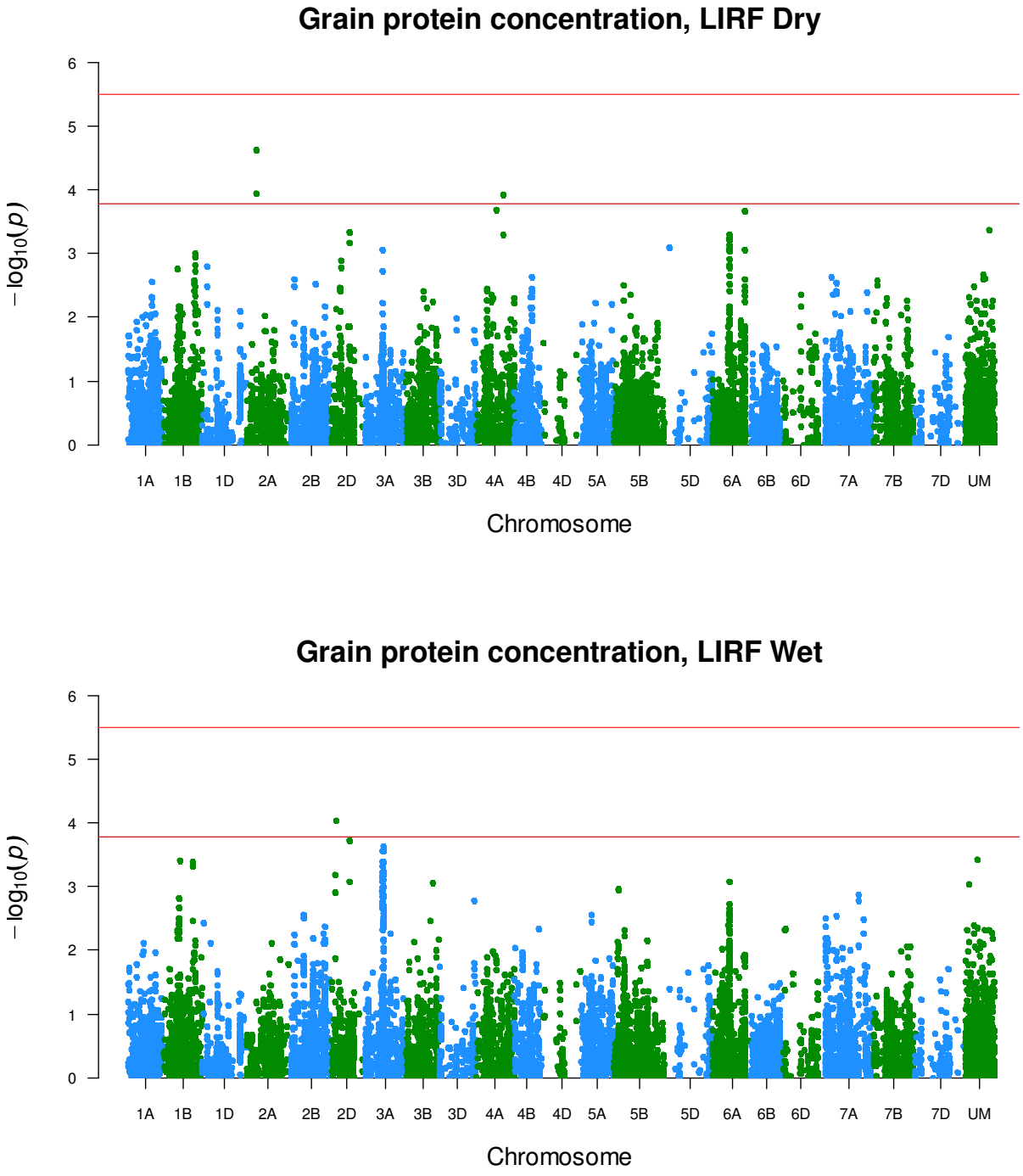
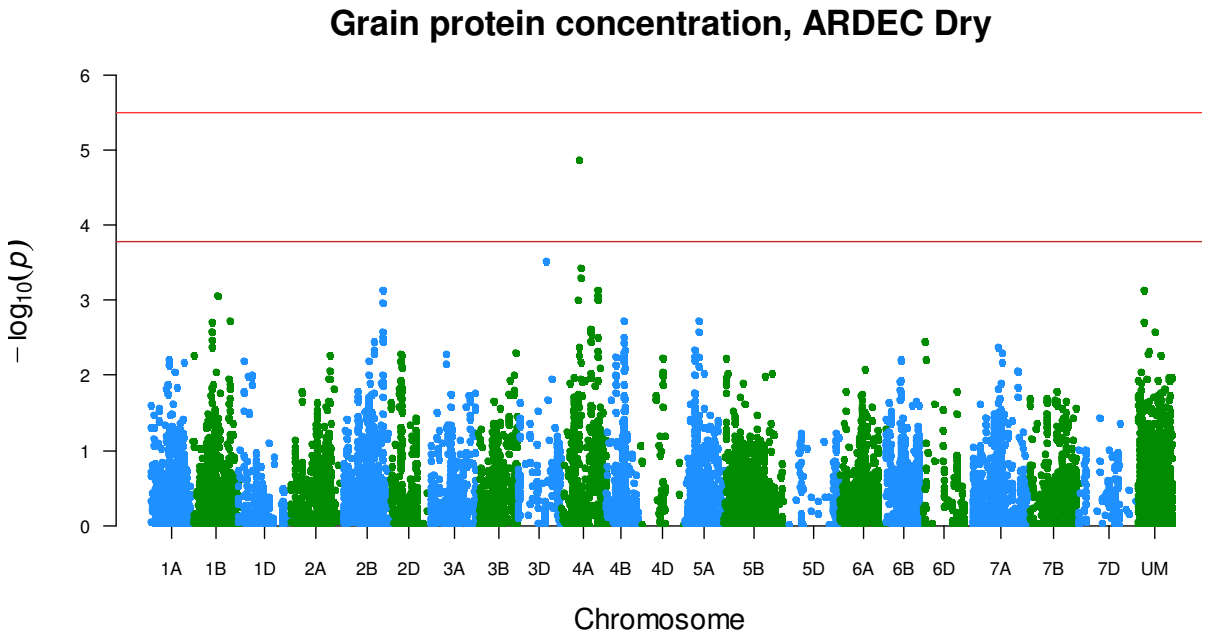


Figure 11. (Continued)



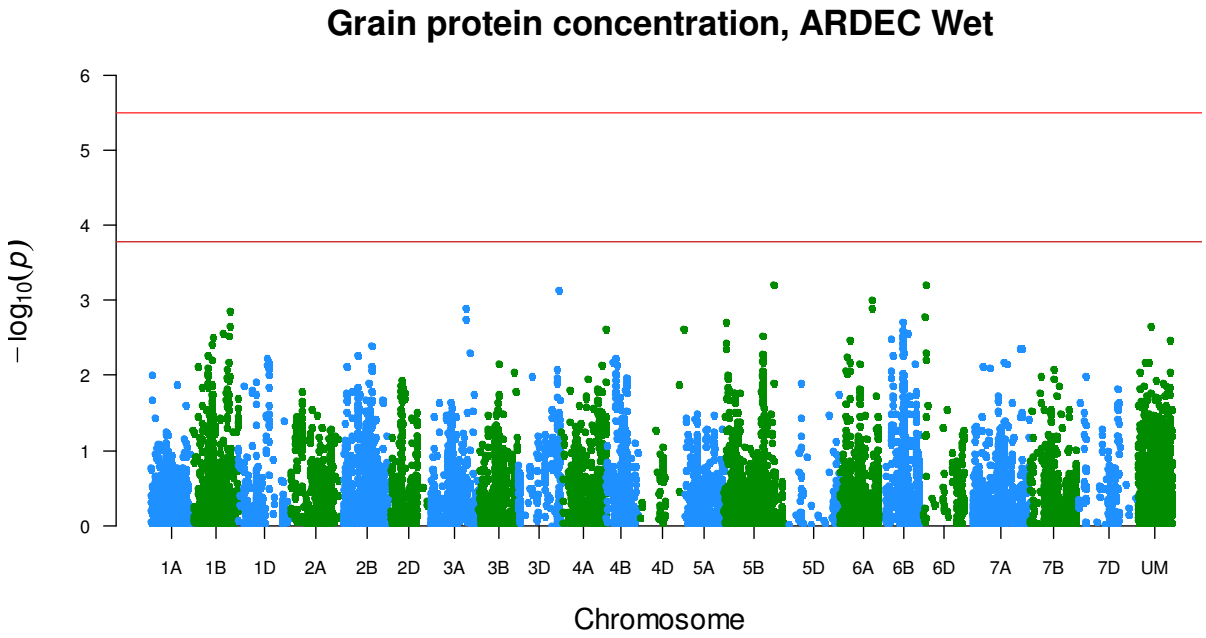
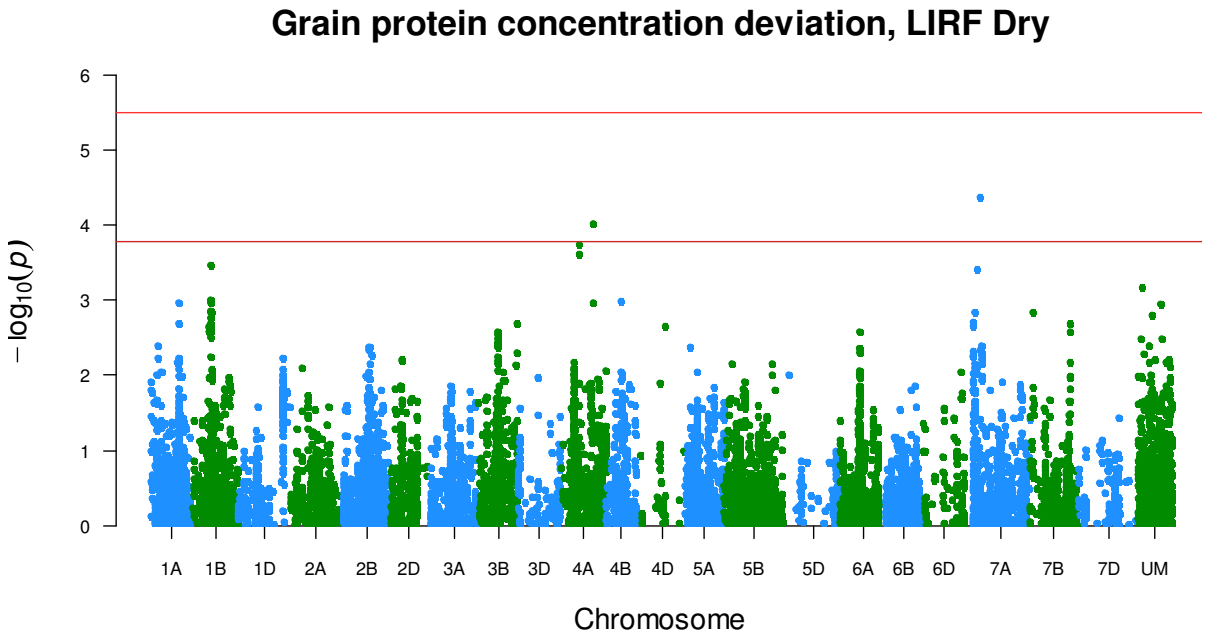


Figure 11. (Continued)



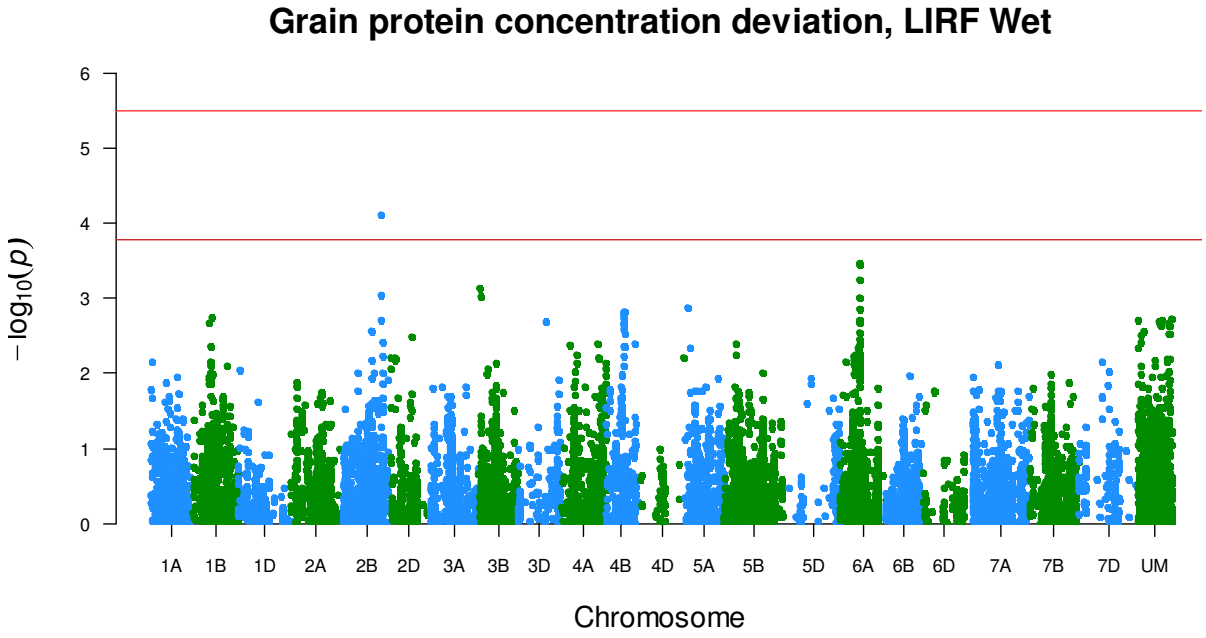
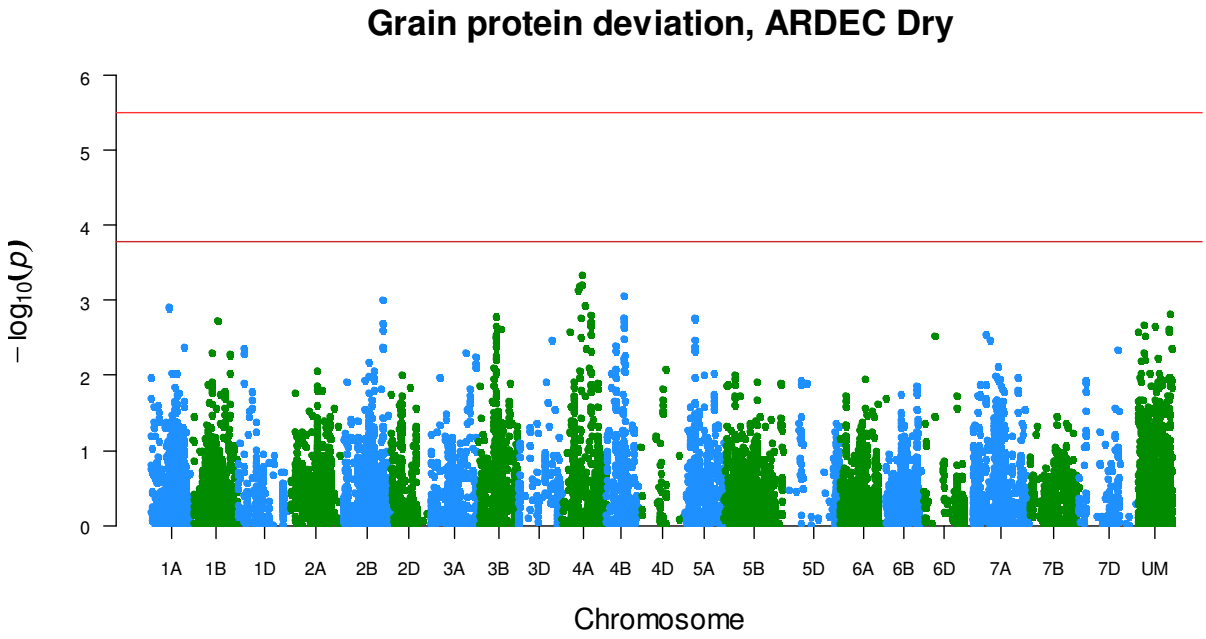


Figure 11. (Continued)



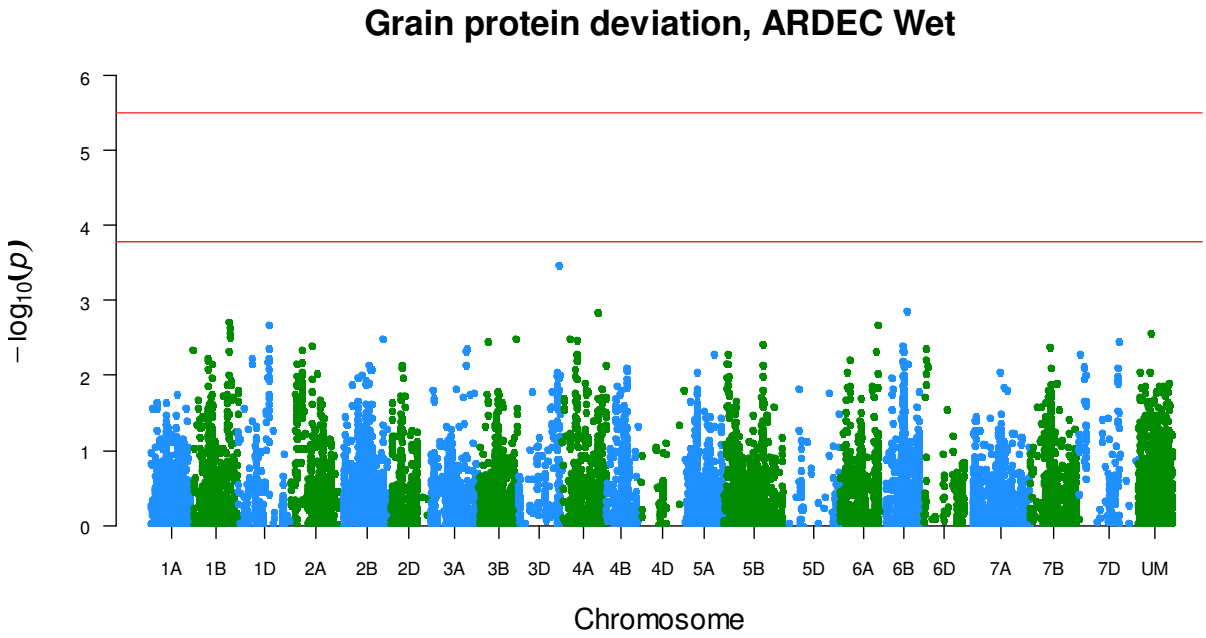
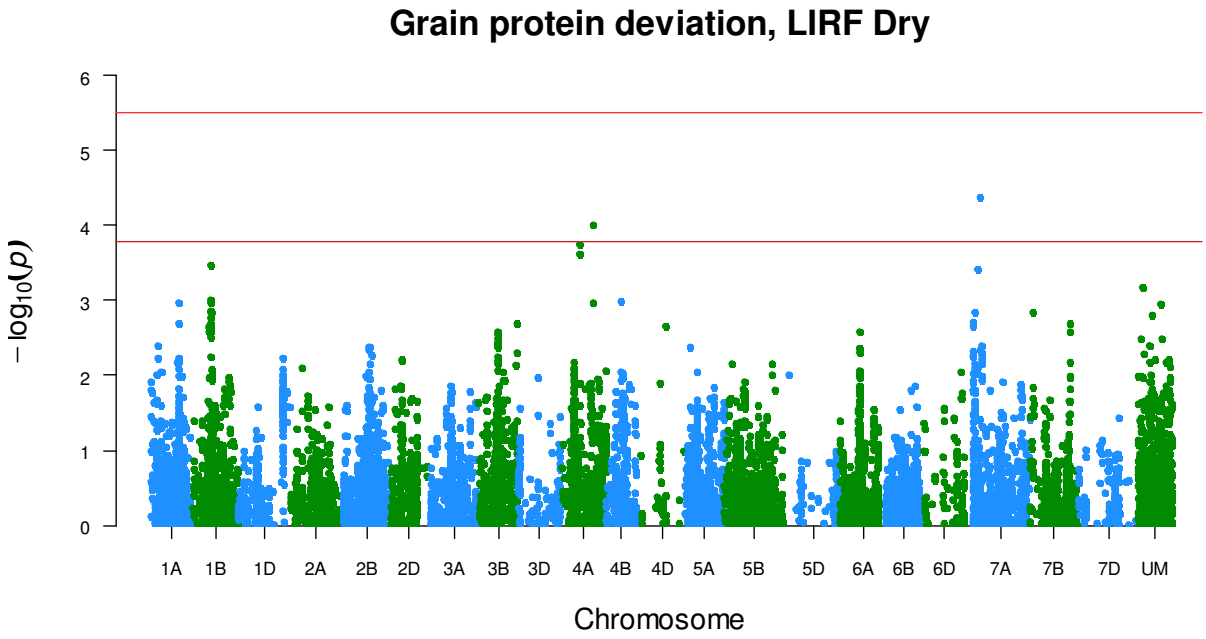


Figure 11. (Continued)



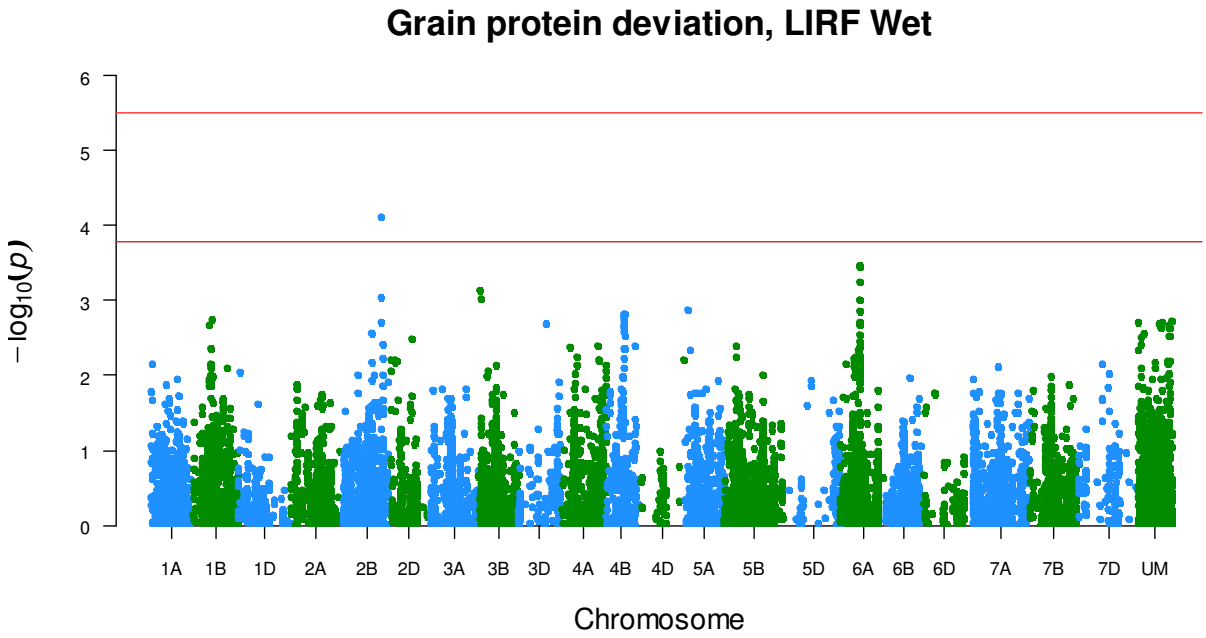
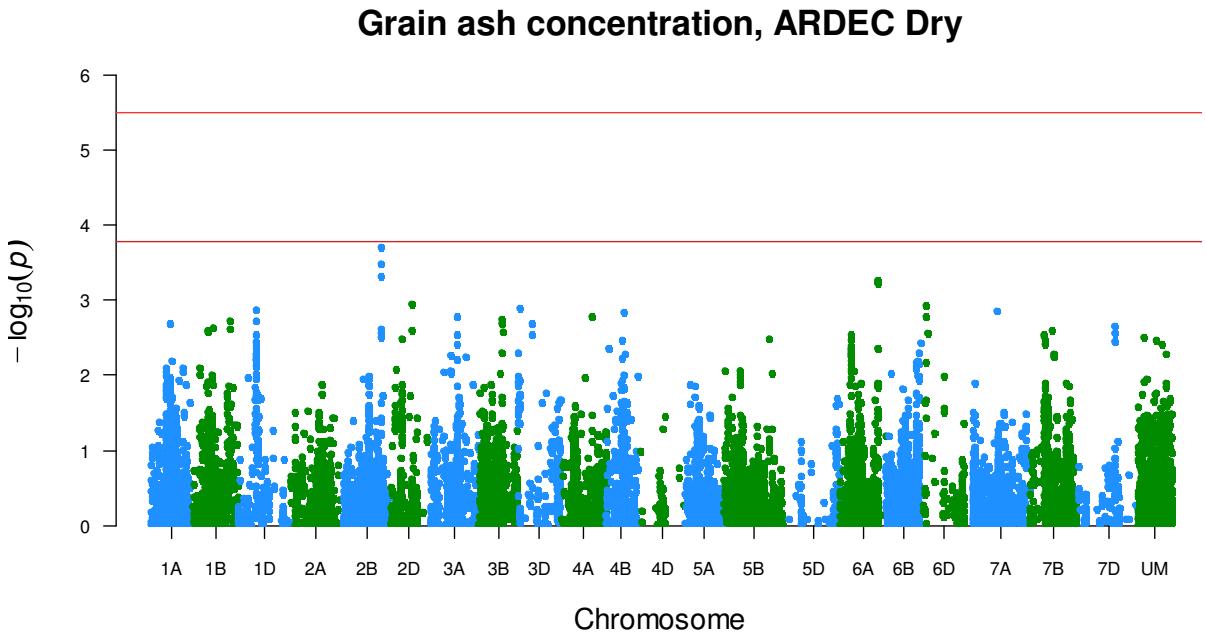


