DISSERTATION

ASSOCIATION MAPPING FOR YIELD, YIELD COMPONENTS AND DROUGHT TOLERANCE-RELATED TRAITS IN SPRING WHEAT GROWN UNDER RAINFED AND IRRIGATED CONDITIONS

Submitted by

Erena Aka Edae

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

Summer 2013

Doctoral Committee:

Advisor: Patrick Byrne
Co-Advisor: Scott Haley

William Black
Eric Storlie

ABSTRACT

ASSOCIATION MAPPING FOR YIELD, YIELD COMPONENTS AND DROUGHT TOLERANCE-RELATED TRAITS IN SPRING WHEAT GROWN UNDER RAINFED AND IRRIGATED CONDITIONS

Genome-wide association mapping shows promise for identifying quantitative trait loci (QTL) for many traits including drought stress tolerance. Candidate gene analysis also has been used to identify functional single nucleotide polymorphisms (SNPs) that can be associated with important traits. In 2010 and 2011, we evaluated an International maize and wheat improvement center (CIMMYT) spring wheat association mapping panel under rainfed and full irrigation conditions in Greeley, CO, and Melkassa, Ethiopia (total of five environments) for grain yield and its components, canopy spectral reflectance, and several other phenological or drought-related traits. A total of 287 lines were genotyped with Diversity Array Technology (DArT) markers to identify associations with measured traits under different moisture regimes. Significant differences among lines were observed for most traits within each environment and across environments. Best linear unbiased predictors (BLUPs) of each line were used to calculate marker-trait associations using 1863 markers with a mixed linear model with population structure and a kinship-matrix included as covariates. Three drought responsive candidate genes (Dehydration-Responsive Element Binding 1A, *DREB1A*; Enhanced Response to abscisic acid (ABA), *ERA1*; and Fructan 1-exohydrolase, *1-FEH*), were amplified using genome-specific primers and sequenced from 126 lines to identify single nucleotide polymorphisms (SNPs) within the candidate genes and determine their association with measured traits. For genome wide association mapping, the highest number of stable associations was obtained for kernel

hardness followed by grain volume weight (test weight), an important trait under drought stress conditions. The most stable marker-trait association was obtained for grain yield on chromosome 2DS. All marker-trait associations for above-ground biomass were environment-specific. Multi-trait marker-trait association for grain yield and other traits such as harvest index, final biomass, thousand kernel weight, plant height and flag leaf length were detected on chromosome 5B. A grain yield QTL was again co-localized with harvest index QTL on chromosome 1BS. Normalized difference vegetation index (NDVI) shared QTL region with a harvest index QTL on chromosome 1AL, while green leaf area shared a QTL with harvest index on chromosomes 5A. For drought tolerance candidate genes, SNPs within *DREB1A* gene were associated with final biomass, spikelets per spike, days to heading and NDVI. The *1-FEH* gene amplified from the A genome showed associations with grain yield, final biomass, NDVI, green leaf area, kernel number per spike and spike length. However, *1-FEH* from the B genome was associated with traits such as days to heading, days to maturity, thousand kernel weight and test weight. The *ERA1* gene from the B genome was associated with spike m^{-2}, harvest index, grain filling duration, leaf senescence, flag leaf width, plant height and spike length, whereas *ERA1* from the D genome was associated with kernel weight per spike, flag leaf width, leaf senescence, kernel number per spike and harvest index. In general, each candidate gene had effects on multiple traits under both rainfed and irrigated conditions. Both genome wide and candidate gene approaches showed that most of the measured traits are controlled by several QTL/genes with minor effects. QTL/genes with pleotropic effects were also detected. Therefore, the information generated by this study might be used in marker-assisted selection to improve drought tolerance of wheat.

AKNOWLEDGMENTS

I would like to thank my advisors Dr. Patrick Byrne and Dr. Scott Haley for their extraordinary guidance and mentoring throughout my study at Colorado State University. I would like to thank my committee members Dr. William Black and Dr. Eric Storlie for their valuable comments during writing of this thesis. Dr. Philip Chapman also deserves a credit for his advice on statistical data analyses.