Figure 11. (Continued)



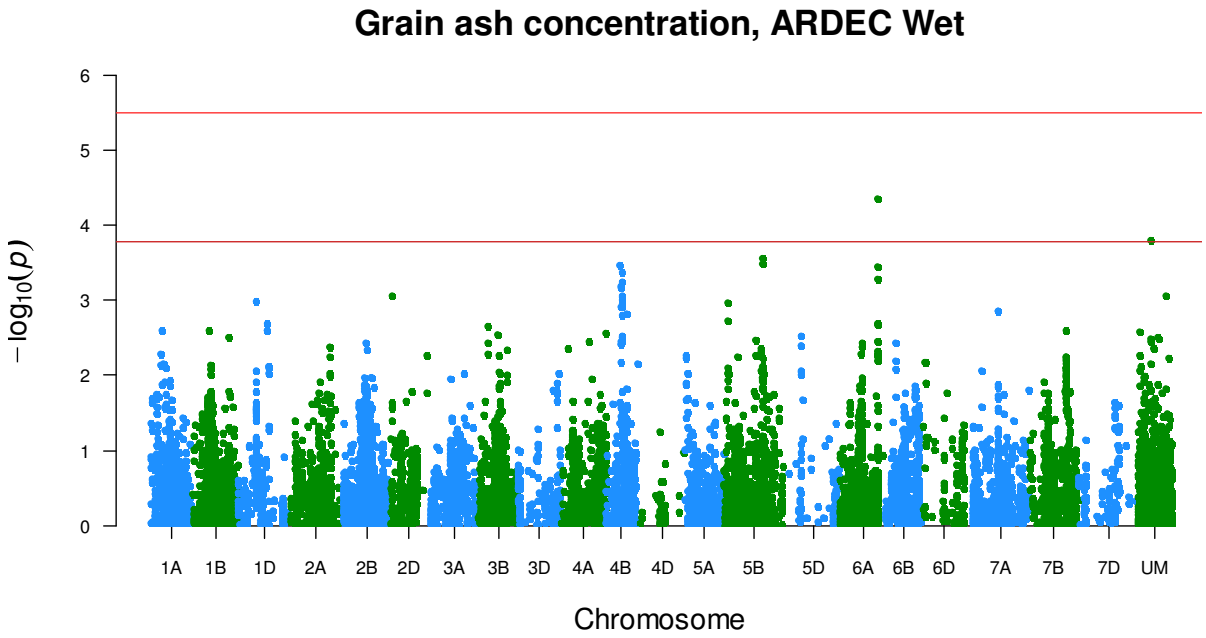
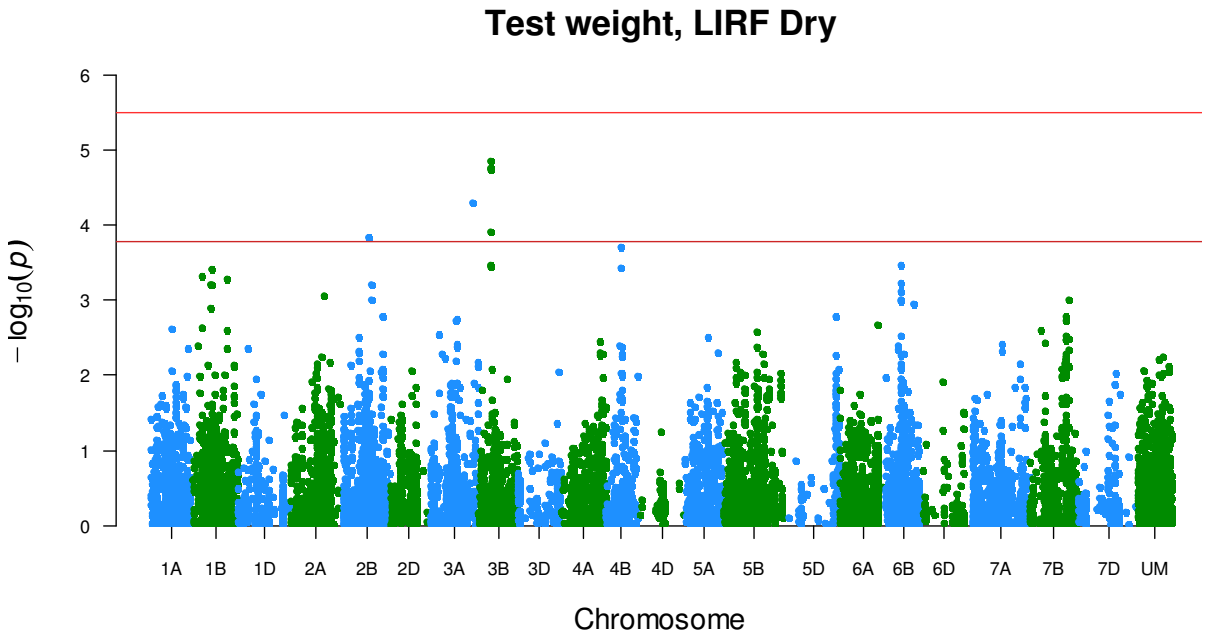


Figure 11. (Continued)



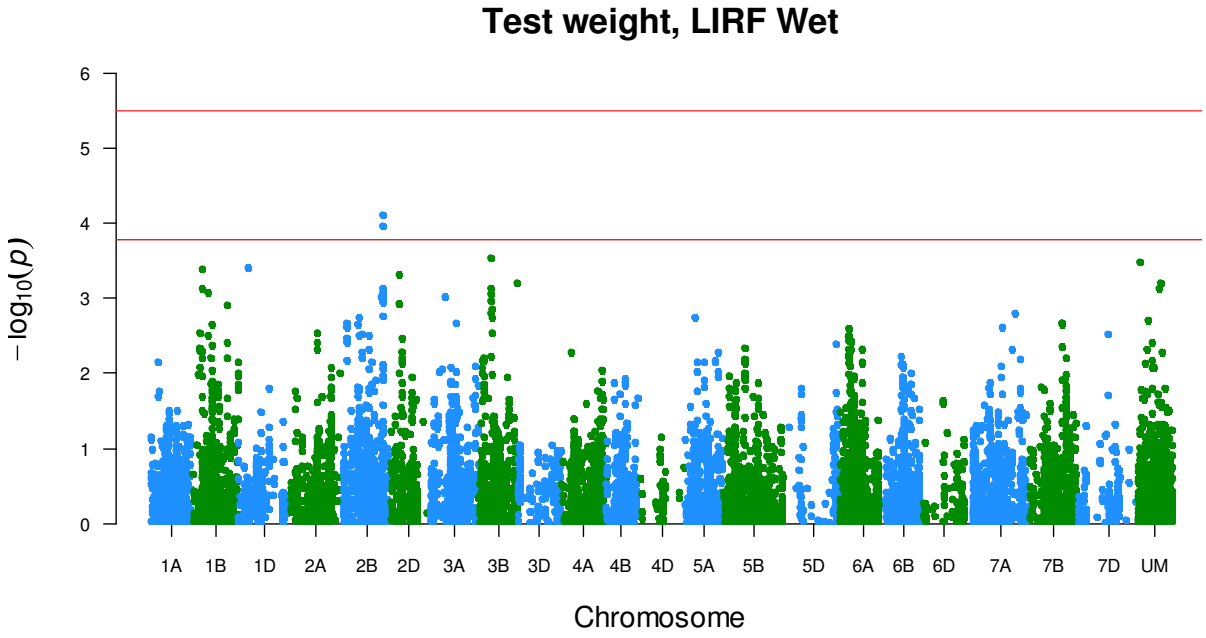
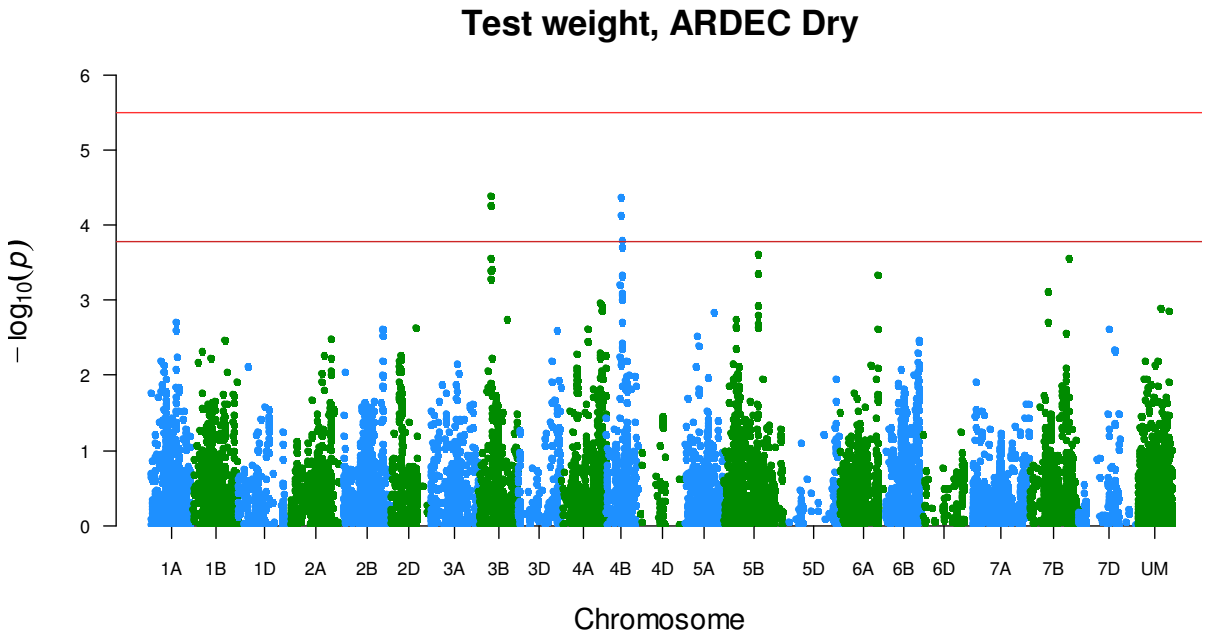


Figure 11. (Continued)



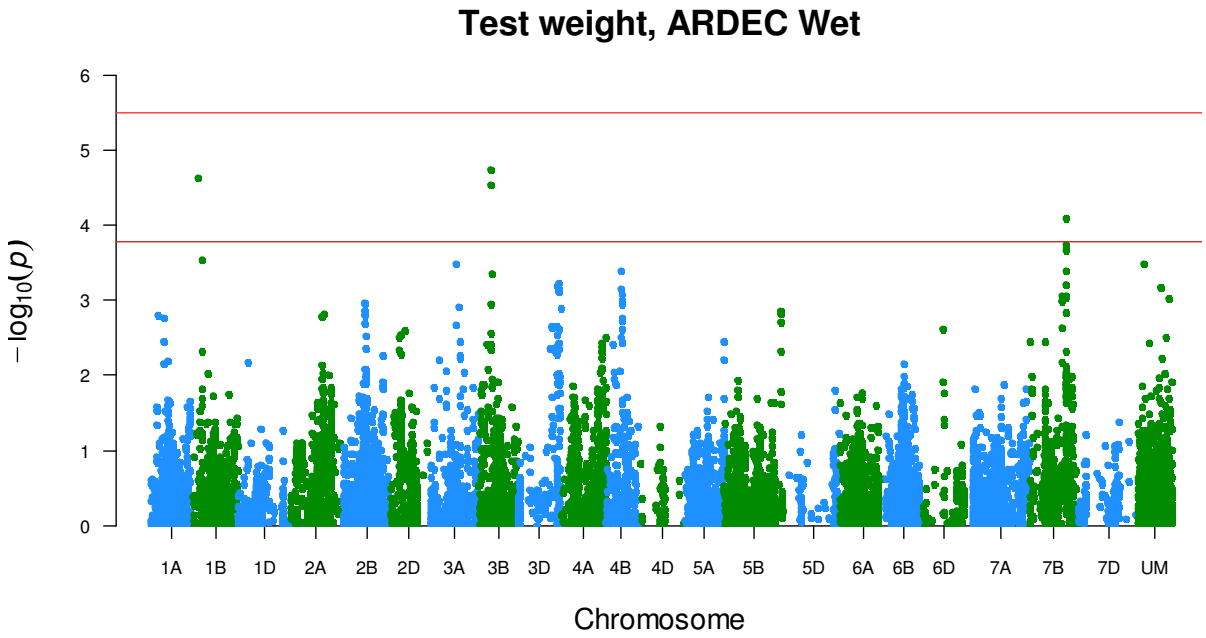


Figure 11. (Continued)

Grain yield

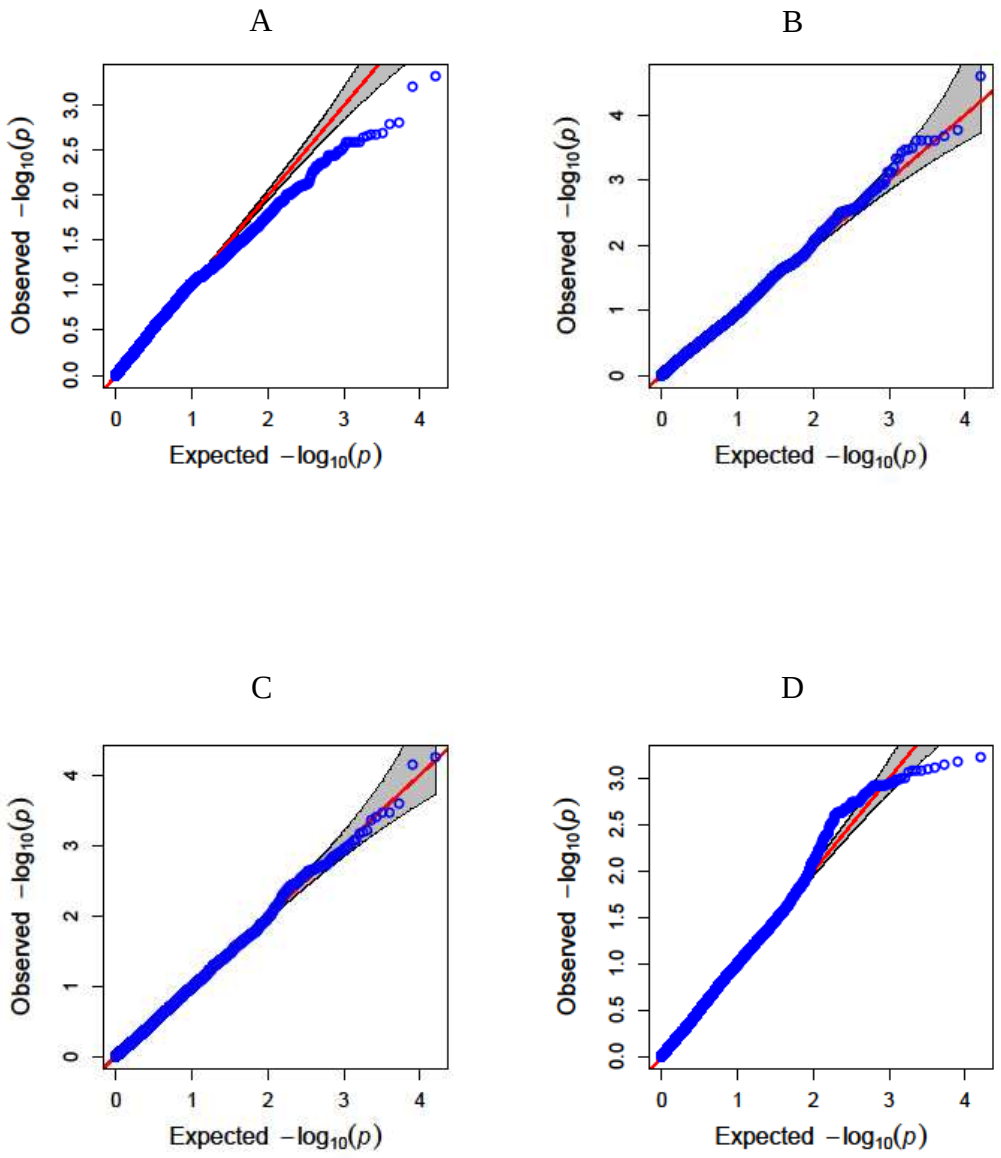


Figure 12. QQ plots for five traits investigated in four environments: A, ARDEC Dry; B, ARDEC wet; C, LIRF Dry; D, LIRF Wet.

Grain protein concentration

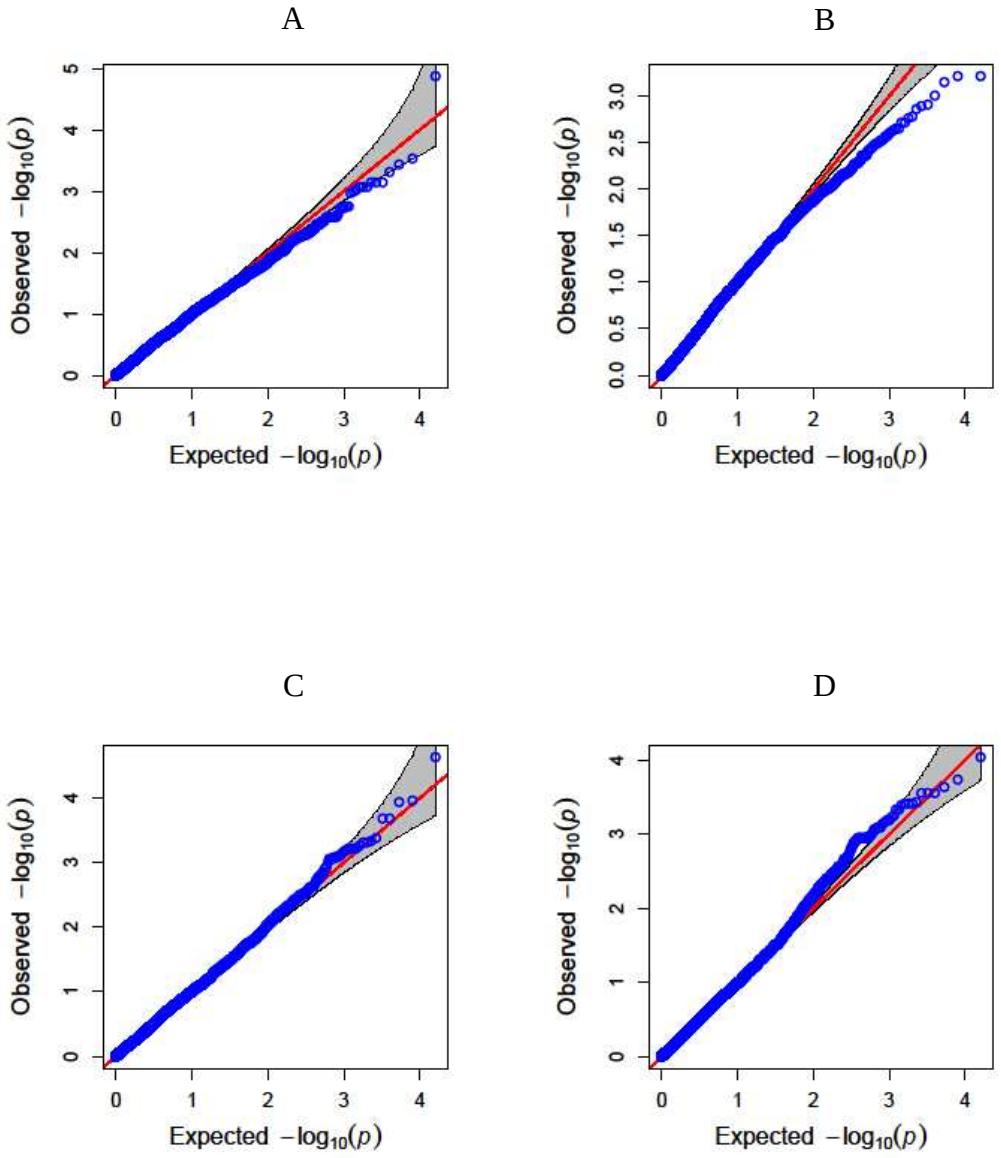


Figure 12. (Continued)

Grain protein deviation

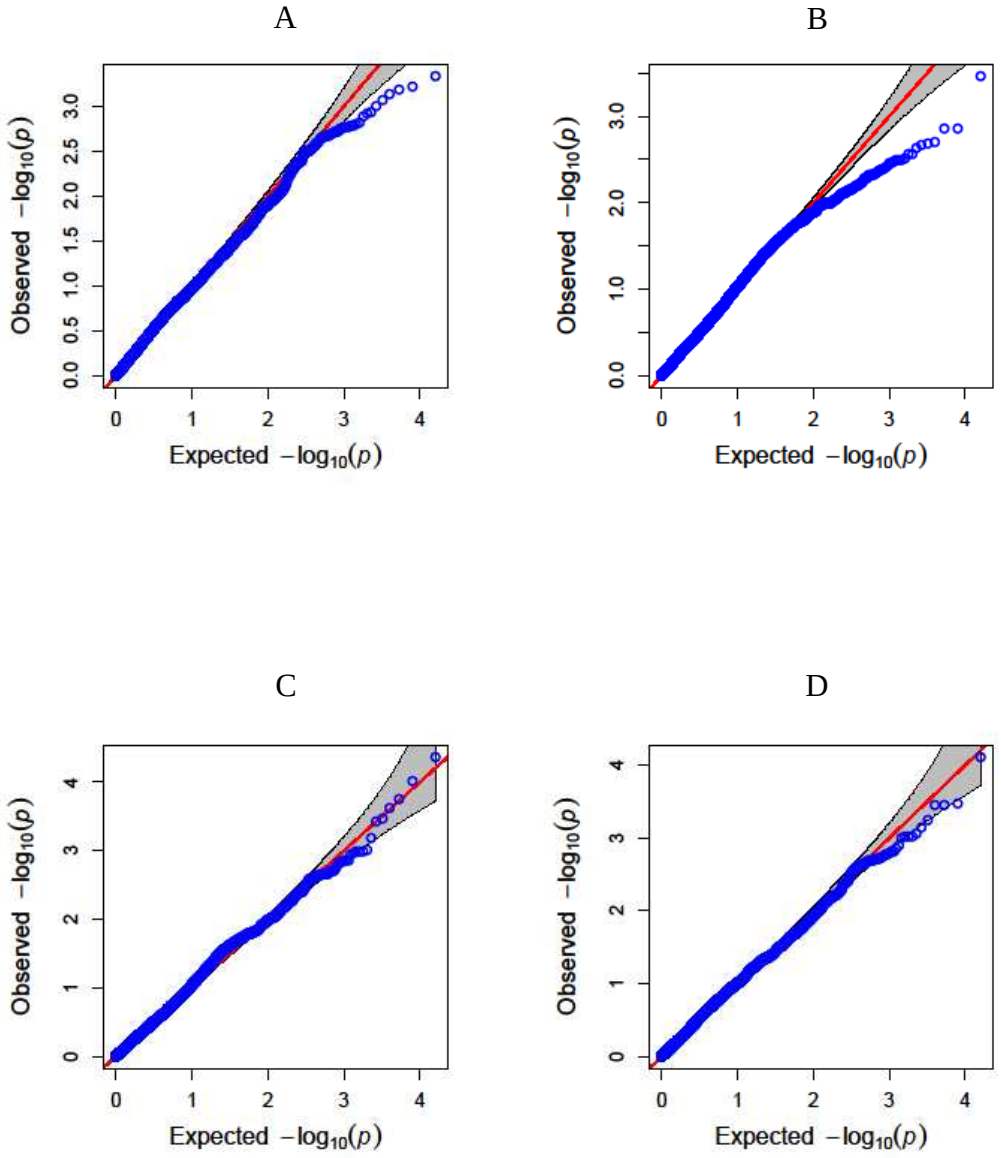


Figure 12. (Continued)

Grain ash concentration

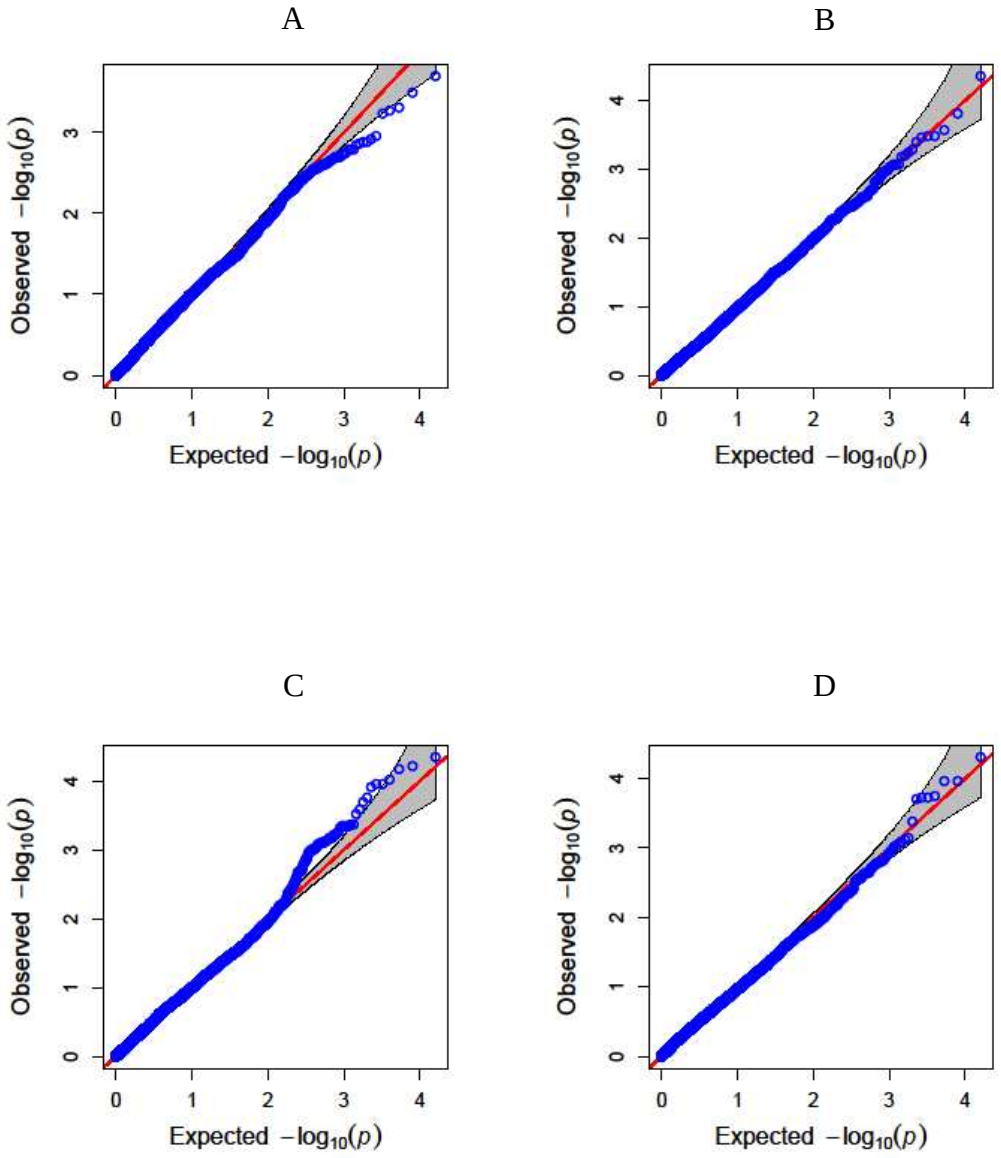


Figure 12. (Continued)

Test weight

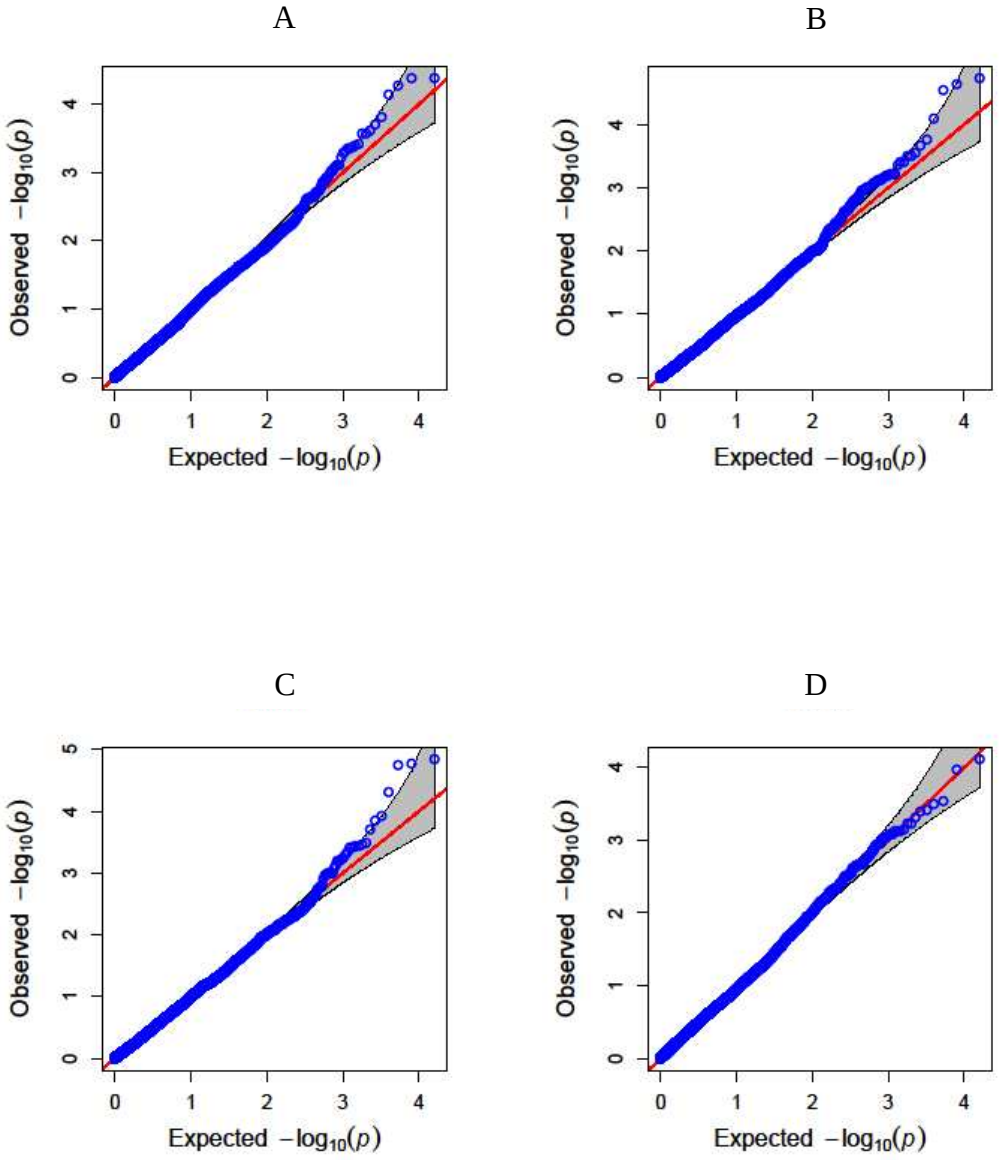


Figure 12. (Continued)

Table 34. Strong marker-trait associations in the Hard Winter Wheat Association Mapping detected at $P < 1.67 \times 10^{-4}$ for five traits in four environments.

Trait	Env.	SNP	Chr.	Position	*P*-value	FDR	MAF	Allelic Effect	R^2
Gy	ARDEC Wet	wsnp_Ex_c29382_38422739	UM	-	2.50×10^{-5}	0.40116	0.33	-253.85	0.061
	LIRF Dry	BS00021704_51	6A	140.9	5.51×10^{-5}	0.57096	0.20	130.75	0.052
		Excalibur_c4518_2931	6A	140.9	7.11×10^{-5}	0.57096	0.21	-125.91	0.050
Gpc	ARDEC Dry	wsnp_Ex_c19207_28125072	4A	69.6	1.32×10^{-5}	0.21186	0.17	1.98	0.062
	LIRF Dry	BS00021706_51	2A	47.2	2.30×10^{-5}	0.36967	0.37	-2.17	0.048
		Kukri_c14598_614	2A	47.2	0.00011	0.58118	0.38	-2.00	0.040
		BS00064369_51	4A	118.7	0.00012	0.58118	0.28	2.11	0.040
	LIRF Wet	D_GBUVHFX02JKG4A_54	2D	22.5	9.01×10^{-5}	0.44258	0.36	-2.33	0.040
Gpd	LIRF Dry	wsnp_CAP11_rep_c4066_1921894	7A	62.7	4.27×10^{-5}	0.68503	0.09	-3.01	0.053
		BS00064369_51	4A	118.7	9.65×10^{-5}	0.72781	0.28	1.80	0.048
	LIRF Wet	BS00039187_51	2B	152.6	7.64×10^{-5}	0.90722	0.31	1.96	0.052
Gac	ARDEC Wet	Kukri_c66671_183	6A	140.9	4.37×10^{-5}	0.70185	0.09	0.39	0.054
		GENE-2559_63	UM	-	0.00016	0.92887	0.19	-0.28	0.046
	LIRF Dry	wsnp_Ra_c15715_24192817	4B	59.9	4.45×10^{-5}	0.27892	0.13	-0.29	0.053
		BS00064935_51	4B	63.2	6.00×10^{-5}	0.27892	0.13	0.29	0.051
		Excalibur_c26226_458	4B	59.9	6.71×10^{-5}	0.27892	0.12	-0.29	0.050
		IAAV8654	4B	63.2	9.50×10^{-5}	0.27892	0.12	0.28	0.048
		BS00064884_51	4B	63.2	0.00011	0.27892	0.11	-0.28	0.047
		Ex_c7593_754	4B	63.2	0.00011	0.27892	0.12	-0.28	0.047
		BS00065555_51	4B	63.2	0.00012	0.27892	0.13	0.27	0.047
	LIRF Wet	BobWhite_c19554_544	2B	107.5	4.84×10^{-5}	0.45243	0.05	0.42	0.052
		BS00088489_51	2B	108.0	0.00011	0.45243	0.05	0.39	0.046
		GENE-1125_32	UM	-	0.00011	0.45243	0.05	0.39	0.046

Table 34. (Continued)

Trait	Env.	SNP	Chr.	Position	*P*-value	FDR	maf	Allelic Effect	R^2
Tw	ARDEC Dry	BS00047114_51	3B	48.1	4.10×10^{-5}	0.29529	0.24	0.60	0.058
		BS00065934_51	3B	45.9	5.52×10^{-5}	0.29529	0.25	0.59	0.056
		Excalibur_c56787_95	4B	58.1	4.24×10^{-5}	0.29529	0.44	0.57	0.058
		CAP11_c3631_75	4B	59.9	7.46×10^{-5}	0.29957	0.36	-0.57	0.054
		RAC875_c48283_1574	4B	63.0	0.00016	0.48778	0.07	-0.91	0.049
	ARDEC Wet	BS00105846_51	1B	28.8	2.29×10^{-5}	0.15064	0.41	-0.52	0.055
		BS00047114_51	3B	48.1	1.82×10^{-5}	0.15064	0.25	0.57	0.056
		BS00065934_51	3B	45.9	2.81×10^{-5}	0.15064	0.26	0.55	0.053
		Kukri_c16814_103	7B	133.6	7.95×10^{-5}	0.31918	0.22	-0.58	0.047
	LIRF Dry	Excalibur_c76665_98	2B	109.2	0.00015	0.39087	0.21	-0.52	0.042
		IAAV3851	3A	169.9	4.96×10^{-5}	0.1993	0.17	0.55	0.048
		BS00047114_51	3B	48.1	1.41×10^{-5}	0.09412	0.25	0.51	0.056
		BS00065934_51	3B	45.9	1.73×10^{-5}	0.09412	0.26	0.50	0.054
		RAC875_c5799_224	3B	45.9	1.76×10^{-5}	0.09412	0.22	0.53	0.054
		RAC875_c43028_62	3B	47.1	0.00012	0.38586	0.13	-0.54	0.043
	LIRF Wet	RFL_Contig1115_407	2B	161.4	7.58×10^{-5}	0.78829	0.20	-1.17	0.044
		BS00056645_51	2B	161.4	0.00011	0.78829	0.16	1.25	0.042

Env., environment; SNP, single nucleotide polymorphism; Chr, chromosome; FDR, false discovery rate; MAF, minor allele frequency; R^2, phenotypic variation explained by SNP.

Number of MTA detected at different level of significance

A summary of the number of MTA detected for five evaluated traits in all environments is given in Table 35. Considering the criterion of $P<0.001$, there were 307 significant MTA detected in the four environments for the five traits. Meanwhile, considering the criterion of $P < 1.67 \times 10^{-4}$, there were only 40 significant MTA detected in the four environments for the five traits. The highest number of MTA was recorded for Tw (17), followed by Gac (12) and Gpc (5), with the least for Gy (3) and Gpd (3). This suggests that different statistical approaches can enormously affect the number of significant MTA.

Table 35. Number of MTA detected at $P< 0.001$ and at $P < 1.67 \times 10^{-4}$.

Traits	**Number of MTA detected**			
	MTA ($P< 0.001$)	**MTA ($P < 1.67 \times 10^{-4}$)**	**Bonferroni ($P < 3.11 \times 10^{-6}$)**	**FDR < 0.05**
Grain yield	43	3	0	0
Grain protein concentration	69	5	0	0
Grain protein deviation	26	3	0	0
Grain ash concentration	75	12	0	0
Test weight	94	17	0	0
Total	307	40	0	0

Allelic effect

The range of allelic effects of MTA ($P<0.001$) was approximately -443 to 446 kg ha^{-1} for Gy, -3.6 to 4.4 g kg^{-1} for Gpc, -3.0 to 2.2 for Gpd, -0.35 to 0.42 g kg^{-1} for Gac, and -1.2 to 1.4 kg hL^{-1} for Tw (Table 33). The range of allelic effects of strong MTA ($P < 1.67 \times 10^{-4}$) for Gy, Gpc, Gpc, Gac, and Tw was approximately -254 to 131 kg ha^{-1}, -2.3 to 2.1 g kg^{-1}, -3.1 to 2.0, -0.29 to 0.42 g kg^{-1}, and -1.2 to 1.3 kg hL^{-1}, respectively (Table 34).

Phenotypic variation explained by MTA was approximately the same among SNP. The values for Gy, Gpc, Gpc, Gac, and Tw were 3.0 to 6.0%, 3.0 to 6.0%, 3.0 to 5.0%, 3.0 to 5.0%, and 3.0

to 6.0%, respectively (Table 33). Similarly, phenotypic variation explained by strong MTA was approximately the same among SNP, more or less 5% (Table 34). This suggests that the SNP had almost the same proportion of effect, thus, making it challenging to identify the most promising candidate genes for these evaluated traits.

4.3. DISCUSSION

Trait means, correlation, and heritability

The mapping panel used in this study was a collection of 299 genotypes representative of the U.S. Great Plains region developed by public and private breeders. The purpose of this study was to characterize the HWWAMP with the intention of identifying MTA/genes underlying grain quality traits under contrasting soil moisture levels.

The measured phenotypic traits responded differently to water deficit. Grain yield was reduced about 50% under drought conditions. This finding agrees with the previous report (Chapter 2 of this dissertation) and other studies (El-Feki et al. 2013; Edae et al. 2014). Water deficit during grain filling reduces sucrose and starch accumulation in grain (Ahmadi & Baker 2001), resulting in grain yield reduction.

Under drought condition, Gpc increased. This result agrees with many previous reports (El-Feki et al. 2013; Balla et al. 2011). Foulkes et al. (2001) reported that the restricted availability of water by late grain filling consistently decreased green canopy area and that drought after flowering accelerated canopy senescence. Drought stress changes the rates of the synthesis of starch components and starch accumulation (Ahmadi & Baker 2001; Jenner & Rathjen 1975). The reduced sucrose and starch content of wheat grain (Ahmadi & Baker 2001) results in higher

final grain protein concentration (Weightman et al. 2008). In addition, drought and heat stresses are two events that often occur simultaneously in the field (Aprile et al. 2013). Triboï et al. (2003) reported that the rate of N (indication of protein) accumulation increased with an increase in post-anthesis temperature. It is suggested that under drought conditions, Gpc increases mainly due to higher rates of accumulation of grain nitrogen and lower rates of accumulation of carbohydrates.

Grain yield and Gpc are two major targets in wheat breeding programs. Grain protein deviation, derived from regression between Gpc and Gy, could be used to identify genotypes having higher or lower Gpc than expected based on Gy (Monaghan et al. 2001). The mean post-anthesis N uptake was notably and significantly associated with mean Gpd (Bogard et al. 2010). In the current study, Gpd showed the same trend as Gpc in responding to water stress. Both Gpd and Gpc were higher in the dry treatments compared to the wet treatments. Therefore, the possibility of using Gpd, the deviation from the Gpc–Gy relationship, was evaluated. Guttieri et al. (2015) suggested that Gpd was repeatable across environments and selection for increased Gpd may be effective.

Grain ash, consisting of the inorganic materials or minerals, is a minor constituent of wheat grain. Grain ash is mainly concentrated in the bran (Shelton et al. 2008). High heat and water stress during grain filling increased Gac (Ahmed & Fayyaz-ul-Hassan 2015). Gac was positively correlated with the drought susceptibility index under severe water deficit (Bogale & Tesfaye 2011). Bogale & Tesfaye (2011) also suggested that drought-susceptible wheat genotypes in a field experiment accumulated more minerals in the kernels during grain-filling.

Correlations between pairs of evaluated traits were analyzed for all four environments (Table 30 and 31). Grain yield was significantly and negatively correlated with Gpc and Gac. The finding

agreed with previous reports (El-Feki et al. 2013; Bilgin et al. 2010; Hrušková & Švec 2009). In contrast, Gy has a positive correlation with Tw. Bilgin et al. (2010) found a positive correlation between these traits (r=0.381, P<0.001). In the current study, Gpc, Gpd, and Gac are positively correlated with each other in all environments. These findings agree with two previous studies in the RIL and BC populations (Chapter 2 and 3 of this dissertation). El-Feki et al. (2013) found that Gpc and Gac had a significantly positive correlation that was stable across four environments (r=0.31, 0.43, 052, and 0.59) in the CO940610/Platte DH population.

Heritability estimates of the five investigated traits are represented in Table 32. The abnormal heritability estimates were observed for Gy (-0.31), Gpc (0.08) and Gpd (-0.06) in ARDEC Dry. This might be caused by experimental errors and/or estimation method. The entries in these experiments were not replicated, with the exception of two checks (Hatcher and Settler CL) with 15 replicates for each. The environmental variance was determined based on the variation of each check. The heritability estimates for all evaluated traits varied highly over environments, from low to high. Except for the abnormal values stated above, the lowest estimate was for Gpd (0.41) in ARDEC Wet and the highest for Gpc (0.92) in LIRF Wet. The variation in estimates is probably caused by environmental factors as well as the estimation method used, which relies heavily on the phenotypic variance of just two cultivars. The high heritability estimates, however, indicates that a large part of the expression of these traits is genetically controlled, making progress from selection easier to achieve in a breeding program. Clarke et al. (2009) reported that heritability estimates among environments for Gpc ranged from 0.43 to 0.84, thus suggesting that Gpc was moderately to highly heritable and complexly inherited. Barnard et al. (2002) estimated board-sense heritability of Gpc at 0.57, concluding the inheritance of wheat quality traits is complex due to the polygenic control of most of these characteristics. However,

the heritability estimates of the traits in this current study varied from environment to environment, indicating more environmental influence on heritability estimates of the traits. Therefore, this study suggested that the selection of environment would influence the success of breeding for improved traits. Heritability estimates are influenced by the number of replications, years and locations tested, and by the genetics of the populations (Clarke et al. 2009).

Marker-trait associations

Grain yield and Gpc have an inverse relationship (Wang et al. 2012; Kibite & Evans 1984; Pleijel 1999). Of three markers associated with Gy, two are located on chromosome 6A. In the previous study of the RIL population (Chapter 2 of this dissertation), only markers on chromosomes 1B and 7B were detected in association with Gy. In a spring wheat association mapping panel grown at LIRF and in Ethiopia, Edae et al. (2014) found that Gy MTA were detected only on chromosomes 1BS, 2DS, 5B (73 and 76.4 cM) and 7B. However, other researchers have found QTL for grain yield located on chromosome 6A (Heidari et al. 2011; Kirigwi et al. 2007; Simmonds et al. 2014).

Only four SNP on chromosomes 2A, 2D, and 4A were highly significantly ($P < 1.67x10^{-4}$) associated with Gpc in the current study. These MTA were not repeated across environments. In the previous study in the RIL population (Chapter 2 of this dissertation), there was only one marker *Xgwm569* on chromosome 7B detected (*P*-value = 0.075) that was associated with Gpc. El-Feki et al. (2013) detected QTL for Gpc on chromosomes 5B, 6A, 6B, 7B, and 7D in the DH population. This indicates that the MTA detected for Gpc did not consistently occur across populations and environments. However, many others have found QTL associated with Gpc on chromosomes 2A, 2D, and 4A (Wang et al. 2012; Prasad et al. 2003; Sun et al. 2008; Turner et al. 2004; Elangovan et al. 2011), indicating the MTA found in the current study have precedents.