I wish to extend my thanks to research associates Scott Reid, Emily Hudson-Arns, Scott Seifert, Victoria Valdez, John Stromberger and Rebecca Kottke for their support on field and laboratory data generation.

I am grateful to the financial support from Beachell-Borlaug International Scholarship Program and Alliance for Graduate Education and the Professoriate (AGEP).

Finally, I extend my special appreciation to my wife, Almaz Bulcha, and my sons, Naol and Sena Erena for their understanding and encouragement during my study at Colorado State University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

1.0 LITERATURE REVIEW

1.1 Wheat production and importance

Hexaploid wheat (*Triticum aestivum* L.) (2n=6x=42) has a large genome size of about 17,300 Mb which is approximately 35 times and 110 times larger than that of rice (*Oryza sativa* L.) and Arabidopsis, respectively (Hussain and Rivandi, 2007). Hexaploid wheat is an allopolyploid (AABBDD) formed first through hybridization of *Triticum urartu* (2n=2x=14, AA) with an unknown source of the B genome, despite speculation tending toward *Aegilops speltoides* (2n=2x=14, BB), and subsequently hybridization with *Aegilops tauschii* (2n=2x=14, DD) (Daud and Gustafson, 1996; Devos and Gale 1997). Repetitive DNA elements account for approximately 90% of the wheat genome, and transposable elements make up 80% of this (Wanjugi et al., 2009).

Wheat is the most widely adapted major crop and is grown on a larger land area than any other crop worldwide (Reynolds et al. 2011; Munns and Richards, 2007). Wheat is the third most important cereal crop next to only maize (*Zea mays* L.) and rice in annual production (Graybosch and Peterson, 2010). One-fifth of the total calories of the world's population comes from wheat (FAO, 2010), making wheat an important component of food security at the global level.

1.2 Drought and wheat

Drought in agriculture refers to water deficit in the root zone of plants and results in yield reduction during the crop life cycle (Rampino et al., 2006; Passioura, 2007; Nevo and Chen, 2010; Ji et al., 2010). Therefore, drought tolerance is defined as the ability of plants to survive and reproduce under water deficit conditions (Fleury et al., 2010). There are three components of

drought resistance viz. dehydration avoidance, dehydration tolerance and dehydration escape. Dehydration avoidance is the ability of the plant to maintain its hydration state whereas dehydration tolerance refers to a plants' ability to function after dehydration (Blum, 2011). Dehydration avoidance strategies in plants are a deep rooting system to access water, efficient use of available water and matching rainfall through life cycle modification (Salekdeh et al., 2009). In crop plant drought resistance, dehydration avoidance is a more common and effective mechanism than dehydration tolerance. The escape mechanism has been used in crop improvement efficiently through selection for a shortened crop cycle to develop early maturing varieties that escape terminal moisture stress. The disadvantage of the escape mechanism is that it is associated with a yield penalty under optimum growing conditions. Moreover, breeders for well-developed agricultural regions have already optimized crop flowering time to match the growing environments (Passioura, 2007).

Drought stress is usually unpredictable in its timing, duration and intensity. Plant response to drought stress is complex as it involves a number of physiobiochemical processes at the cellular level and different interacting component traits with different responses at the whole plant level (Witcombe et al., 2008; Kadam, 2012). Hence, drought tolerance is a complex trait with low heritability, quantitative in nature and having a high level of genotype by environment (GxE) interaction. Further, plant phenology and morphological traits such as plant height and tillering can confound plant responses to drought (Fleury et al. 2010). Drought is also commonly accompanied by heat stress and the simultaneous occurrence of these two abiotic stresses under field conditions can have significantly greater effects on crop productivity than individual stress effects (Salekdeh et al., 2009).

Plant breeding has improved crop performance under drought conditions in the past (Cattivelli et al., 2008). However, previous progress in genetic gain of yield is not enough to meet the higher demand for food products as a result of world population increase in the face of changing climate. Currently, there is a great interest to increase crop productivity under drought conditions through combining knowledge gained on physiological traits, drought tolerance genetic control and the target environments (Blum, 2011). The success of physiological trait-based breeding for drought tolerance depends on the genetic correlation of the trait with final yield, extent of genetic variability, level of heritability and extent of GxE interactions (Mir et al., 2012). With the availability of desired traits at hand, precise phenotyping in target drought environments is a key to accurately associate the massive genotypic data available today with phenotypic expression of a trait (Salekdeh et al., 2009).

Drought stress seriously limits wheat productivity around the world. Wheat is grown under a wide range of environmental conditions, but it is best adapted to temperate regions where rainfall is 30-90 cm (Hussain and Rivandi, 2007). Wheat is also the major cereal grown in dry regions of the temperate zone. Nearly 50% of the area sown to wheat is affected by drought on an annual basis (Trethowan and Reynolds, 2007) and it can cause up to 50% yield reduction in comparison to yield under full irrigation (Nezhadi et al., 2012). Winter wheat is commonly grown in the Great Plains following a fallow period, where soil moisture stored during the fallow period is used for winter wheat production (Dhuyvetter et al., 1996). Although the soil moisture stored during the fallow period is often sufficient for vegetative stage growth and development of wheat plants, post-anthesis drought stress often limits wheat productivity in the Great Plains (Mulat, 2004).

Wheat is one of the major cereal crops grown in Ethiopia, and it ranks fourth after teff (*Eragrostis tef*), maize and sorghum (*Sorghum bicolor*) in area coverage (Bayeh, 2010). Wheat is grown in Ethiopia mainly in humid or sub-humid agro-ecological zones, and the average national yield is typically below East African and world yield averages (Schneider and Anderson, 2010). Drought stress both at early growth stages and during the grain filling stage are among the factors contributing to the low productivity of wheat in Ethiopia.

Genetic studies conducted under water-limited environments have identified quantitative trait loci (QTL) underlying yield and yield component traits of wheat (El-Feki, 2010; McIntyre et al., 2010; Pinto et al., 2010; Kirigwi et al., 2007). Many chromosomal regions with minor effects have been involved in controlling yield, but repeatable QTL across environments and different backgrounds are rare, if indeed there are any. This situation has undermined the transferability of QTL information into practice in plant breeding programs to increase yield genetic gain under water-limited environments. Therefore, focusing on the identification and utilization of genomic regions for traits related to drought tolerance (e.g., root traits, reproductive traits) may be a more feasible strategy than yield per se approaches.

1.3 Molecular markers and QTL mapping in wheat

1.3.1 Molecular markers

Marker-assisted selection (MAS) may accelerate the variety development process in plant breeding. Several marker systems have been used for QTL mapping for different crop species. Both bi-allelic and multi-allelic co-dominant markers are suitable for estimating linkage disequilibrium (LD). Simple sequence repeats (SSRs) and restriction fragment length polymorphisms (RFLPs) are co-dominant markers that have been widely used for QTL mapping (Bryan et al., 1997; Landjeva et al., 2007). Among dominant markers, amplified fragment length

polymorphisms (AFLPs) and randomly amplified polymorphic DNAs (RAPDs) have been used successfully in QTL mapping despite their low statistical power in relation to co-dominant markers (Abdurakhmonov and Abdukarimov, 2008). More recently, however, Single Nucleotide Polymorphism (SNPs) and Diversity Array Technology (DArT) markers have been widely utilized for genome-wide scanning of QTL in many crop plants. The development of sequencing technologies has allowed the discovery of several fold greater numbers of SNPs than DArT markers in many crop species (Poland et al., 2012). These marker systems are inexpensive per data point and simultaneously assay several thousand loci in a single assay.

Diversity arrays technology is a hybridization-based alternative similar to a microarray platform to detect the presence versus absence of individual DNA fragments in genomic representations generated by complexity reduction methods from samples of genomic DNA (Jaccoud et al., 2001). The applicability of DArT for hexaploid wheat has been tested by Akbari et al. (2006) by comparing with SSR, RFLP and AFLP markers in terms of distribution along chromosomes, segregation distortion, level of polymorphism frequency and reproducibility of markers. Generally, the increase of ploidy level did not negatively affect the application of DArT markers for hexaploid wheat. The data quality for wheat was also similar to the quality of DArT data previously generated for barley (*Hordeum vulgare* L.) and other species. There was no significant difference in the distribution of the SSR markers and DArT markers among the seven homoeologous chromosome groups of wheat. However, there was a statistically significant deficit of DArT markers on the D genome and a greater tendency to map to gene-rich telomeric regions than SSR and AFLP markers (Akbari et al., 2006).