In an association study of wheat protein composition, Plessis et al. (2013) detected markers associated with grain protein loci on chromosomes 2DS and 4AL.

Grain protein deviation is a deviation from regression between grain protein and grain yield. Gpd was highly positively correlated with Gpc (Table 30 and 31). Gpd is also an index for selecting both Gpc and/or Gy in breeding programs (Monaghan et al. 2001; Guttieri et al. 2015; Mosleth et al. 2015; Bogard et al. 2008). There were three MTA detected for Gpd on chromosomes 2B, 4A, and 7A in the current study. A SNP on chromosome 2B, *BS00064369_51*, was associated with Gpd and Gpc in LIRF Dry is an indicator of the close association between Gpc and Gpd, making it a promising SNP for grain protein selection. Monaghan et al. (2001) suggested that Gpd is a criterion for selection of both high Gpc and Gy in breeding programs. Gpd is an interesting potential target in breeding programs as it appears to be relatively robust across different environments (Bogard et al. 2010). However, MTA for Gpd was not robust in our study.

In the current study, there were 10 SNP associated with Gac detected on chromosomes 2B, 4B, and 6A. These MTA were not repeated across environments. El-Feki et al. (2013) did not find any QTL associated with Gac on these chromosomes in the CO940610/Platte DH population. The previous study in the RIL population (Chapter 2 of this dissertation) did not detect any markers associated with Gac on chromosomes 1B, 6B, and 7B. However, Zhang et al. (2008) reported some QTL for Gac on chromosomes 1B and 6A in a population of 93 RILs of durum wheat, where Kofa contributed the allele for high Gac for both QTL. Gac is an important quality parameter for the durum milling industry, and must not exceed 0.9% for first grade commercial semolina (Troccoli et al. 2000).

Test weight was the trait having the highest number of MTA among the five traits investigated. There were a total of 17 strong MTA in the four environments. Of these 17 MTA, two SNP,

BS00047114_51 and *BS00065934_51*, on chromosome 3B were repeated across three environments, ARDEC Dry, ARDEC Wet, and LIRF Dry. These 17 MTA were detected on 1B, 2B, 3A, 3B, 4B, and 7B; but most of them were on chromosome 3B (8) and 4B (3). The previous study in the RIL population (Chapter 2 of this dissertation) detected markers on chromosomes 1B, 6B and 7B having association with Tw, and showing repeatability across two environments, rainfed and fully irrigated. El-Feki et al. (2013) also found QTL for Tw on chromosomes 1B, 6B, 7A, and 7D, but none of them was repeated across environments. In an association study in a spring wheat AM panel grown at LIRF and in Ethiopia, Edae et al. (2014) obtained the second highest number of MTA for Tw on 2DL, 3BS, 4A, 4BL and 7BL, which comprised the location of MTA detected in three out of the total four environments. (Elangovan et al. 2011) identified 26 QTL for Tw on 16 chromosomes, the largest number of QTL detected for any of their evaluated traits. These results suggest that these MTA for Tw are repeatable, making them promising targets for marker-based selection in breeding programs.

Association mapping has been used successfully to detect QTL in wheat for end-use quality traits such as grain protein (Plessis et al. 2013; Zheng et al. 2009), kernel size and milling quality (Breseghello & Sorrells 2006); and for grain yield and yield components (Edae et al. 2014; Maccaferri et al. 2011; Neumann et al. 2011; Dodig et al. 2012; Sukumaran et al. 2015). However, in the current study, allelic effect estimates of each SNP were very small for five investigated traits. The allelic effect variation among SNP was also very small. This suggests that the effect of each SNP was not outstanding and relatively similar, thus, presenting difficulties in identifying the true MTA for the traits and finding candidate genes. With the same experiments (germplasm, locations, years, and treatments), but different traits investigated, Awad (2015) and Grogan (2015) encountered similar problems of few MTA detected and small size of allelic

effects. Some possible reasons can help explain this result. The investigated traits are complex traits, thus largely affected by environmental factors. The genetic diversity for the investigated traits may be low, especially in related elite germplasm, leading to more difficulties in finding an outstanding candidate genes or MTA. The experimental errors in the studies may be large. These may include low homogeneity of field trial locations (e.g., for soil moisture or soil chemical and physical properties), low precision of trait measurement, too few replicates, and a less than optimum experimental design. Statistical analysis can likely be improved as methods for GWAS continue to evolve. There were no permutations used for MTA in the current study. The low marker density across the genome could also contribute to the problem.

Conclusions

- Using different statistical criteria resulted in enormously different numbers of significant MTA.
- No MTA (using Gao's approach) were detected for grain quality traits on chromosomes 1B, 6B and 7B, except for one MTA for Tw on 7B.
- Effective size of MTA detected was not large.
- Grain protein deviation can be used to select favorable alleles for both grain protein concentration and yield.
- The same two SNP were significant for Tw over three of four environments. These should be further investigated.

CHAPTER 5: GENERAL CONCLUSIONS

A previous study found many QTL for grain quality traits in a population of 185 CO940610/Platte DH lines (El-Feki 2010). Three chromosome regions on 1B.1, 6B.1, and 7B were of interest because clusters of QTL for grain quality traits co-located on these regions. El-Feki reported on chromosome region 1B.1, QTL for Gpc in one environment, QTL for Kw in two environments, and QTL for Kd in one environment. On chromosome 6B.1, he found QTL for Gpc in two environments. On chromosome 7B, he found QTL for Gpc in three environments and QTL for Gac in three environments. In particular, at locus *Glu-B1* on 1B and *Xwmc182a* on 6B, CO940610 contributed the higher value Gpc allele, while at locus *Xwmc182b* on 7B, Platte contributed the higher value Gpc allele. Therefore, the overall goal of this dissertation research was to validate the selected QTL detected in the CO940610/Platte DH population for grain quality traits using three mapping populations, which included 186 CO940610/Platte RIL (RIL population), 35 CO940610/Platte BC_3F_2 lines (BC population), and 299 hard winter wheat association mapping panel (AM population), under contrasting soil moisture conditions.

MTA for Gpc, Kw, and Kd on chromosome 1B.1

One MTA for Gpc was detected in the BC population in the LIRF Dry environment, and Platte contributed the higher value Gpc allele. In contrast, El-Feki found that CO940610 contributed the higher value Gpc allele at this locus. However, this study used ETS marker (*Bx7-MAR*) representative of this locus, while El-Feki used a protein band.

MTA for Kw were found in the RIL population (two environments) and in the BC population (two environments), but was not tested in the AM population. This MTA was consistent with El-Feki's findings, and was robust across environments and populations.

MTA for Kd were presented in the RIL population (one environment) and in the BC population (one environment), but was not studied in the AM population. This MTA agreed with El-Feki's findings. It was robust across populations, but not across environments.

MTA for Gpc on chromosome 6B.1

An MTA for Gpc (*Xwmc182a*-Gpc association) was detected in the BC population (two environments), but not in the RIL and AM populations. This MTA conformed to El-Feki's finding, although it was not robust across populations.

MTA for Gpc and Gac on chromosome 7B

One MTA for Gpc (*Xwmc182b*-Gpc association) was detected in the BC in the ARDEC Wet environment, but not in the RIL and AM populations. This MTA conformed to El-Feki (2010), however, it was not robust across environments and populations.

An MTA for Gac was found in the RIL population and robust across two environments, and in the BC population in the ARDEC Wet environment, but not in the AM population. These results conformed to El-Feki (2010). The MTA for Gac was likely robust across populations and environments.

Linkage map construction on 6B.1 and 7B

Linkage maps constructed in the CO940610/Platte RIL and DH populations were mostly consistent.

Four SSR markers (*Xbarc136, Xwmc397, Xwmc182a,* and *Xbarc198)* were used to constructed linkage map 6B.1. Span of the linkage map was 12.2 cM in El-Feki and 9.5 cM in this study. However, marker orders are slightly different.

Eleven SSR markers were mapped to linkage group on 7B. Order of two markers, *Xgwm569* and

Xgwm606, was reversed in comparison with El-Feki (2010). Three markers, *Xwmc426, Xwmc17a,* and *Xwmc364* were added in comparison with the previous study. The other six markers (*Xgwm76, Xwmc182b, Xgwm573, Xgwm46, Xgwm333,* and *Barc278*) had exactly the same order as in El-Feki, although span of the linkages was much different. The linkage span in the current study was 41.6 cM, while it was 19.5 cM in El-Feki (2010).

Two linkage groups around *Xwmc182a* and *Xwmc182b*

The linkage group around *Xwmc182a* on 6B.1 contained three markers (*Xwmc182a, Barc198,* and *Xwmc397*), and spanned 2.3 cM. These markers were associated with Tw, Kw, Kd, and Kh in the RIL population. Marker *Xwmc182a* was associated with Gpc in the DH and BC populations.

The linkage group around *Xwmc182b* on 7B contained four markers (*Xwmc182b, Xwmc426, Xgwm573,* and *Xwmc17a*), and spanned 1.9 cM. Three of these markers (*Xwmc426, Xgwm573,* and *Xwmc17a*) were associated with Gac, which was robust across two environments. Grain ash concentration was stably and positively correlated with Gpc. Marker *Xwmc182b* was associated with Gpc in the DH and BC population.

Therefore, these tightly linked markers located around *Xwmc182a* on 6B.1 and *Xwmc182b* on 7B may be useful for breeding programs.

Marker-assisted selection

Three markers *Bx7-MAR, Xwmc182a,* and *Xmwc182b* were used to develop marker-assisted backcross (MABC) lines. Through the time-consuming process of crossing and selection, 35 BC_3F_2 were obtained, then tested in three environments. The 35 BC_3F_2 lines were classified in eight allelic combinations, and it was hypothesized that the combinations PL-PL-CO and CO-CO-

PL were the lowest and highest Gpc, respectively. The results from testing in the field trials showed Gpc of the CO-CO-PL was higher than PL-PL-CO in the ARDEC Wet environment. This indicates that the finding conformed to (El-Feki 2010) and that CO-CO-PL lines may be useful.

Genome-wide association study

Genome-wide association study (GWAS) is a powerful tool to detect genomic regions controlling complex traits, but this study found relatively few MTA based on *P*-value $<1.67x10^{-4}$ as suggested by Gao et al. (2008). Using different statistical criteria resulted in enormously different numbers of significant MTA; this is a contentious issue and a best method has not been established.

No MTA (using Gao's approach) were detected for Gpc, Gac, Kw, and Kd on chromosomes 1B.1, 6B.1 and 7B, which were detected in the DH population (El-Feki 2010). However, the same two SNP (BS00105846_51 and BS00047114_51 on chromosome 3B) were associated with Tw, and robust over three of four environments.

GWAS could be improved by using a larger and more diverse population, having denser markers, improving the genetic map, reducing experimental errors, and improving statistical analysis.

Future works

- Two linkage groups around *Xwmc182a* and *Xwmc182b* may be valuable for grain quality traits and should be validated before their usein breeding programs.
- The MABC lines with desirable allelic combinations need further investigation.
- Two SNP (BS00105846_51 and BS00047114_51 on chromosome 3B) associated with Tw should be further investigated to detect candidate genes underlying the markers.

REFERENCES

Abdurakhmonov, I.Y. & Abdukarimov, A., 2008. Application of association mapping to understanding the genetic diversity of plant germplasm resources. *International Journal of Plant Genomics*, p.DOI: 10.115/2008/574927.

Abiola, O. et al., 2003. The nature and identification of quantitative trait loci: a community's view. *Nature Reviews Genetics*, 4(11), pp.911–916.

Adhikari, T.B. et al., 2011. Association Mapping of Quantitative Resistance to Phaeosphaeria nodorum in Spring Wheat Landraces from the USDA National Small Grains Collection. *Phytopathology*, 101(11), pp.1301–1310.

AgriPro Biosciences Inc., 1998. Registeration of Longhorn.

Ahmadi, A. & Baker, D.A., 2001. The effect of water stress on the activities of key regulatory enzymes of the sucrose to starch pathway in wheat. *Plant Growth Regulation*, 35, pp.81–91.

Ahmed, M. & Fayyaz-ul-Hassan, 2015. Response of Spring Wheat (Triticum aestivum L.) Quality Traits and Yield to Sowing Date. *PloS ONE*, 10(4), p.e0126097.

Ahuja, I. et al., 2010. Plant molecular stress responses face climate change. *Trends in Plant Science*, 15(12), pp.664–674.

Alam, M.A. et al., 2011. Powdery Mildew Resistance Genes in Wheat: Identification and Genetic Analysis. *Journal of Molecular Biology Research*, 1(1), pp.20–39.

Alexander, W.L., Smith, E.L. & Dhanasobhan, C., 1984. A comparison of yield and yield component selection in winter wheat. *Euphytica*, 33, pp.953–961.

Al-Saleh, A. & Brennan, C.S., 2012. Bread Wheat Quality: Some Physical, Chemical and Rheological Characteristics of Syrian and English Bread Wheat Samples. *Foods*, 1, pp.3–17.

Amiri, R., Bahraminejad, S. & Jalali-honarmand, S., 2013. Effect of terminal drought stress on grain yield and some morphological traits in 80 bread wheat genotypes. *International Journal of Agriculture and Crop Sciences*, 5(10), pp.1145–1153.

Ammiraju, J.S.S. et al., 2001. Identification of inter simple sequence repeat (ISSR) markers associated with seed size in wheat. *Theoretical and Applied Genetics*, 102, pp.726–732.

Anderson, J.A., Sorrells, M.E. & Tanksley, S.D., 1993. RFLP Analysis of Genomic Regions Associated with Resistance to Preharvest Sprouting in Wheat. *Crop Science*, 33, pp.453–459.

Ansorge, W.J., 2009. Next-generation DNA sequencing techniques. *New Biotechnology*, 25(4), pp.195–203.

Appelbee, M.-J. et al., 2009. Novel allelic variants encoded at the Glu-D3 locus in bread wheat. *Journal of Cereal Science*, 49(2), pp.254–261.

Aprile, A. et al., 2013. Different stress responsive strategies to drought and heat in two durum wheat cultivars with contrasting water use efficiency. *BMC genomics*, 14(821).

Araki, E., Miura, H. & Sawada, S., 1999. Identification of genetic loci affecting amylose content and agronomic traits on chromosome 4A of wheat. *Theoretical and applied genetics*, 98, pp.977–984.

Asins, M.J., 2002. Present and future of quantitative trait locus analysis in plant breeding. *Plant Breeding*, 121, pp.281–291.

Awad, W., 2015. Genome-Wide Association Study and Drought Tolerance Evaluation of a Winter Wheat Association Mapping Panel. *Doctoral Dissertation, Colorado State University, Colorado, U.S.A.*

Aykut Tonk, F., Ilker, E. & Tosun, M., 2010. A study to incorporate high protein content from tetraploid wheat (T. turgidum dicoccoides) to hexaploid wheat (T. aestivum vulgare). *Turkish Journal of Field Crops*, 15(1), pp.69–72.

Baenziger, P.S. et al., 2012. Registration of "NI04421" Hard Red Winter Wheat. *Journal of Plant Registrations*, 6(1), pp.54–59.

Balla, K. et al., 2011. Quality of winter wheat in relation to heat and drought shock after anthesis. *Czech Journal of Food Sciences*, 29(2), pp.117–128.

Barakat, M.N. et al., 2011. Identification of new microsatellite marker linked to the grain filling rate as indicator for heat tolerance genes in F2 wheat population. *Australian Journal of Crop Science*, 5(2), pp.104–110.

Barnard, A., Labuschagne, M. & Niekerk, H. Van, 2002. Heritability estimates of bread wheat quality traits in the Western Cape province of South Africa. *Euphytica*, 127, pp.115–122.

Basten, C.J., Weir, B.S. & Zeng, Z.-B., 2005. QTL Cartographer: Version 1 . 17. *Manual*.

Benjamini, Y. & Hochberg, Y., 1995. Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. *Journal of Royal Statistical Society*, 57(1), pp.289–300.

Bennett, D. et al., 2012. Detection of two major grain yield QTL in bread wheat (Triticum aestivum L.) under heat, drought and high yield potential environments. *Theoretical and Applied Genetics*, pp.doi: 10.1007/s00122–012–1927–2.

Berman, M. et al., 1996. Image Analysis of Whole Grains to Screen for Flour-Milling Yield in Wheat Breeding. *Cereal Chemistry*, 73(3), pp.323–327.

Bettge, A.D., Morris, C.F. & Greenblatt, G.A., 1995. Assessing genotypic softness in single wheat kernels using starch granule-associated friabilin as a biochemical marker. *Euphytica*, 86, pp.65–72.

Bilgin, O. et al., 2010. Variation and Heritability for Some Semolina Characteristics and Grain Yield and Their Relations in Durum Wheat (Triticum durum Desf.). *World Journal of Agricultural Sciences*, 6(3), pp.301–308.

Blanco, A. et al., 2002. Detection of grain protein content QTLs across environments in tetraploid wheats. *Plant Molecular Biology*, 48, pp.615–623.

Blanco, A., Simeone, R. & Gadaleta, A., 2006. Detection of QTLs for grain protein content in durum wheat. *Theoretical and Applied Genetics*, 112, pp.1195–1204.

Blum, A., 2011. Drought resistance - is it really a complex trait? *Functional Plant Biology*, 38, pp.753–757.

Blum, A., 1988. Plant breeding for stress environments. *CRC Press*.

Blum, A. et al., 1989. Yield stability and canopy temperature of wheat genotypes under drought-stress. *Field Crops Research*, 22, pp.289–296.

Bogale, A. & Tesfaye, K., 2011. Relationship Between Kernel Ash content, Water Use Efficiency and Yield in Durum Wheat under Water Deficit Induced at Different Growth Stages. *African Journal of Basic & Applied Sciences*, 3(3), pp.80–86.

Bogard, M. et al., 2010. Deviation from the grain protein concentration-grain yield negative relationship is highly correlated to post-anthesis N uptake in winter wheat. *Journal of Experimental Botany*, 61(15), pp.4303–4312.

Bogard, M. et al., 2008. Genetic analysis of grain protein deviation in wheat. , pp.1–3.

Bohn, M. et al., 2001. Re-evaluation of the prospects of marker-assisted selection for improving insect resistance against Diatraea spp. in tropical maize by cross validation and independent validation. *Theoretical and Applied Genetics*, 103, pp.1059–1067.

Bonferroni, C.E., 1935. Il calcolo delle assicurazion su gruppi di teste, chapter "Studi in Onore del Professore Salvatore ortu Carboni." *Rome*, pp.13–60.

Bonferroni, C.E., 1936. Teoria statistica delle classi e calcolo delle probabilitá. *Pubblicazioni del Istituto Superiore di Scienze Economiche e Commercial di Firenze*, 8, pp.3–62.

Boyer, J.S., 1982. Plant productivity and environment. *Science*, 218(4571), pp.443–448.

Bradbury, P.J. et al., 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics (Oxford, England)*, 23(19), pp.2633–2635.

Branlard, G., Dardevet, M. & Saccomano, R., 2001. Genetic diversity of wheat storage proteins and bread wheat quality. *Euphytica*, 119, pp.59–67.

Breseghello, F. et al., 2005. Genetic loci related to kernel quality differences between a soft and a hard wheat cultivar. *Crop Science*, 45, pp.1685–1695.

Breseghello, F. & Sorrells, M.E., 2006. Association mapping of kernel size and milling quality in wheat (Triticum aestivum L.) cultivars. *Genetics*, 172, pp.1165–1177.

Breseghello, F. & Sorrells, M.E., 2007. QTL analysis of kernel size and shape in two hexaploid wheat mapping populations. *Field Crops Research*, 101, pp.172–179.

Broman, K.W. et al., 2003. R/qtl: QTL mapping in experimental crosses. *Bioinformatics*, 19(7), pp.889–890.

Brown, B. et al., 2005. Nitrogen Management for Hard Wheat Protein Enhancement. *Pacific Northwest Extension Publications*, 578, pp.1–14.

Bryan, G.J. et al., 1999. Polymorphic simple sequence repeat markers in chloroplast genomes of Solanaceous plants. *Theoretical and Applied Genetics*, 99, pp.859–867.

Buntjer, J.B., Sørensen, A.P. & Peleman, J.D., 2005. Haplotype diversity: The link between statistical and biological association. *Trends in Plant Science*, 10(10), pp.466–471.

Butow, B.J. et al., 2004. Dissemination of the highly expressed Bx7 glutenin subunit (Glu-B1al allele) in wheat as revealed by novel PCR markers and RP-HPLC. *Theoretical and applied genetics*, 109, pp.1525–1535.

Campbell, B.T. et al., 2003. Identification of QTLs and environmental interactions associated with agronomic traits on chromosome 3A of wheat. *Crop Science*, 43, pp.1493–1505.

Campbell, K.G. et al., 1999. Quantitative trait loci associated with kernel traits in a soft x hard wheat cross. *Crop Science*, 39, pp.1184–1195.

Cane, K., Spackman, M. & Eagles, H.A., 2004. Puroindoline genes and their effects on grain quality traits in southern Australian wheat cultivars. *Australian Journal of Agricultural Research*, 55, pp.89–95.

Canè, M.A. et al., 2014. Association mapping for root architectural traits in durum wheat seedlings as related to agronomic performance. *Molecular Breeding*, 34, pp.1629–1645.

Cavagnaro, P.F. et al., 2010. Genome-wide characterization of simple sequence repeats in cucumber (Cucumis sativus L.). *BMC genomics*, 11(569).

Chantret, N. et al., 2005. Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (Triticum and Aegilops). *The Plant cell*, 17, pp.1033–1045.

Chastain, T.G., 2003. Yield Components and Crop Yield. *Oregon State University*.

Chaves, M.M., Maroco, J.P. & Pereira, J.S., 2003. Understanding plant responses to drought—from genes to the whole plant. *Functional Plant Biology*, 30, pp.239–264.

Chee, P.W. et al., 2001. Evaluation of a High Grain Protein QTL from Triticum turgidum L. var. dicoccoides in an Adapted Durum Wheat Background. *Crop Science*, 41, pp.295–301.

Chen, F. et al., 2007. Prevalence of a novel puroindoline b allele in Yunnan endemic wheats (Triticum aestivum ssp. yunnanense King). *Euphytica*, 156, pp.39–46.

Chu, C.G. et al., 2008. Whole genome mapping in a wheat doubled haploid population using SSRs and TRAPs and the identification of QTL for agronomic traits. *Molecular Breeding*, 22, pp.251–266.

Clarke, F.R. et al., 2009. Protein concentration inheritance and selection in durum wheat. *Canadian Journal of Plant Science*, 89, pp.601–612.

Collard, B.C.Y. et al., 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, 142, pp.169–196.

Cosson, P., Decroocq, V. & Revers, F., 2014. Development and characterization of 96 microsatellite markers suitable for QTL mapping and accession control in an Arabidopsis core collection. *Plant methods*, 10(2).

Crossa, J. et al., 2007. Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. *Genetics*, 177, pp.1889–1913.

Cui, F. et al., 2011. Wheat kernel dimensions: How do they contribute to kernel weight at an individual QTL level? *Journal of Genetics*, 90(3), pp.409–425.

Cuthbert, J.L. et al., 2008. Molecular mapping of quantitative trait loci for yield and yield components in spring wheat (Triticum aestivum L.). *Theoretical and Applied Genetics*, 117, pp.595–608.

Daniel, C. & Triboı, E., 2002. Changes in wheat protein aggregation during grain development : effects of temperatures and water stress. *European Journal of Agronomy*, 16, pp.1–12.

Daud, H.M. & Gustafson, J.P., 1996. Molecular evidence for Triticum speltoides as a B-genome progenitor of wheat (Triticum aestivum). *Genome*, 39, pp.543–548.

Dholakia, B.B. et al., 2003. Molecular marker analysis of kernel size and shape in bread wheat.

Plant Breeding, 122, pp.392–395.

Dholakia, B.B. et al., 2001. Molecular marker analysis of protein content using PCR-based markers in wheat. *Biochemical Genetics*, 39(9/10), pp.325–338.

Dodig, D. et al., 2012. Genetic and association mapping study of wheat agronomic traits under contrasting water regimes. *International Journal of Molecular Sciences*, 13, pp.6167–6188.

Doğan, R., 2009. The correlation and path coefficient analysis for yield and some yield components of durum wheat (Triticum turgidum var. durum L.) in west Anatolia conditions. *Pakistan Journal of Botany*, 41(3), pp.1081–1089.

Dreher, K. et al., 2003. Money matters (I): Costs of field and laboratory procedures associated with conventional and marker-assisted maize breeding at CIMMYT. *Molecular Breeding*, 11, pp.221–234.

Edae, E.A., 2013. Association Mapping for Yield, Yield Components and Drought Tolerance-Related Traits in Spring Wheat Grown under Rainfend and Irrigated Coditions. *Doctoral Dissertation, Colorado State University, Colorado, U.S.A*.

Edae, E.A. et al., 2014. Genome-wide association mapping of yield and yield components of spring wheat under contrasting moisture regimes. *Theoretical and applied genetics*, 127, pp.791–807.

Edwards, A. et al., 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American journal of human genetics*, 49, pp.746–756.

Elangovan, M. et al., 2011. Mapping QTL associated with agronomic traits in bread wheat (Triticum aestivum L .). *Journal of Wheat Research*, 3(1), pp.14–23.

El-Feki, W.M., 2010. Mapping Quantitative Trati Loci for Bread Making Quality and Agronomic Traits in Winter Wheat under Different Soil Moisture Levels. *Doctoral Dissertation, Colorado State University, Colorado, U.S.A*.

El-Feki, W.M. et al., 2013. Quantitative trait locus mapping for end-use quality traits in hard winter wheat under contrasting soil moisture levels. *Crop Science*, 53, pp.1953–1967.

El-Feki, W.M. et al., 2015. Registration of CO940610/"Platte" Wheat Doubled Haploid Mapping Population. *Journal of Plant Registrations*, 9, pp.419–423.

Endelman, J.B., 2011. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *The Plant Genome*, 4, pp.250–255.

Falconer, S.D. & Mackay, F.C.T., 1996. Introduction to Quantitative Genetics, 4th edition. *TIG*, 12(7).

FAO, 2013. http://faostat3.fao.org/browse/Q/*/E.

Fjell, D.L., Paulsen, G.M. & Walter, T.L., 1985. Relationship among planted and harvested kernel weights and grain yield and protein percentage in winter wheat. *Euphytica*, 34, pp.751–757.

Fleury, D. et al., 2010. Genetic and genomic tools to improve drought tolerance in wheat. *Journal of Experimental Botany*, 61(12), pp.3211–3222.

Foulkes, M.J., Scott, R.K. & Sylvester-Bradley, R., 2002. The ability of wheat cultivars to withstand drought in UK conditions: formation of grain yield. *Journal of Agricultural Science*, 138, pp.153–169.

Foulkes, M.J., Scott, R.K. & Sylvester-Bradley, R., 2001. The ability of wheat cultivars to withstand drought in UK conditions: resource capture. *Journal of Agricultural Science*, 138, pp.1–16.

Gale, M.D. & Devos, K.M., 1997. Comparative genetics in the grasses. *Plant Molecular Biology*, 35, pp.3–15.

Gao, X., Starmer, J. & Martin, E.R., 2008. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genetic Epidemiology*, 32, pp.361–369.

Ghavami, F. et al., 2011. Mixed Model Association Mapping for Fusarium Head Blight Resistance in Tunisian-Derived Durum Wheat Populations. *Genes|Genomes|Genetics*, 1, pp.209–218.

Gibson, G. & Muse, S. V, 2009. A Primer of Genome Science. *North Corolina State University, Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts, 01375*.

Giura, A. & Saulescu, N.N., 1996. Chromosomal location of genes controlling grain size in a large grained selection of wheat (Triticum aestivum L.). *Euphytica*, 89, pp.77–80.

Graybosch, R.A. & Peterson, C.J., 2010. Genetic improvement in winter wheat yields in the Great Plains of North America, 1959-2008. *Crop Science*, 50, pp.1882–1890.

Griffiths, A., Miller, J. & Suzuki, D., 2000. An Introduction Genetic Analysis.

Grogan, S., 2015. Dissecting drought tolerance in winter wheat using phenotypic and genetic analyses of agronomic and spectral traits. *Doctoral Dissertation, Colorado State University, Colorado, U.S.A.*

Groos, C. et al., 2003. Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. *Theoretical and applied genetics*, 106, pp.1032–1040.

Groos, C., Bervas, E. & Charmet, G., 2004. Genetic analysis of grain protein content, grain hardness and dough rheology in a hard×hard bread wheat progeny. *Journal of Cereal*

Science, 40, pp.93–100.

Gupta, P.K. et al., 1999. Molecular markers and their applications in wheat breeding. *Plant Breeding*, 118, pp.369–390.

Gupta, P.K. et al., 2011. QTL analysis, association mapping and marker-assisted selection for some quality traits in bread wheat - An overview of the work done at CCS University, Meerut. *Journal of Wheat Research*, 3(2), pp.1–11.

Guttieri, M.J. et al., 2000. End-Use Quality of Six Hard Red Spring Wheat Cultivars at Different Irrigation Levels. *Crop Science*, 40, pp.631–635.

Guttieri, M.J. et al., 2015. Variation for Grain Mineral Concentration in a Diversity Panel of Current and Historical Great Plains Hard Winter Wheat Germplasm. *Crop Science*, 55, pp.1035–1052.

Haider, N., 2013. The origin of the B-genome of bread wheat (Triticum aestivum L.). *Russian Journal of Genetics*, 49(3), pp.263–274.

Haley, S.D. et al., 2005. Registration of "Hatcher" Wheat. *Crop Science*, 45, pp.2654–2656.

Haley, S.D. et al., 2007. Registration of "Ripper" Wheat. *Journal of Plant Registrations*, 1, pp.1–6.

Haley, S.D. et al., 2011. Registration of "Snowmass" Wheat. *Journal of Plant Registrations*, 5(1), pp.87–90.

Hartl, D. & Jones, E., 2001. Genetics: Analysis of genes and genomes. *8th Edition*.

Hayden, M.J. et al., 2008. Multiplex-ready PCR: A new method for multiplexed SSR and SNP genotyping. *BMC genomics*, 9(80).

Heidari, B. et al., 2011. Mapping QTL for grain yield, yield components, and spike features in a doubled haploid population of bread wheat. *Genome*, 54, pp.517–527.

Hill, J., Becker, H.C. & Tigerstedt, P.M.A., 2012. Quantitavie and Ecological Aspects of Plant Breeding.

Holland, J.B., 2007. Genetic architecture of complex traits in plants. *Current Opinion in Plant Biology*, 10, pp.156–161.

Hrušková, M. & Švec, I., 2009. Wheat hardness in relation to other quality factors. *Czech Journal of Food Sciences*, 27(4), pp.240–248.

Huang, X.Q. et al., 2003. Advanced backcross QTL analysis for the identification of quantitative trait loci alleles from wild relatives of wheat (Triticum aestivum L.). *Theoretical and applied genetics*, 106, pp.1379–1389.

Huang, X.Q. et al., 2006. Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (Triticum aestivum L.). *Theoretical and Applied Genetics*, 113, pp.753–766.

Ibrahim, S.E. et al., 2012. Comparison of QTLs for drought tolerance traits between two advanced backcross populations of spring wheat. *International Journal of AgriScience*, 2(3), pp.216–227.

Igrejas, G. et al., 2002. Mapping QTLs for grain hardness and puroindoline content in wheat (Triticum aestivum L.). *Theoretical and Applied Genetics*, 106, pp.19–27.

Jackson, E.A., Holt, L.M. & Payne, P.I., 1983. Characterisation of high molecular weight gliadin and low-molecular-weight glutenin subunits of wheat endosperm by two-dimensional electrophoresis and the chromosomal localisation of their controlling genes. *Theoretical and applied genetics*, 66(1), pp.29–37.

Jacob, H.J. et al., 1991. Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat. *Cell*, 67, pp.213–224.

Jahufer, M.Z.Z. et al., 2003. DNA fingerprinting and genetic relationships among white clover cultivars. *Proceedings of the New Zealand Grassland Association*, 65, pp.163–169.

Jansen, R.C., 1993. Interval mapping of multiple quantitative trait loci. *Genetics*, 135, pp.205–211.

Jansen, R.C. & Stam, P., 1994. High resolution of quantitative traits into multiple loci via interval mapping. *Genetics*, 136, pp.1447–1455.

Jenner, C. & Rathjen, A., 1975. Factors Regulating the Accumulation of Starch in Ripening Wheat Grain. *Australian Journal of Plant Physiology*, 2(3), pp.311–322.

Jiaqin, S. et al., 2009. Unraveling the complex trait of crop yield with quantitative trait loci mapping in Brassica napus. *Genetics*, 182, pp.851–861.

Jolly, C.J. et al., 1993. Characterisation of the wheat Mr 15000 "grain-softness protein" and analysis of the relationship between its accumulation in the whole seed and grain softness. *Theoretical and Applied Genetics*, 86, pp.589–597.

Jones, C. & Olson-Rutz, K., 2012. Practices To Increase Wheat Grain Protein. *Montana State University Extension, EBO206*, pp.1–12.

Jones, N., Ougham, H. & Thomas, H., 1997. Markers and mapping: We are all geneticists now. *New Phytologist*, 137, pp.165–177.

Kadam, S. et al., 2012. Genomic associations for drought tolerance on the short arm of wheat chromosome 4B. *Functional and Integrative Genomics*, 12, pp.447–464.

Katti, M. V, Ranjekar, P.K. & Gupta, V.S., 2001. Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Molecular biology and evolution*, 18(7), pp.1161–1167.

Kearsey, M., 1998. The principles of QTL analysis (a minimal mathematics approach). *Journal of Experimental Botany*, 49(327), pp.1619–1623.

Khan, I.A. et al., 2000. Development of PCR-Based markers for a high Grain Protein Content gene from Triticum turgidum ssp. dicoccoides transferred to bread wheat. *Crop Science*, 40, pp.518–524.

Kibite, S. & Evans, L.E., 1984. Causes of negative correlations between grain yield and grain protein concentration in common wheat. *Euphytica*, 33, pp.801–810.

Kilic, H. & Tacettin, Y., 2010. The Effect of Drought Stress on Grain Yield, Yield Components and some Quality Traits of Durum Wheat (Triticum turgidum ssp. durum). *Notulae Botanicae Horti Agrobotanici ClujNapoca*, 38(1), pp.164–170.

Kirigwi, F.M. et al., 2007. Markers associated with a QTL for grain yield in wheat under drought. *Molecular Breeding*, 20, pp.401–413.

Knapp, S.J., Stroup, W.W. & Ross, W.M., 1985. Exact confidence intervals for heritability on a progeny mean basis. *Crop Science*, 25, pp.192–194.

Kosmolak, F.G., 1978. Grinding Time—a Screening Test for Kernel Hardness in Wheat. *Canadian Journal of Plant Science*, 58, pp.415–420.

Kuchel, H. et al., 2007. Genetic dissection of grain yield in bread wheat. I. QTL analysis. *Theoretical and Applied Genetics*, 115, pp.1029–1041.

Kuchel, H. et al., 2006. The genetic control of milling yield, dough rheology and baking quality of wheat. *Theoretical and Applied Genetics*, 112, pp.1487–1495.

Kumar, J. et al., 2011. Introgression of a major gene for high grain protein content in some Indian bread wheat cultivars. *Field Crops Research*, 123, pp.226–233.

Kumar, N. et al., 2007. QTL mapping for yield and yield contributing traits in two mapping populations of bread wheat. *Molecular Breeding*, 19, pp.163–177.

Lagercrantz, U., Ellegren, H. & Andersson, L., 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic acids research*, 21(5), pp.1111–1115.

Lander, E. & Kruglyak, L., 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature genetics*, 11, pp.241–247.

Lander, E.S. et al., 1987. MAPMAKER: An interactive computer package for constructing

primary genetic linkage maps of experimental and natural populations. *Genomics*, 1, pp.174–181.

Lander, E.S. & Botstein, S., 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121, pp.185–199.

Langridge, a et al., 2001. Trends in genetic and genome analyses in wheat: a review. *Australian Journal of Agricultural Research*, 52, pp.1043–1077.

Li, J. & Ji, L., 2005. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity*, 95, pp.221–227.

Li, S. et al., 2007. A intervarietal genetic map and QTL analysis for yield traits in wheat. *Molecular Breeding*, 20, pp.167–178.

Li, Y. et al., 2009. Detection of QTLs for bread-making quality in wheat using a recombinant inbred line population. *Plant Breeding*, 128, pp.235–243.

Lincoln, S.E., Daly, M.J. & Lander, E.S., 1993. Constructing Genetic Linkage Maps with MAPMAKER / EXP Version A Tutorial and Reference Manual. *A Whitehead Institute for Biomedical Research Technical Report, Third Edition*.

Lipka, A.E. et al., 2012. GAPIT: Genome association and prediction integrated tool. *Bioinformatics*, 28(18), pp.2397–2399.

Litt, M. & Luty, J. a, 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American journal of human genetics*, 44, pp.397–401.

Liu, H. Ben, 1997. Statistical Genomics: Linkage, Mapping and QTL Analysis. *CRC Press*.

Maccaferri, M. et al., 2015. A Genome-Wide Association Study of Resistance to Stripe Rust (Puccinia striiformis f. sp. tritici) in a Worldwide Collection of Hexaploid Spring Wheat (Triticum aestivum L.). *Genes|Genomes|Genetics*, 5, pp.449–465.

Maccaferri, M. et al., 2011. Association mapping in durum wheat grown across a broad range of water regimes. *Journal of experimental botany*, 62(2), pp.409–438.

Maccaferri, M. et al., 2010. Association mapping of leaf rust response in durum wheat. *Molecular Breeding*, 26, pp.189–228.

Maccaferri, M. et al., 2008. Quantitative trait loci for grain yield and adaptation of durum wheat (Triticum durum Desf.) across a wide range of water availability. *Genetics*, 178, pp.489–511.

Mammadov, J. et al., 2012. SNP markers and their impact on plant breeding. *International Journal of Plant Genomics*, pp.1–9.

Manley, M., 1995. Wheat hardness by near infrared (NIR) spectroscopy : New insights. *Doctoral Thesis, University of Plymouth*.

Manly, K.F., Cudmore, R.H. & Meer, J.M., 2001. Map Manager QTX, cross-platform software for genetic mapping. *Mammalian Genome*, 12, pp.930–932.

Mann, G. et al., 2009. Genetic control of wheat quality: interactions between chromosomal regions determining protein content and composition, dough rheology, and sponge and dough baking properties. *Theoretical and applied genetics*, 118, pp.1519–1537.

Mares, D.J.. & Campbell, A.W., 2001. Mapping components of flour and noodle colour in Australian wheat. *Australian Journal of Agricultural Research*, 52, pp.1297–1309.

Marshall, D., Ellison, F. & Mares, D., 1984. Effects of grain shape and size on milling yields in wheat. I. Theoretical analysis based on simple geometric models. *Australian Journal of Agricultural Research*, 35, pp.619–630.

Martin, J.M. et al., 2001. Milling and bread-baking traits associated with puroindoline sequence type in hard red spring wheat. *Crop Science*, 41, pp.228–234.

Martin, J.M. et al., 2007. Pilot scale milling characteristics of transgenic isolines of a hard wheat over-expressing puroindolines. *Crop Science*, 47, pp.497–506.

Martin, J.T. et al., 2007. RonL Hard White Wheat. *Agricultural Research Center, Kansas State University*.

Marza, F. et al., 2006. Quantitative trait loci for yield and related traits in the wheat population Ning7840 x Clark. *Theoretical and Applied Genetics*, 112, pp.688–698.

Matsuoka, Y., 2011. Evolution of polyploid triticum wheats under cultivation: The role of domestication, natural hybridization and allopolyploid speciation in their diversification. *Plant and Cell Physiology*, 52(5), pp.750–764.

McCartney, C.A. et al., 2005. Mapping quantitative trait loci controlling agronomic traits in the spring wheat cross RL4452 x "AC Domain." *Genome*, 48, pp.870–883.

McIntyre, C.L. et al., 2010. Molecular detection of genomic regions associated with grain yield and yield-related components in an elite bread wheat cross evaluated under irrigated and rainfed conditions. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, 120, pp.527–41.

Mirbahar, A.A. et al., 2009. Effect of Water Stress on Yield and Yield Components of Wheat (Triticum Aestivum L .) Varieties. *Pakistan Journal of Botany*, 41(3), pp.1303–1310.

Mohammadi, M. et al., 2012. Relationships between grain yield and yield components in bread wheat under different water availability (dryland and supplemental irrigation conditions).

Notulae Botanicae Horti Agrobotanici Cluj-Napoca, 40(1), pp.195–200.

Mohammadi, R. et al., 2015. Field evaluation of durum wheat landraces for prevailing abiotic and biotic stresses in highland rainfed regions of Iran. *The Crop Journal*, pp.1–11.

Mohan, M. et al., 1997. Genome mapping, molecular markers and marker-assisted selection in crop plants. *Molecular Breeding*, 3, pp.87–103.

Monaghan, J.M. et al., 2001. The use of grain protein deviation for identifying wheat cultivars with high grain protein concentration and yield. *Euphytica*, 122, pp.309–317.

Morris, C.F., 2002. Puroindolines: The molecular genetic basis of wheat grain hardness. *Plant Molecular Biology*, 48, pp.633–647.

Mosleth, E.F. et al., 2015. A novel approach to identify genes that determine grain protein deviation in cereals. *Plant Biotechnology Journal*, 13, pp.625–635.

Narasimhamoorthy, B. et al., 2006. Advanced backcross QTL analysis of a hard winter wheat x synthetic wheat population. *Theoretical and Applied Genetics*, 112, pp.787–796.

Nelson, J.C., 1997. QGENE: software for maker-based genomics analysis and breeding. *Molecular Breeding*, 3, pp.239–245.

Nelson, J.C. et al., 2006. Quantitative trait locus analysis of wheat quality traits. *Euphytica*, 149, pp.145–159.

Neumann, K. et al., 2011. Genome-wide association mapping: A case study in bread wheat (Triticum aestivum L.). *Molecular Breeding*, 27, pp.37–58.

Nezhad, Z.K. et al., 2012. QTL analysis for thousand-grain weight under terminal drought stress in bread wheat (Triticum aestivum L.). *Euphytica*, 186, pp.127–138.

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

Osborne, B.G. et al., 2001. The hardness locus in Australian wheat lines. *Australian Journal of Agricultural Research*, 52, pp.1275–1286.

Paillard, S. et al., 2003. An integrative genetic linkage map of winter wheat (Triticum aestivum L.). *Theoretical and applied genetics*, 107, pp.1235–1242.

Parisod, C. et al., 2013. Genome-specific introgression between wheat and its wild relative Aegilops triuncialis. *Journal of Evolutionary Biology*, 26, pp.223–228.

Paterson, A.H., 1996. Plant genome mapping: strategies and applications. *Biotechnology*, VII.

Patil, R.M. et al., 2009. Molecular mapping of QTLs for gluten strength as measured by sedimentation volume and mixograph in durum wheat (Triticum turgidum L. ssp durum).

Journal of Cereal Science, 49, pp.378–386.

Paux, E. et al., 2008. A physical map of the 1-gigabase bread wheat chromosome 3B. *Science*, 322, pp.101–104.

Paux, E. et al., 2006. Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. *Plant Journal*, 48, pp.463–474.

Payne, P.I., 1987. Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. *Annual Review of Plant Physiology*, 38, pp.141–153.

Payne, P.I., Holt, L.M. & Law, C.N., 1982. Structural and genetical studies on the high-molecular-weight subunits of wheat glutenin - Part 1: Allelic variation in subunits amongst varieties of wheat (Triticum aestivum). *Theoretical and Applied Genetics*, 63, pp.229–236.

Peng, J.H. et al., 2009. Microsatellite-based molecular diversity of bread wheat germplasm and association mapping of wheat resistance to the Russian wheat aphid. *Genetica*, 135, pp.95–122.

Perretant, M.R. et al., 2000. QTL analysis of bread-making quality in wheat using a doubled haploid population. *Theoretical and Applied Genetics*, 100, pp.1167–1175.

Saint Pierre, C. et al., 2008. Winter wheat genotypes under different levels of nitrogen and water stress: Changes in grain protein composition. *Journal of Cereal Science*, 47, pp.407–416.

Pleijel, H., 1999. Grain protein accumulation in relation to grain yield of spring wheat (Triticum aestivum L.) grown in open-top chambers with different concentrations of ozone, carbon dioxide and water availability. *Agriculture, Ecosystems & Environment*, 72, pp.265–270.

Plessis, A. et al., 2013. Association study of wheat grain protein composition reveals that gliadin and glutenin composition are trans-regulated by different chromosome regions. *Journal of Experimental Botany*, 64(12), pp.3627–3644.

Porter, K.B., Gilmore, E.C. & Tuleen, N.A., 1980. Registration of TAM 105 wheat. *Crop Science*, 20(1), p.114.

Prasad, M. et al., 1999. A microsatellite marker associated with a QTL for grain protein content on chromosome arm 2DL of bread wheat. *Theoretical and Applied Genetics*, 99, pp.341–345.

Prasad, M. et al., 2003. QTL analysis for grain protein content using SSR markers and validation studies using NILs in bread wheat. *Theoretical and applied genetics*, 106, pp.659–667.

Price, A.L. et al., 2006. Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics*, 38(8), pp.904–909.

Pshenichnikova, T. a. et al., 2008. Mapping of the quantitative trait loci (QTL) associated with grain quality characteristics of the bread wheat grown under different environmental conditions. *Russian Journal of Genetics*, 44(1), pp.74–84.

Quarrie, S.A. et al., 2005. A high-density genetic map of hexaploid wheat (Triticum aestivum L.) from the cross Chinese Spring × SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theoretical and Applied Genetics*, 110, pp.865–880.

Rafalski, A., 2002. Applications of single nucleotide polymorphism in crop genetics. *Current opinion in plant biology*, 5, pp.94–100.

Rafalski, J.A., 2010. Association genetics in crop improvement. *Current Opinion in Plant Biology*, 13, pp.174–180.

Raman, R. et al., 2005. Genetic and in silico comparative mapping of the polyphenol oxidase gene in bread wheat ({\em Triticum aestivum} L.). *Functional and integrative genomics*, 5, pp.185–200.

Ramel, C., 1997. Mini- and microsatellites. *Environmental Health Perspectives*, 105, pp.781–789.

Rampino, P. et al., 2006. Drought stress response in wheat: Physiological and molecular analysis of resistant and sensitive genotypes. *Plant, Cell and Environment*, 29, pp.2143–2152.

Rector, B.G. et al., 2013. Use of Wheat SSRs to Assess Genetic Diversity in Medusahead (Taeniatherum caput-medusae). *Invasive Plant Science and Management*, 6, pp.352–361.

Riede, C.R. & Anderson, J.A., 1996. Linkage of RFLP markers to an aluminum tolerance gene in wheat. *Crop Science*, 36, pp.905–909.

Risch, N., 1992. Genetic linkage: interpreting lod scores. *Science*, 255(5046), pp.803–804.

Risch, N. & Merikangas, K., 1996. The future of genetic studies of complex human diseases. *Science*, 273, pp.1516–1517.

Roder, M.S. et al., 1998. A Microsatellite Map of Wheat. *Genetics*, 149, pp.2007–2023.

Salem, K.F.M., El-Zanaty, A.M. & Esmail, R.M., 2008. Assessing Wheat (Triticum aestivum L.) Genetic Diversity Using Morphological Characters and Microsatallite Markers. *World Journal of Agricultural Sciences*, 4(5), pp.538–544.

Sayed, H. et al., 2002. Segregation distortion in doubled haploid lines of barley (Hordeum vulgare) detected by simple sequence repeat (SSR) markers. *Euphytica*, 125, pp.265–272.

Schwarz, G., 1978. Estimating the dimension of a model. *The Annals of Statistics*, 6(2), pp.461–464.

Sears, R., 1954. The aneuploids of common wheat. *University Archives of the University of Missouri-Columbia, Research Bulletin*, 572.

Sehgal, S.K. et al., 2012. Chromosome arm-specific BAC end sequences permit comparative analysis of homoeologous chromosomes and genomes of polyploid wheat. *BMC Plant Biology*, 12(64).

Seleiman, M. et al., 2011. Productivity, Grain and Dough Quality of Bread Wheat Grown with Different Water Regimes. *Journal of AcroCrop Science*, 2(1), pp.11–17.

Shelton, D., Martin, G. & Peterson, B., 2008. Wheat and Flour Testing Methods : A guide to Understanding Wheat and Flour Quality: Version 2.

Shewry, P.R., 2009. Wheat. *Journal of Experimental Botany*, 60(6), pp.1537–1553.

Simmonds, J. et al., 2014. Identification and independent validation of a stable yield and thousand grain weight QTL on chromosome 6A of hexaploid wheat (Triticum aestivum L.). *BMC plant biology*, 14(191).

Sleper, D.A. & Poehlman, J.M., 2006. Breeding Field Crops. *Blackwell Publishing, Fifth Edition*, p.225.

Somers, D.J., Isaac, P. & Edwards, K., 2004. A high-density microsatellite consensus map for bread wheat (Triticum aestivum L.). *Theoretical and applied genetics*, 109, pp.1105–1114.

Song, Q.J. et al., 2005. Development and mapping of microsatellite (SSR) markers in wheat. *Theoretical and Applied Genetics*, 110, pp.550–560.

Sourdille, P. et al., 2003. An update of the Courtot x Chinese Spring intervarietal molecular marker linkage map for the QTL detection of agronomic traits in wheat. *Theoretical and applied genetics*, 106(3), pp.530–538.

Sourdille, P. et al., 1996. Linkage between RFLP markers and genes affecting kernel hardness in wheat. *Theoretical and Applied Genetics*, 93, pp.580–586.

Stackelberg, M. von, Rensing, S.A. & Reski, R., 2006. Identification of genic moss SSR markers and a comparative analysis of twenty-four algal and plant gene indices reveal species-specific rather than group-specific characteristics of microsatellites. *BMC plant biology*, 6(9).

Stam, P., 1993. Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. *The Plant Journal*, 3(5), pp.739–744.

Sukumaran, S. et al., 2015. Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. *Theoretical and Applied Genetics*, 128, pp.353–363.

Sun, G. et al., 2010. Variation explained in mixed-model association mapping. *Heredity*, 105, pp.333–340.

Sun, H. et al., 2008. Quantitative trait loci (QTLs) for quality traits related to protein and starch in wheat. *Progress in Natural Science*, 18, pp.825–831.

Sun, X. et al., 2010. Mapping quantitative trait loci for quality factors in an inter-class cross of US and Chinese wheat. *Theoretical and applied genetics*, 120, pp.1041–51.

Sun, X.-Y. et al., 2009. QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. *Euphytica*, 165, pp.615–624.

Suprayogi, Y. et al., 2009. Identification and validation of quantitative trait loci for grain protein concentration in adapted Canadian durum wheat populations. *Theoretical and applied genetics*, 119, pp.437–448.

Syed, S.H. & Rivandi, A., 2007. Molecular Breeding for Drought Tolerance in Plants: Wheat Perspective. *Proceedings of Pakistan Academy of Sciences*, 44(1), pp.35–62.

Tanksley, S.D., 1993. Mapping Polygenes. *Annual Review of Genetics*, 27, pp.205–233.

Tautz, D. & Renz, M., 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research*, 12(10), pp.4127–4138.

Tazzini, N., 2015. Gluten: definition, structure, properties, wheat, cereal list. *http://www.tuscany-diet.net/2015/02/08/gluten-properties-structure-wheat-cereal-list/*.

Thornsberry, J.M. et al., 2001. Dwarf8 polymorphisms associate with variation in flowering time. *Nature genetics*, 28, pp.286–289.

Triboï, E., Martre, P. & Triboï-Blondel, A.-M., 2003. Environmentally-induced changes in protein composition in developing grains of wheat are related to changes in total protein content. *Journal of experimental botany*, 54(388), pp.1731–1742.

Troccoli, A. et al., 2000. Mini Review: Durum Wheat Quality: A Multidisciplinary Concept. *Journal of Cereal Science*, 32, pp.99–113.

Tsilo, T.J. et al., 2010. Genome mapping of kernel characteristics in hard red spring wheat breeding lines. *Theoretical and Applied Genetics*, 121, pp.717–730.

Turner, A.S. et al., 2004. New quantitative trait loci influencing grain texture and protein content in bread wheat. *Journal of Cereal Science*, 40, pp.51–60.

Turner, N., 1986. Adaptation to Water Deficits: a Changing Perspective. *Australian Journal of Plant Physiology*, 13, p.175190.

Tyagi, S. et al., 2014. Interval mapping and meta-QTL analysis of grain traits in common wheat

(Triticum aestivum L.). *Euphytica*, 201, pp.367–380.

UN, 2013. http://www.un.org/en/development/desa/news/population/un-report-world-population-projected-to-reach-9-6-billion-by-2050.html.

Utz, H.F. & Melchinger, A.E., 1996. PLABQTL: A program for composite interval mapping of QTL. *Universitaet Hohenheim, D-70593 Stuttgart, Germany*.

Veraverbeke, W.S. & Delcour, J.A., 2002. Wheat protein composition and properties of wheat glutenin in relation to breadmaking functionality. *Critical reviews in food science and nutrition*, 42(3), pp.179–208.

Vishwakarma, M.K. et al., 2014. Introgression of the high grain protein gene Gpc-B1 in an elite wheat variety of Indo-Gangetic Plains through marker assisted backcross breeding. *Current Plant Biology*, 1, pp.60–67.

Visscher, P.M., 2008. Sizing up human height variation. *Nature genetics*, 40(5), pp.489–490.

Wall, J.S., 1979. The role of wheat protein in determining baking quality. *In: Recent advances in the biochemistry of cereals. Ed Laidman, D.L. and Wyn Jones R.G. London Academy Pressl*, pp.275–311.

Wang, L. et al., 2012. Conditional QTL mapping of protein content in wheat with respect to grain yield and its components. *Journal of Genetics*, 91, pp.303–312.

Wang, S. et al., 2014. Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnology Journal*, 12, pp.787–796.

Wanjugi, H. et al., 2009. Rapid development of PCR-based genome-specific repetitive DNA junction markers in wheat. *Genome*, 52, pp.576–587.

Weightman, R.M. et al., 2008. Effects of drought and the presence of the 1BL/1RS translocation on grain vitreosity, hardness and protein content in winter wheat. *Journal of Cereal Science*, 47, pp.457–468.

Wen, S. et al., 2012. Structural genes of wheat and barley 5-methylcytosine DNA glycosylases and their potential applications for human health. *Proceedings of the National Academy of Sciences of the United States of America*, 109(50), pp.20543–20548.

Wiersma, J.J. et al., 2001. Recurrent selection for kernel weight in spring wheat. *Crop Science*, 41, pp.999–1005.

Winter, P. & Kahl, G., 1995. Molecular marker technologies for plant improvement. *World Journal of Microbiology & Biotechnology*, 11, pp.438–448.

Wu, X., Chang, X. & Jing, R., 2012. Genetic insight into yield-associated traits of wheat grown

in multiple rain-fed environments. *PLoS ONE*, 7(2), p.e31249.

Xu, Y. et al., 1997. Chromosomal regions associated with segregation distortion of molecular markers in F2, backcross, doubled haploid, and recombinant inbred populations in rice (Oryza sativa L.). *Molecular and General Genetics*, 253, pp.535–545.

Xu, Y., 2010. Molecular Plant Breeding. *International Maize and Wheat Improvement Center*.

Xu, Z.Z., Zhou, G.S. & Shimizu, H., 2009. Effects of soil drought with nocturnal warming on leaf stomatal traits and mesophyll cell ultrastructure of a perennial grass. *Crop Science*, 49, pp.1843–1851.

Yildirim, A. et al., 2013. Marker-assisted breeding of a durum wheat cultivar for γ-gliadin and LMW-glutenin proteins affecting pasta quality. *Turkish Journal of Agriculture and Forestry*, 37, pp.527–533.

Young, N.D., 1994. Constructing a plant genetic linkage map with DNA markers. *In: I. K.W. Ronald & L. Phillips (Eds.), DNA-based markers in plants. Kluwer, Dordrecht/Boston/London.*, pp.39–57.

Yu, J. et al., 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature genetics*, 38(2), pp.203–208.

Yu, J. & Buckler, E.S., 2006. Genetic association mapping and genome organization of maize. *Current Opinion in Biotechnology*, 17, pp.155–160.

Yu, L.X. et al., 2011. Association mapping and gene-gene interaction for stem rust resistance in CIMMYT spring wheat germplasm. *Theoretical and Applied Genetics*, 123, pp.1257–1268.

Zadoks, J.C., Chang, T.T. & Konzak, C.F., 1974. A decimal code for the growth stages of cereals. *Weed Research*, 14, pp.415–421.

Zane, L., Bargelloni, L. & Patarnello, T., 2002. Strategies for microsatellite isolation: A review. *Molecular Ecology*, 11, pp.1–16.

Zeng, Z.B., 1994. Precision mapping of quantitative trait loci. *Genetics*, 136, pp.1457–1468.

Zhang, H. et al., 2013. Conditional and unconditional QTL mapping of drought-tolerance-related traits of wheat seedling using two related RIL populations. *Journal of Genetics*, 92(2), pp.213–231.

Zhang, J. et al., 2014. Genetic Dissection of QTL Associated with Grain Yield in Diverse Environments. *Agronomy*, 4, pp.556–578.

Zhang, S. et al., 2014. Development of highly polymorphic simple sequence repeat markers using genome-wide microsatellite variant analysis in Foxtail millet [Setaria italica (L.) P. Beauv.]. *BMC genomics*, 15(78).

Zhang, W. et al., 2008. QTL analysis of pasta quality using a composite microsatellite and SNP map of durum wheat. *Theoretical and Applied Genetics*, 117, pp.1361–1377.

Zhang, Y.L. et al., 2009. QTL mapping for milling, gluten quality, and flour pasting properties in a recombinant inbred line population derived from a Chinese soft x hard wheat cross. *Crop & Pasture Science*, 60, pp.587–597.

Zhang, Z. et al., 2010. Mixed linear model approach adapted for genome-wide association studies. *Nature genetics*, 42(4), pp.355–360.

Zhao, C.-X., Zhu, R.-L. & Liu, Y., 2014. Simple sequence repeats in bryophyte mitochondrial genomes. *Mitochondrial DNA*, pp.1–7.

Zheng, S. et al., 2009. Association analysis reveals effects of wheat glutenin alleles and rye translocations on dough-mixing properties. *Journal of Cereal Science*, 50, pp.283–290.

Zhu, C. et al., 2008. Status and Prospects of Association Mapping in Plants. *The Plant Genome*, 1(1), pp.5–20.

APPENDIX

Table 36. Precipitation and irrigation from January to July 15 for 13 environments

Environments	Populations†	Precipitation (mm)	Irrigation (mm)	Total (mm)
2007/08 Fort Collins Dry	DH	65	70	135
2007/08 Fort Collins Wet	DH	65	178	243
2008/09 Greeley Dry	DH	118	77	195
2008/09 Greeley Wet	DH	118	141	259
2009/10 Akron rainfed	RIL	207	0	207
2009/10 Greeley irrigated	RIL	251	195	446
2011/12 Greeley Dry	AM	82	102	184
2011/12 Greeley Wet	AM	82	335	417
2012/13 Fort Collins Dry	AM	124	0	124
2012/13 Fort Collins Wet	AM	124	222	364
2012/13 Greeley Dry	BC	105	38	143
2012/13 Greeley Wet	BC	105	114	219
2012/13 Fort Collins Wet	BC	124	216	340

Data obtained from http://www.coagmet.colostate.edu/, accessed on May 6th 2015.

† DH, 185 CO940610/Platte doubled haploid lines; RIL, 186 CO940610/Platte recombinant inbred lines; AM, Hard Winter Wheat Association Mapping Panel of 299 lines; BC, 35 CO940610/Platte BC_3F_2 lines.

Table 37. Monthly maximum and minimum temperature (°C) and precipitation (mm) from January to July 15 for seven location-years, where experiments involved were conducted.

		Fort Collins 2008	Greeley 2009	Akron 2010	Greeley 2010	Greeley 2012	Fort Collins 2013	Greeley 2013
		DH†	DH	RIL	RIL	AM	AM, BC	BC
January	Tmax‡	13.8	n/a	13.3	11.2	18.0	17.8	15.0
	Tmin	-19.3	n/a	-25.6	-26.7	-16.7	-22.8	-25.6
	Prec	0.0	n/a	0.0	0.8	0.3	0.0	1.0
February	Tmax	17.2	n/a	10.6	10.3	18.3	15.6	16.9
	Tmin	-15.6	n/a	-19.2	-19.2	-15.8	-16.9	-19.2
	Prec	0.0	n/a	7.6	6.6	10.7	2.5	6.1
March	Tmax	23.1	n/a	29.4	27.7	27.9	23.9	25.2
	Tmin	-12.8	n/a	-7.8	-10.9	-17.1	-15.0	-16.6
	Prec	11.4	n/a	20.3	7.1	0.5	5.1	8.4
April	Tmax	27.7	n/a	28.9	26.7	30.7	26.1	27.1
	Tmin	-8.3	n/a	-4.4	-7.1	-4.4	-14.3	-15.3
	Prec	40.1	n/a	45.7	84.3	23.1	30.5	19.6
May	Tmax	30.3	27.7	33.9	34.1	33.2	30.0	31.1
	Tmin	-7.3	4.4	-3.9	-2.6	0.6	-11.1	-6.9
	Prec	10.4	21.3	43.2	50.0	28.2	35.6	39.9
June	Tmax	32.8	33.1	36.7	35.7	37.4	36.7	36.6
	Tmin	7.3	6.5	6.1	6.7	8.6	3.9	4.6
	Prec	0.0	87.4	58.4	80.5	19.3	12.7	8.4
July	Tmax	36.4	35.9	37.8	35.6	37.6	36.7	37.4
	Tmin	8.7	8.7	8.9	9.7	12.4	10.6	10.2
	Prec	12.2	48.0	48.3	41.7	41.7	38.1	28.7

Data obtained from http://www.coagmet.colostate.edu/, accessed on November 3rd 2015.

† DH, 185 CO940610/Platte doubled haploid lines; RIL, 186 CO940610/Platte recombinant inbred lines; AM, Hard Winter Wheat Association Mapping Panel of 299 lines; BC, 35 CO940610/Platte BC_3F_2 lines.

‡ Tmax or Tmin, maximum or minimum temperature, respectively; Prec, precipitation.

n/a, data are not available for the whole month

Colored cells indicate that data are not available for one or more days in the month: Red, 7th to 23th; Yellow, 20th to 31st; Green, 1st to 16th; Blue, 1st to 20th; and Purple, 16th.

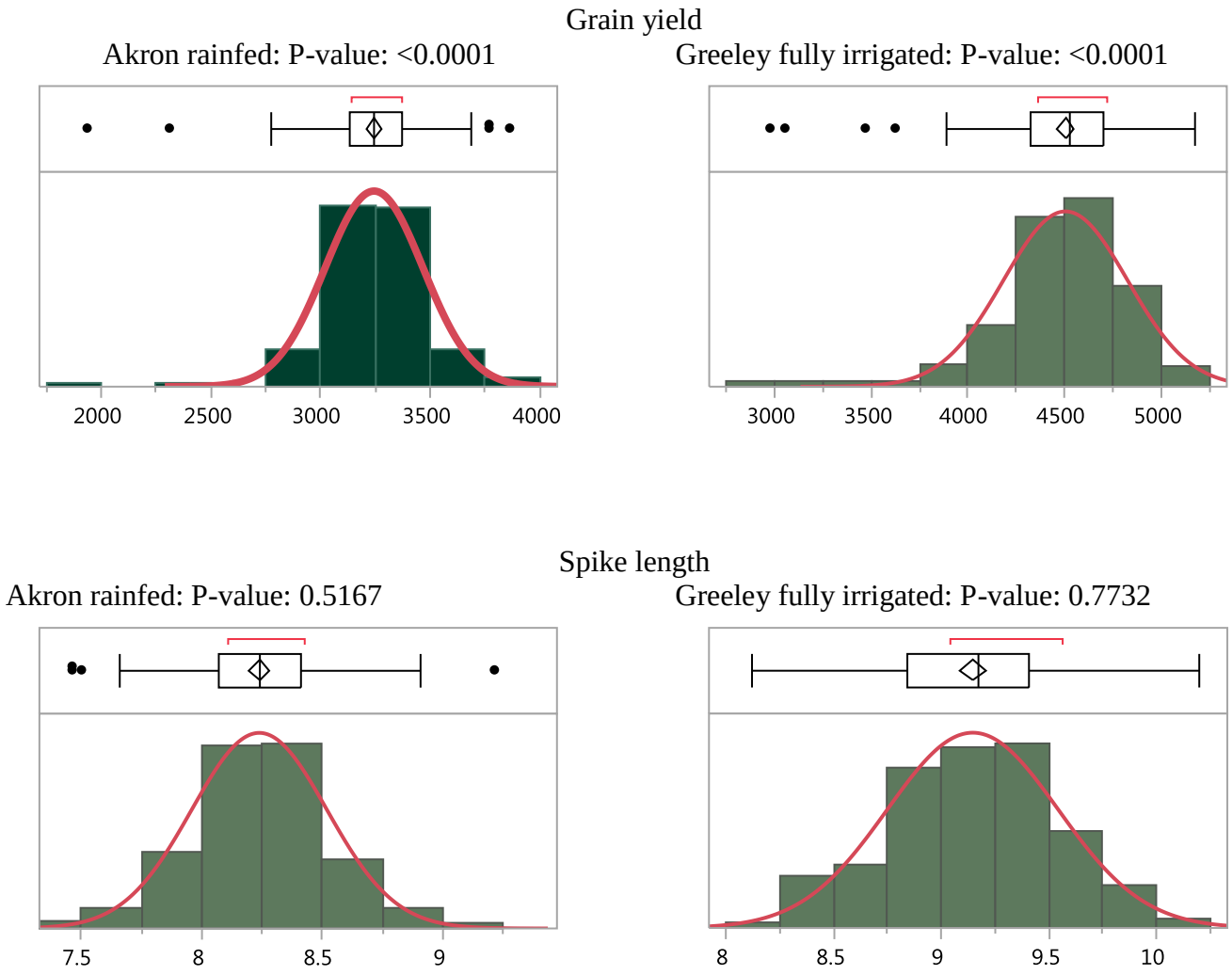
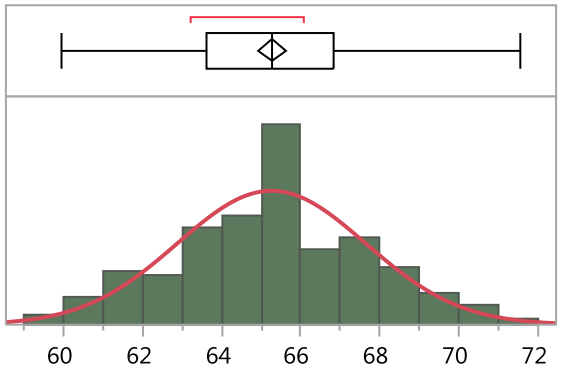


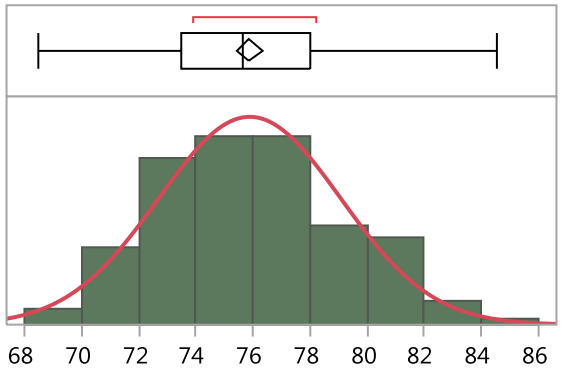
Figure 13. Frequency distributions for the traits for CO940610/Platte RIL population in 2009/10 growing season. *P*-values are for the Shapiro-Wilk test of normality, with * ($P<0.05$) indicating deviation for normality.

Plant height

Akron rainfed: P-value: 0.5321

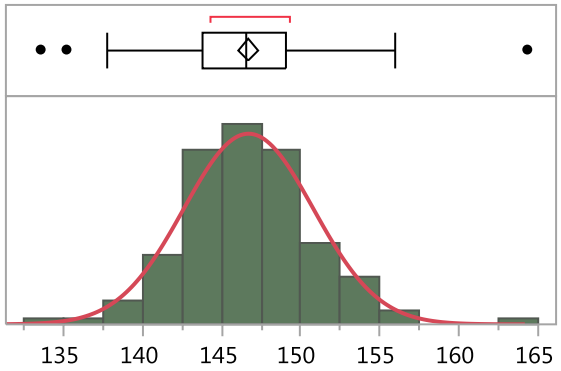


Greeley fully irrigated: P-value: 0.4753



Grain protein concentration

Akron rainfed: P-value: 0.0185



Greeley fully irrigated: P-value: <0.0001

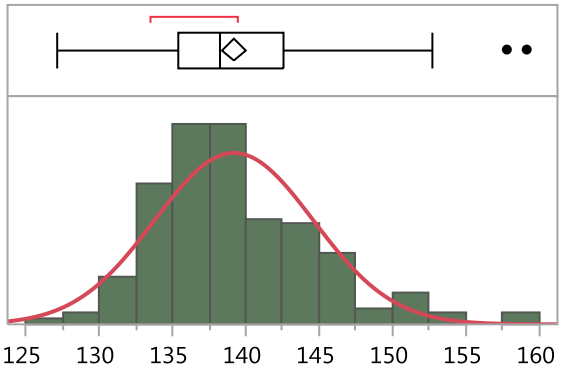


Figure 13. (Continue)

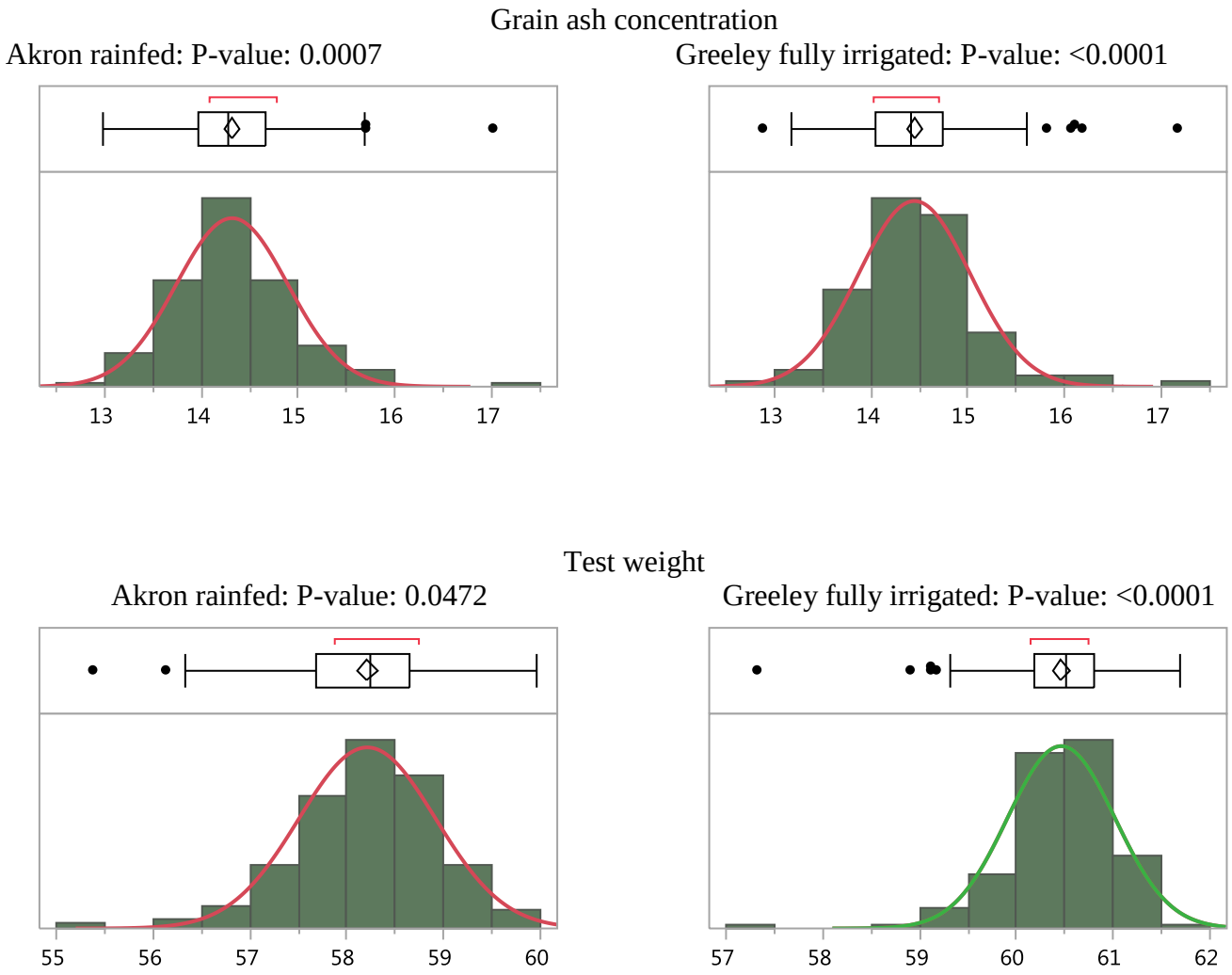


Figure 13. (Continue)

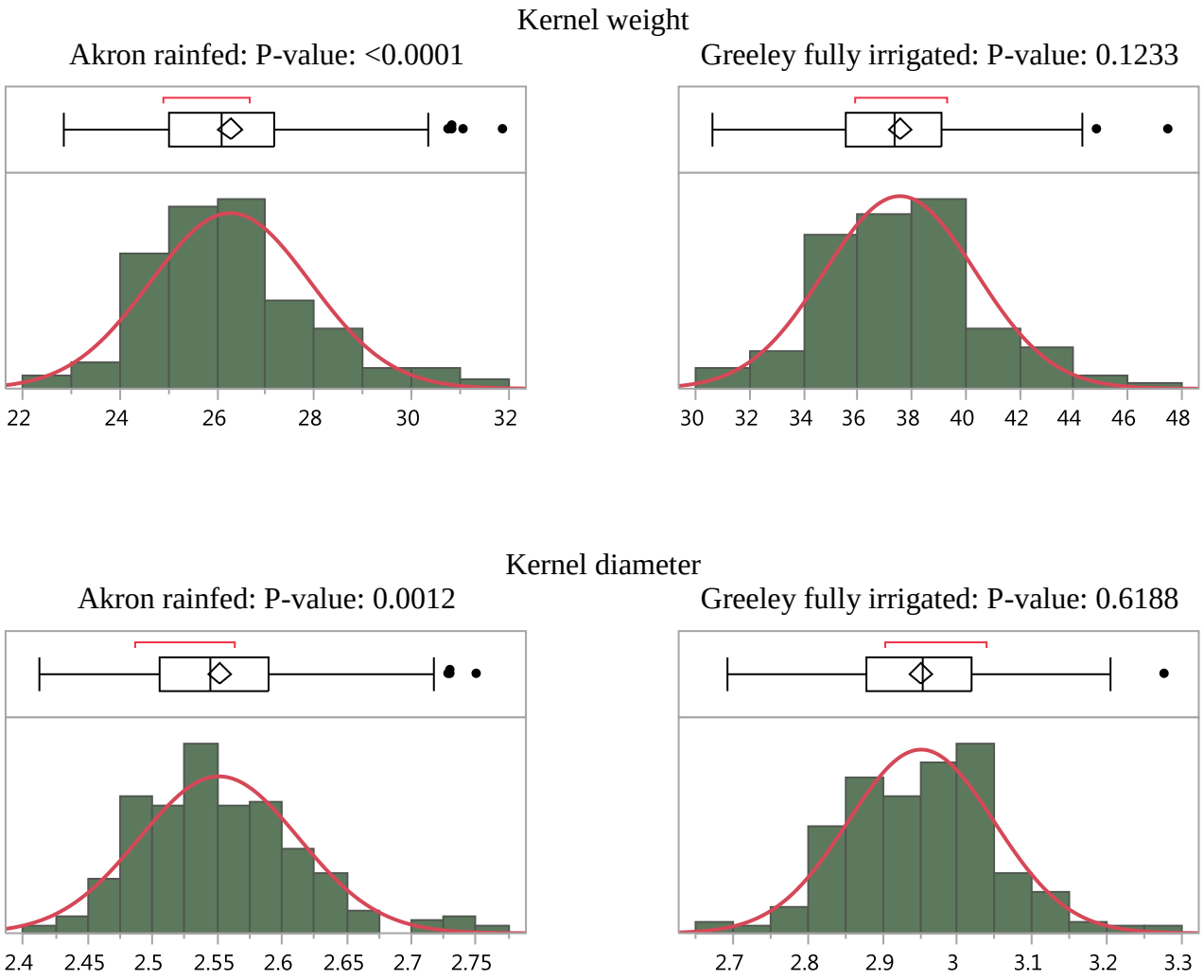


Figure 13. (Continue)

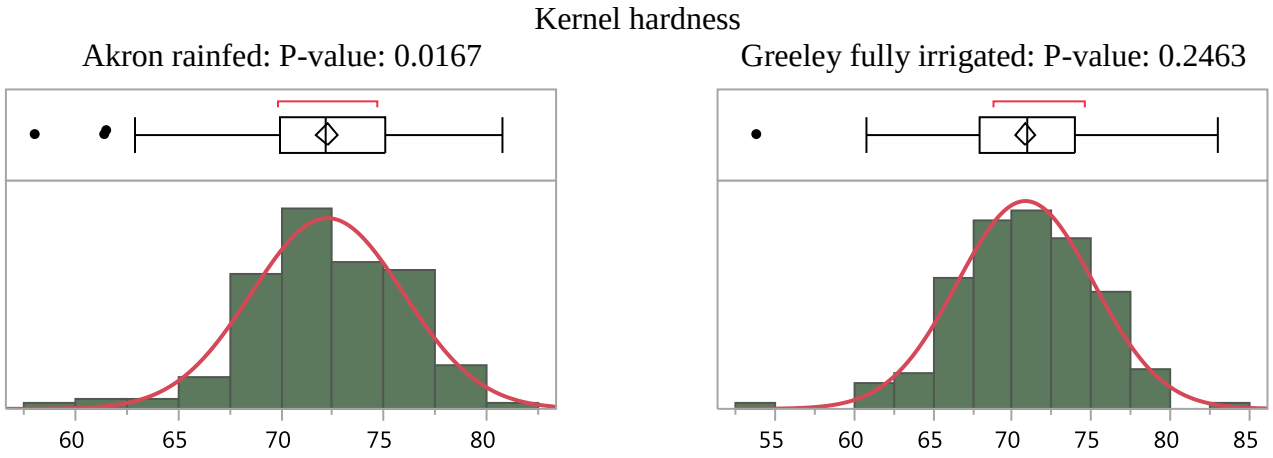


Figure 13. (Continue)

Plant height

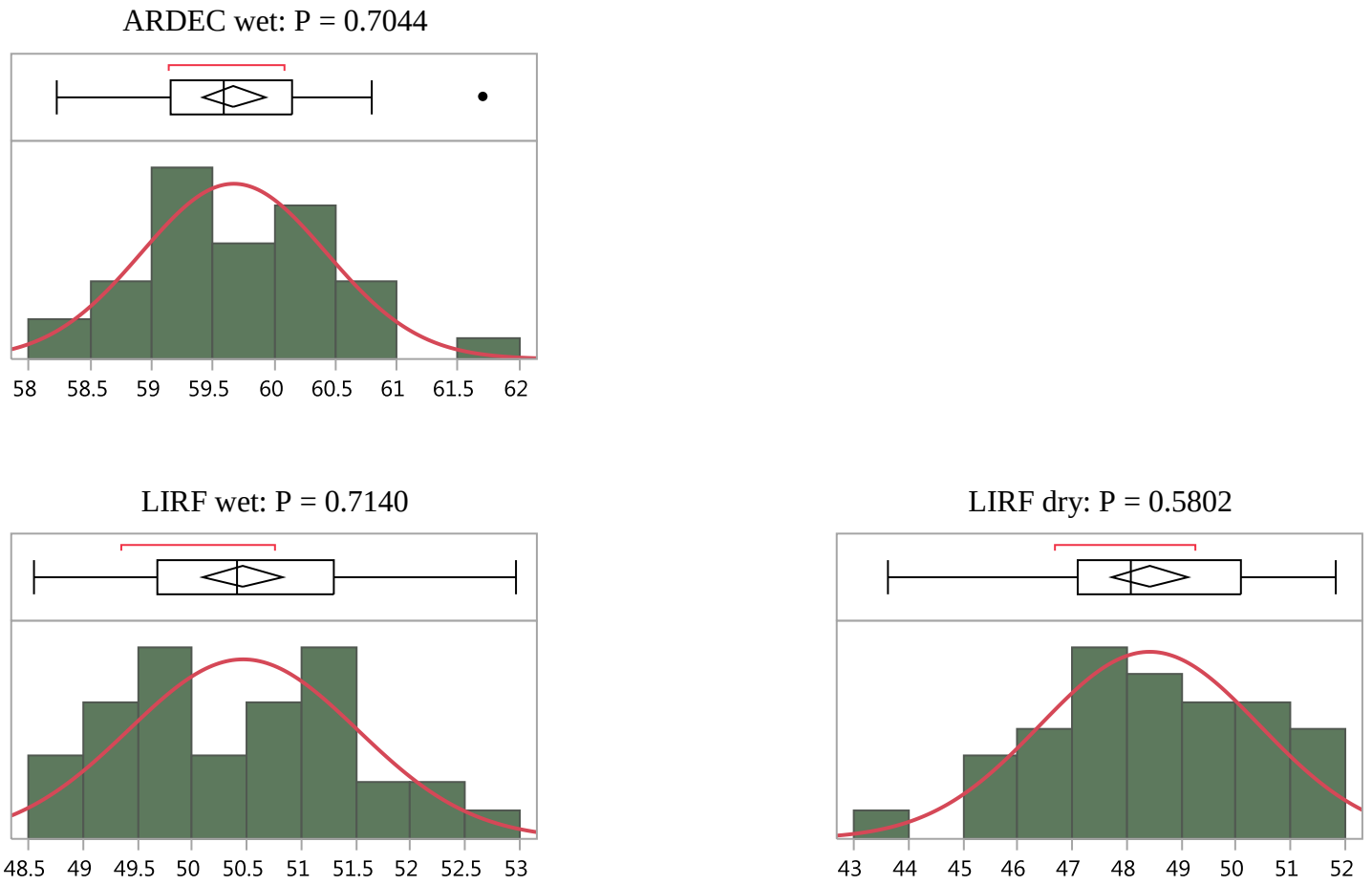
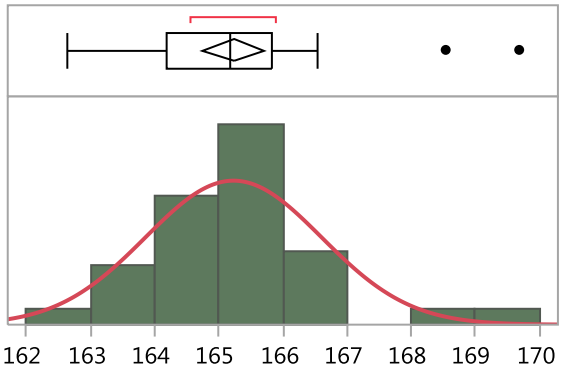


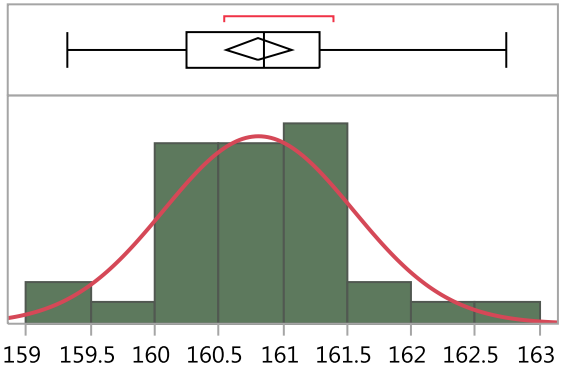
Figure 14. Frequency distributions for seven traits for the CO940610/Platte//Platte backcross populations in the 2013 growing seasons. *P*-values are for the Shapiro-Wilk Test of normality, with * ($P<0.05$) indicating deviation from normality.

Grain protein concentration

ARDEC wet: P = 0.0197*



LIRF wet: P = 0.6162



LIRF dry: P = 0.0295*

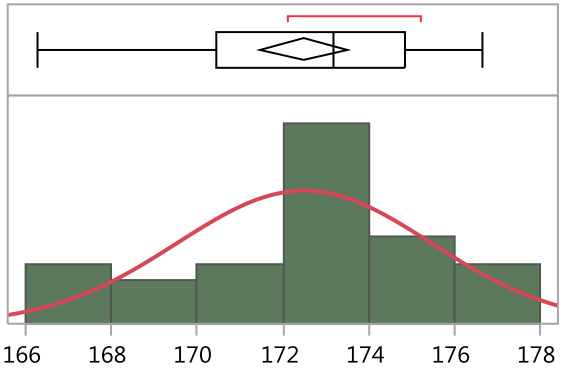
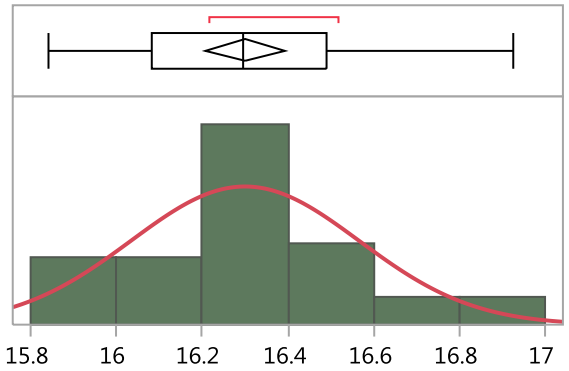


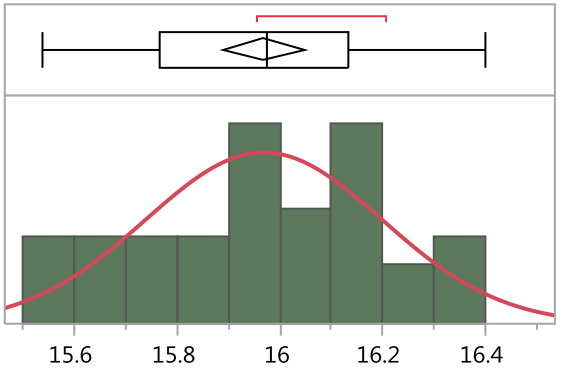
Figure 14. (Continued)

Grain ash concentration

ARDEC wet: P = 0.5597



LIRF wet: P = 0.4468



LIRF dry: P = 0.4406

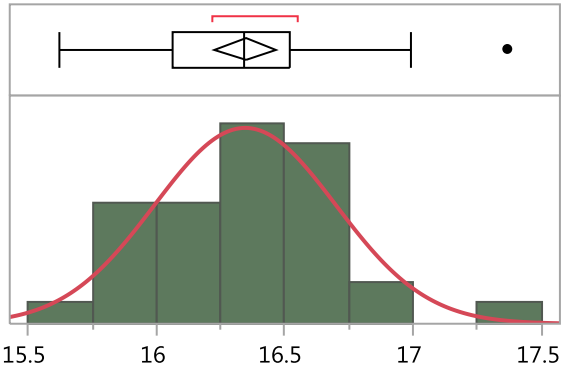


Figure 14. (Continued)

Test weight

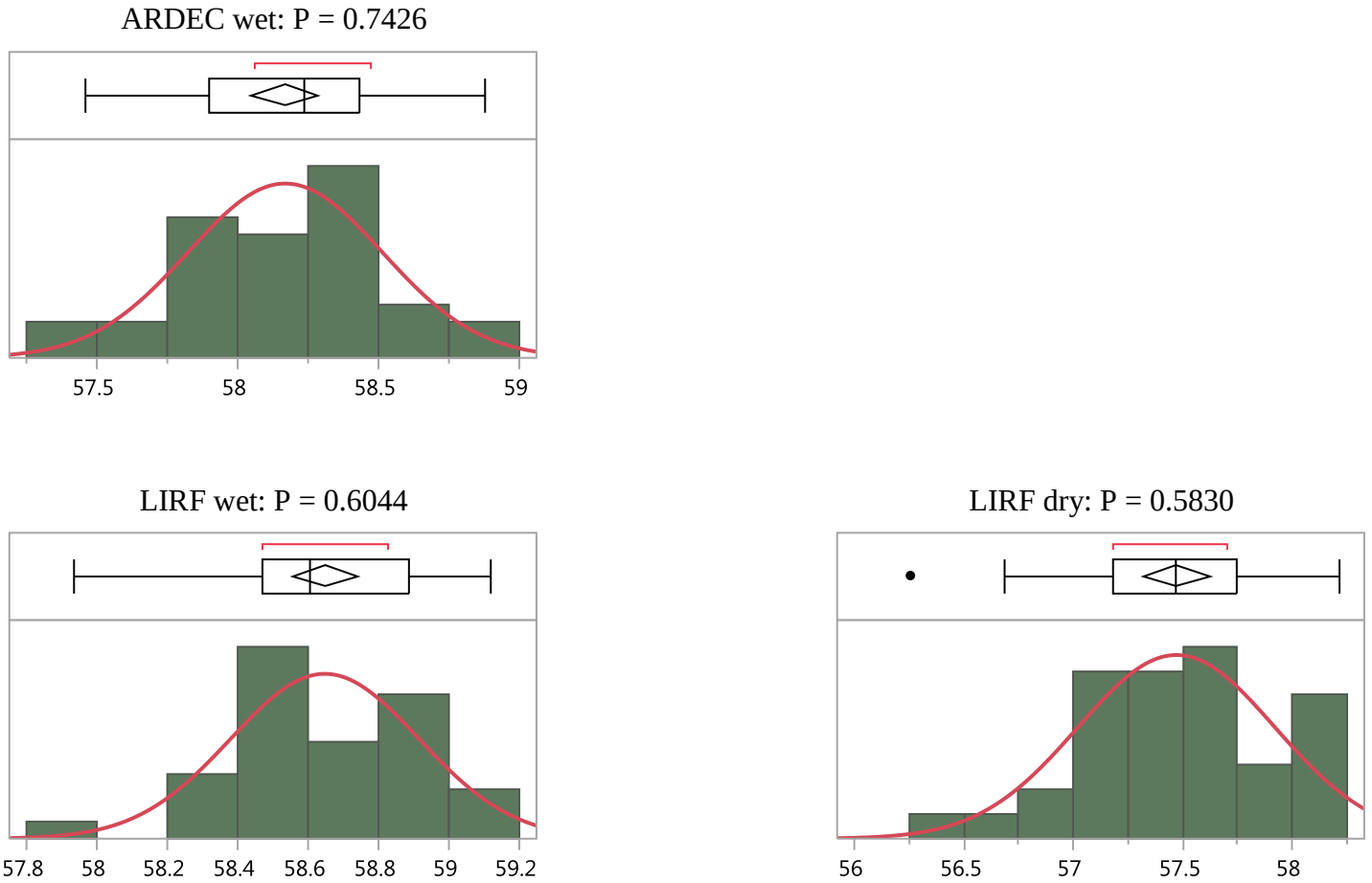
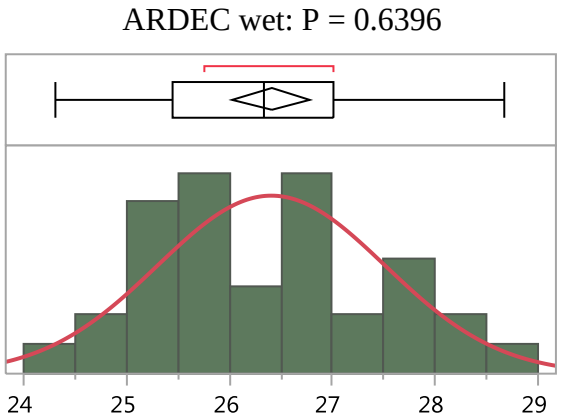
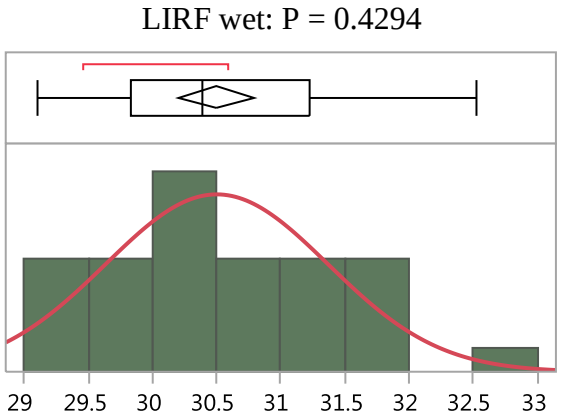


Figure 14. (Continued)

Kernel weight





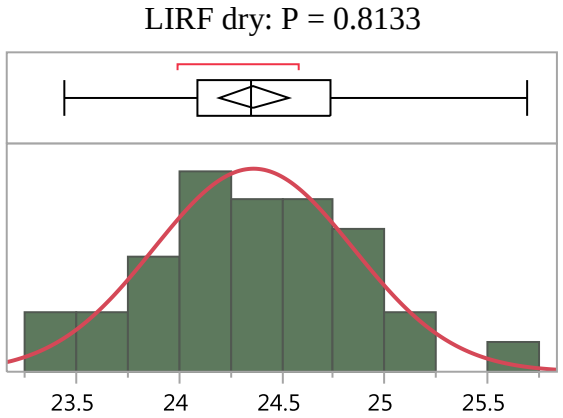


Figure 14. (Continued)

Kernel diameter

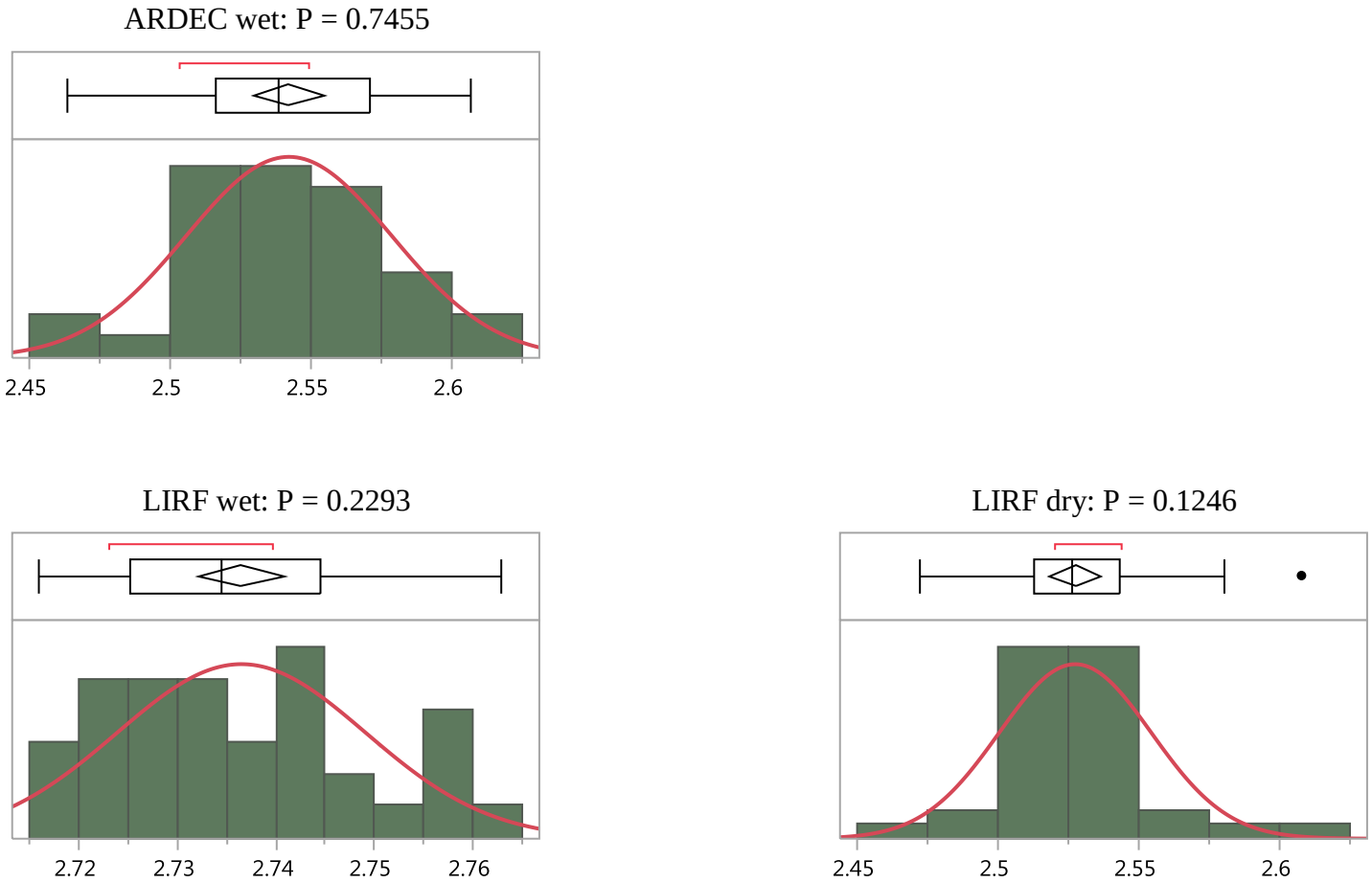
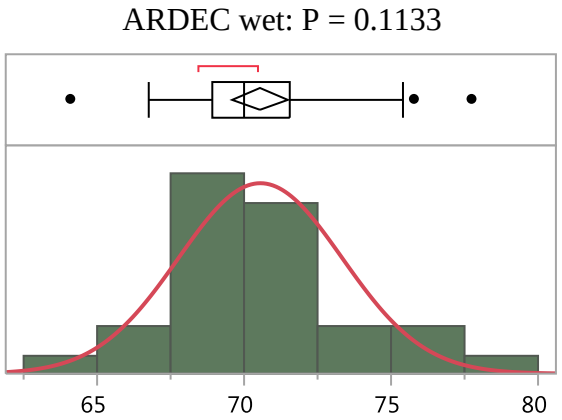
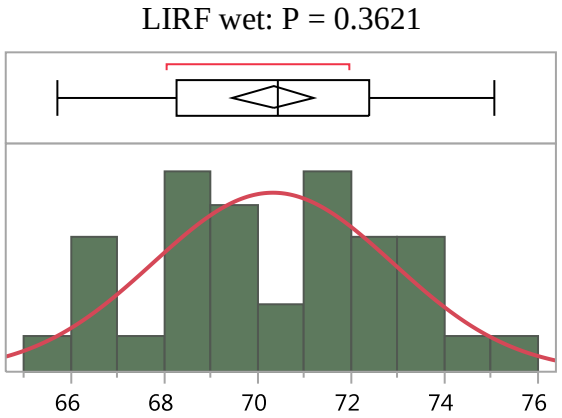


Figure 14. (Continued)

Kernel hardness





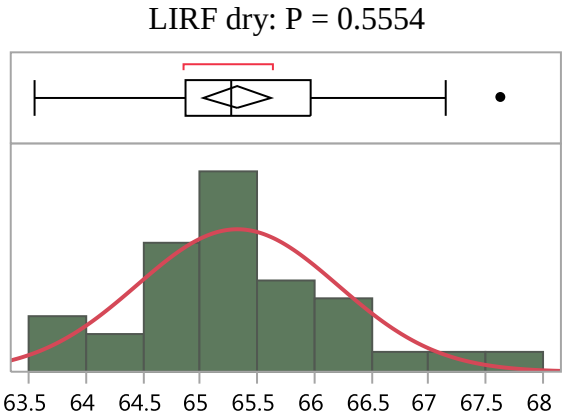


Figure 14. (Continued)

Grain yield

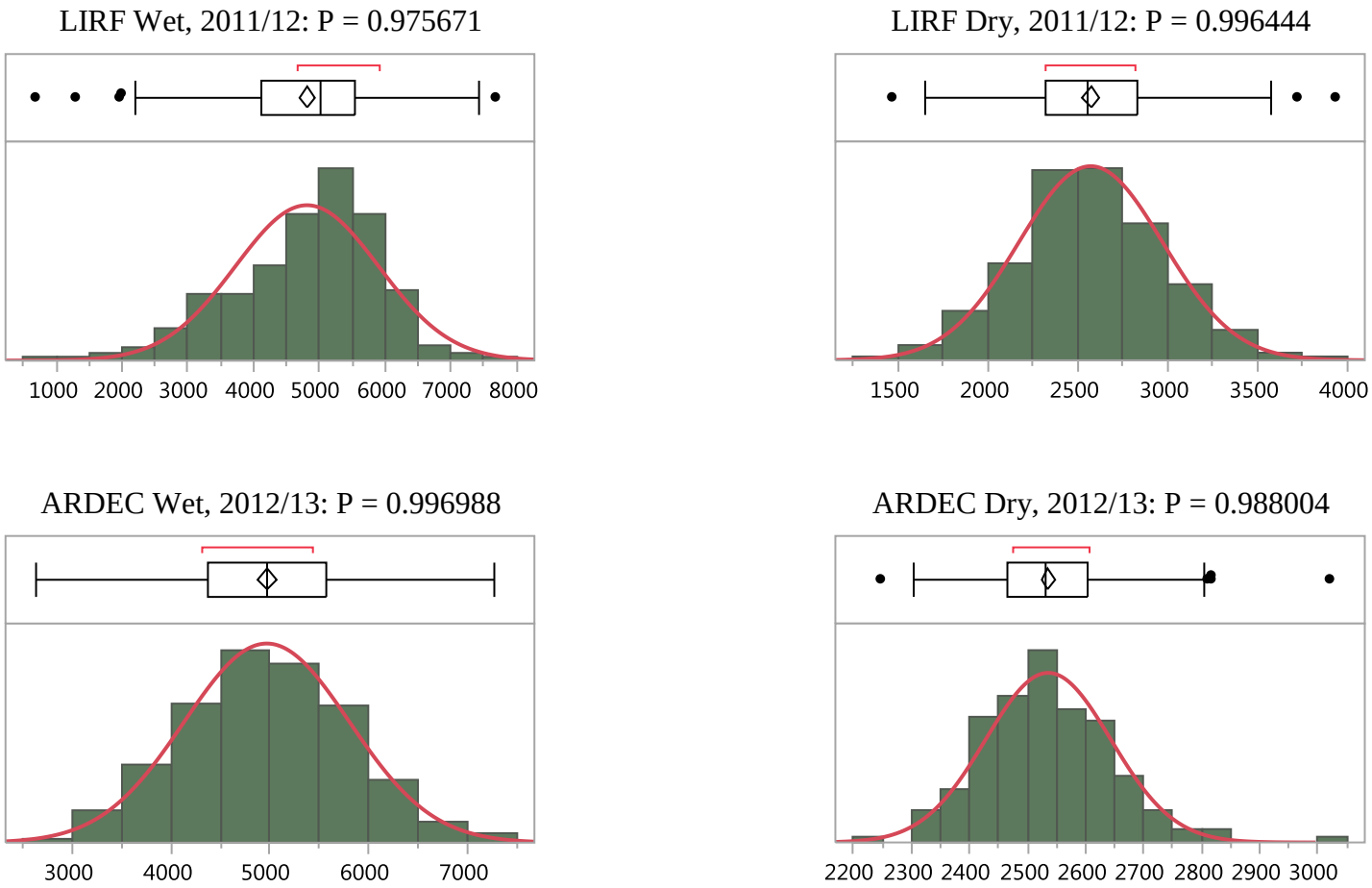


Figure 15. Frequency distributions for five traits investigated in the Hard Winter Wheat Association Mapping Panel. *P*-values are for the Shapiro-Wilk Test of normality, with * ($P<0.05$) indicating deviation from normality.

Grain protein concentration

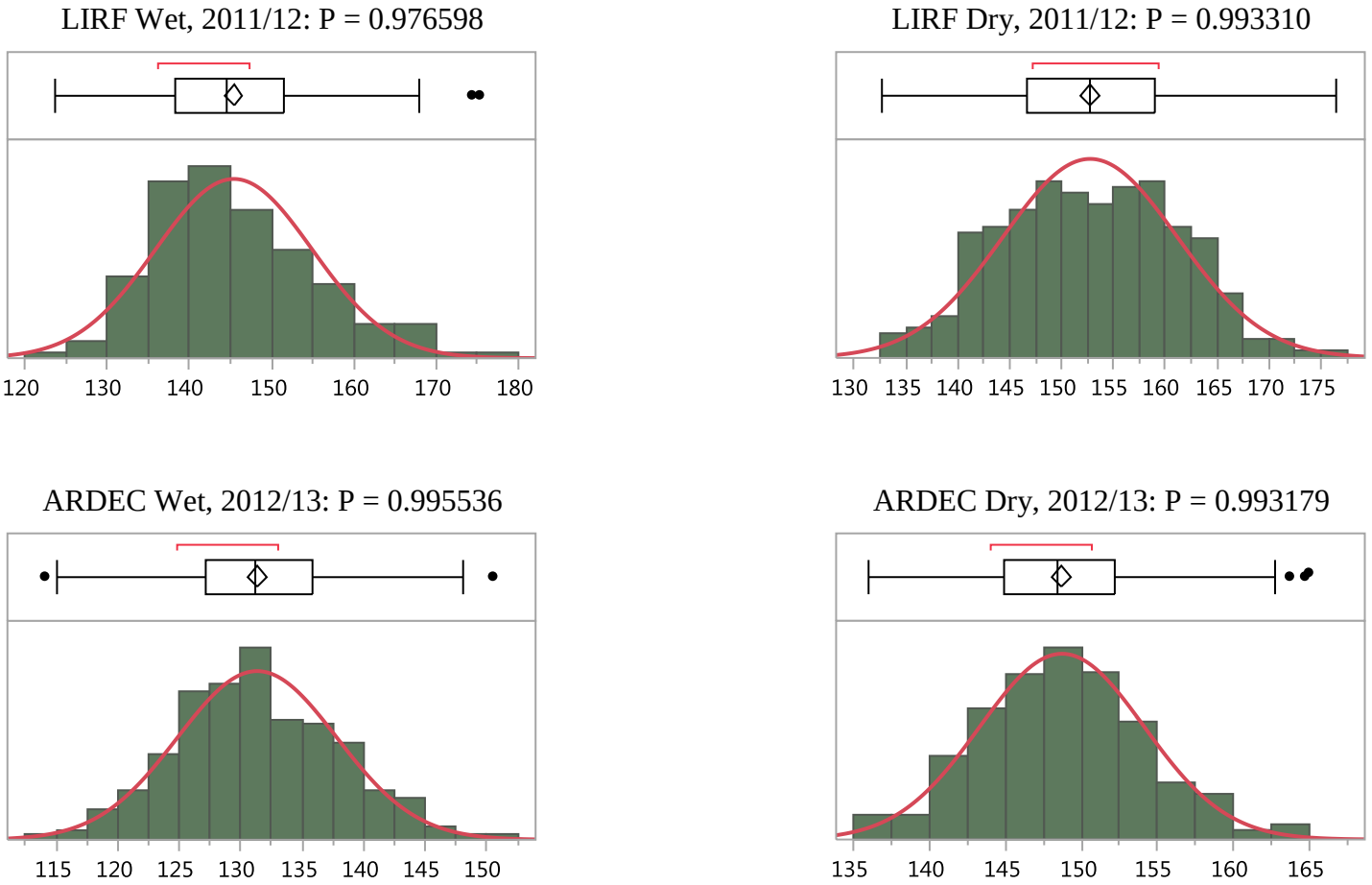
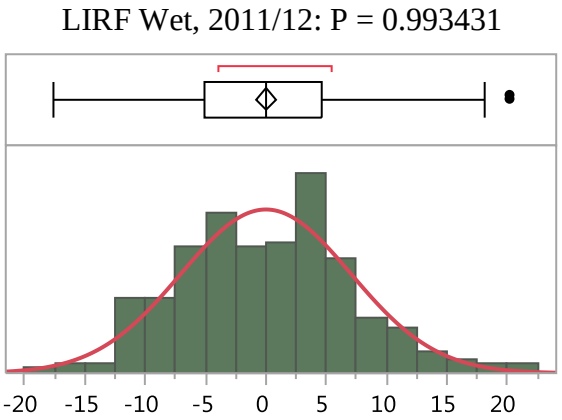
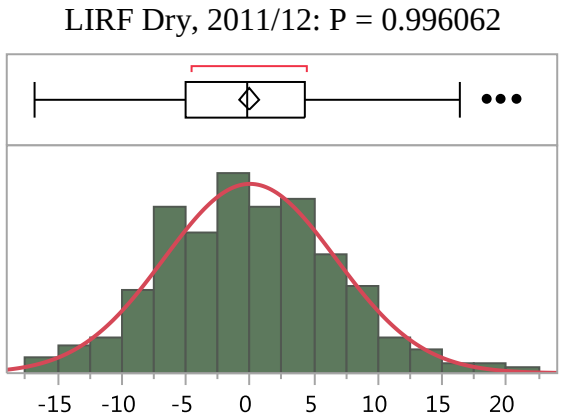
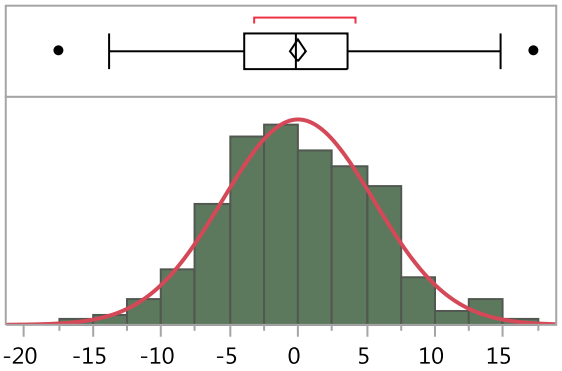


Figure 15. (Continued)

Grain protein deviation







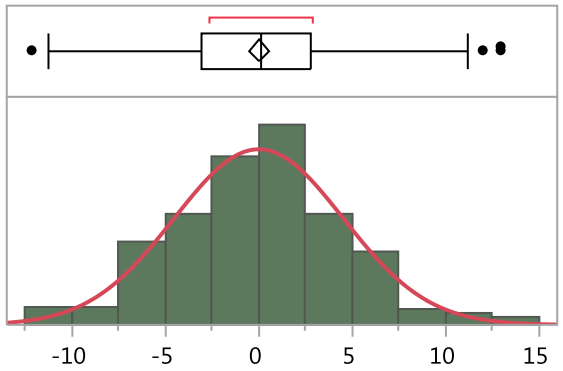


Figure 15. (Continued)

Grain ash concentration

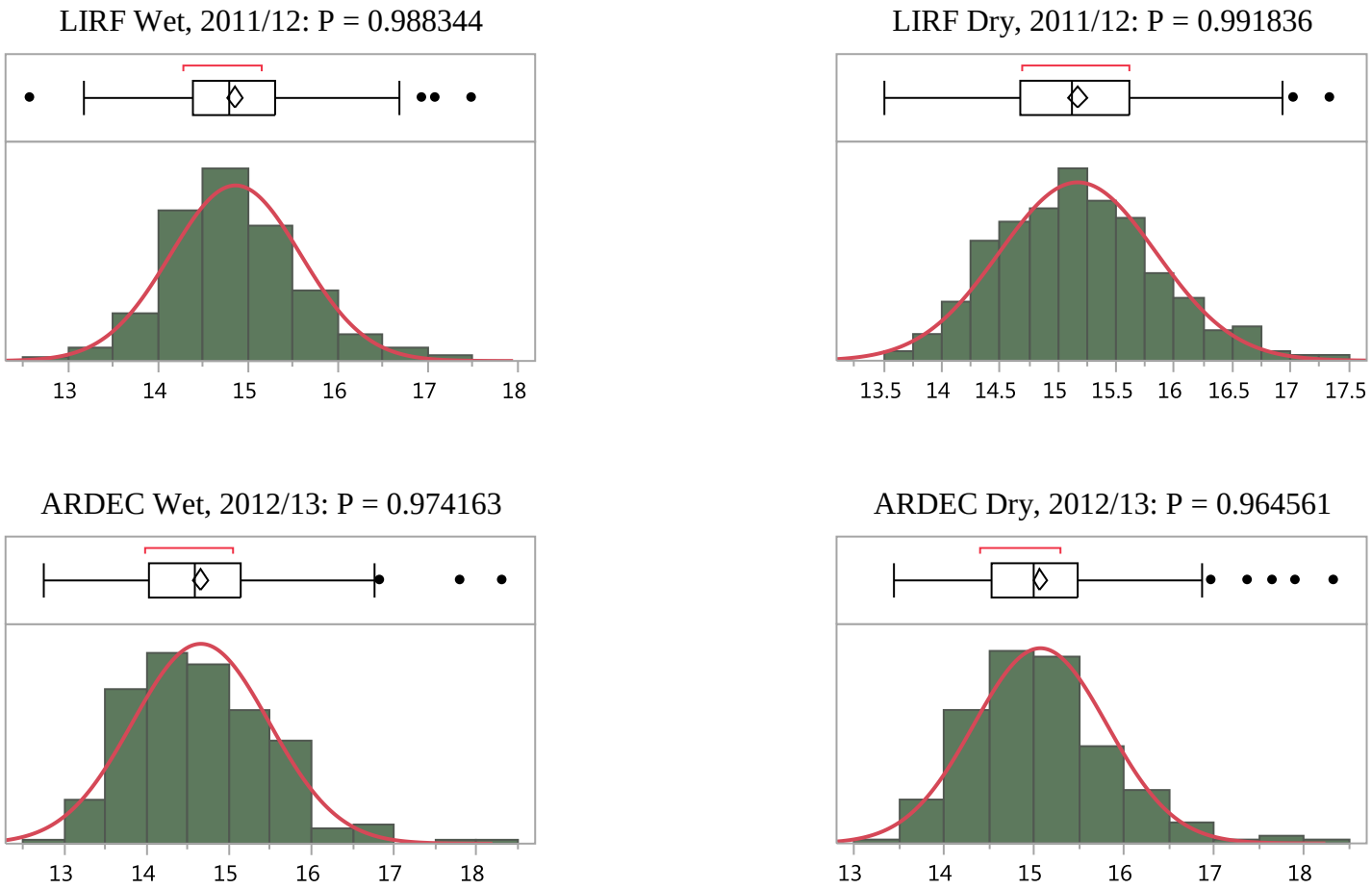
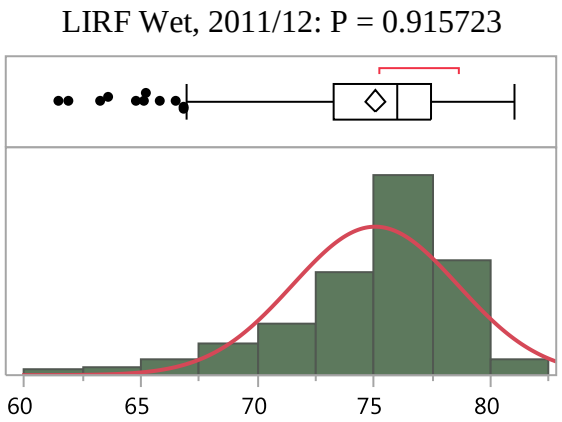
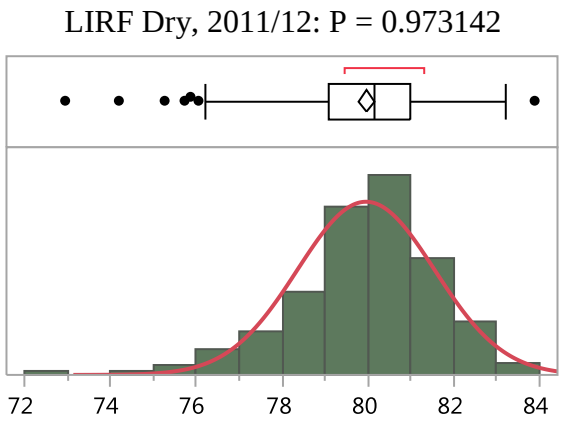
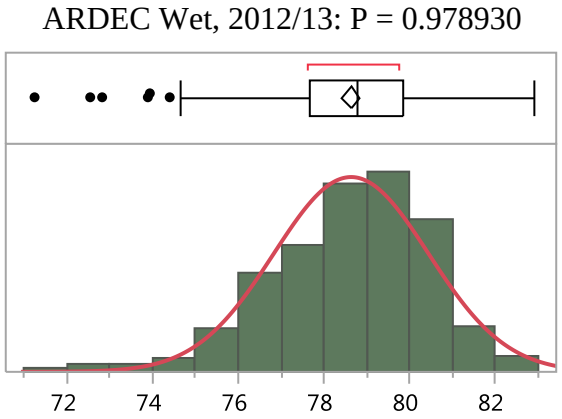


Figure 15. (Continued)

Test weight







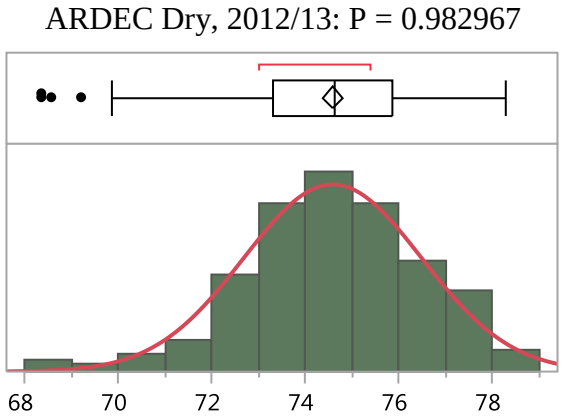


Figure 15. (Continued)