SNP markers are becoming the markers of choice in plant breeding programs for construction of high resolution genetic maps, genome wide association mapping, genomic

selection, and population evolutionary history studies (Aranzana et al., 2005; Zhao et al., 2007; Akhunov et al., 2009). SNPs are generally more abundant, stable, amenable to automation, efficient and cost-effective than other forms of genetic variants (Rafalski, 2002; Akhunov et al., 2009). SNPs can be individually responsible for phenotypic expression of a trait or linked to causative SNPs (Langridge and Fleury, 2011). However, selecting the most suitable set of SNPs which are either causative SNPs or linked to causative SNPs in a cost-effective manner is an important step toward application of molecular markers for crop improvement (McCouch et al., 2010).

1.3.2 Quantitative trait loci mapping (QTL) populations

In crop plants, the standard mapping populations are derived from crosses between two parents which have contrasting characters of a trait under investigation; for example, drought tolerant versus drought susceptible parents. These bi-parental cross populations have been used for determining the number, effect size and chromosomal locations of QTL underlying agriculturally important quantitative traits including grain yield of wheat. Some of the advantages of bi-parental populations include the requirement of relatively fewer markers for genome coverage, no population structure and ability to locate QTL regions along chromosomes (Sorrells and Yu, 2009). The disadvantages of bi-parental population mapping approach are:

1) Only two alleles can be evaluated at a locus.
2) Low mapping resolution due to few recombinations.
3) Longer time required to develop mapping population.

1.4 Association mapping

The classical method of QTL identification is conducted by a bi-parental QTL mapping approach. Association analysis which does not require development of a bi-parental mapping

population is becoming a common method of QTL mapping mainly due to its high resolution, broader allele coverage and cost effectiveness. In this method, diverse lines or cultivars can be used for obtaining information on marker-trait associations. It has the potential to identify QTL associated with a desired trait and even to detect the causal polymorphisms within a gene that are responsible for the difference in two alternative phenotypes (Gupta et al., 2005). The resolution of QTL is high as only closely linked alleles are in LD due to a long history of recombination (Ingvarsson and Street, 2011). Association mapping is also useful for establishing associations between haplotype blocks and traits of interest. However, genomic locations of QTL detected by the association mapping approach need to be inferred from a consensus genetic map and/or physical map for the crop under study. Special mapping populations known as Nested Association Mapping (NAM) populations allow simultaneous QTL detection and chromosomal position determination (Ersoz and Buckler, 2009). However, NAM populations are currently available only for a limited number of crop species like maize. The NAM population in maize was developed by crossing 25 diverse inbred lines to a common reference inbred B73 to produce 25 bi-parental recombinant inbred line families that have one parent in common (Cook et al., 2012).

The steps of association mapping analysis are: (1) selection of a group of individual lines or cultivars with wide genetic diversity to form the mapping population or panel; (2) recording the phenotypic characteristics; (3) genotyping the mapping population with available molecular markers; (4) quantification of the extent of LD for a chromosome and/or a genome using molecular marker data of the mapping panel; (5) assessment of the population structure and kinship (coefficient of relatedness between each pair of individuals); (6) determination of association of phenotypic and genotypic data based on the information gained from LD and

population structure using appropriate statistical methods (Abdurakhmonov and Abdukarimov, 2008).

Association mapping broadly falls into two major classes: (1) genome-wide association mapping, which surveys genetic variation in the whole genome using a large number of markers to detect regions associated with the phenotype (Zhu et al., 2008); and (2) candidate-gene association mapping, which relates within candidate gene polymorphisms with phenotypic variations of the traits. The choice between whole genome scanning and candidate gene approaches depends on the extent of LD in the population and the availability of markers. Although genome-wide association is a promising approach for scanning the entire genome for detecting marker-trait associations with a large number of markers, the candidate gene approach is also important to map targeted genes with known function (Tabor et al., 2002).

The association mapping approach has been used for several crops to identify QTL and also to characterize candidate genes. A review of studies involved with both genome-wide and candidate gene association mapping approaches is presented below.

1.4.1 Genome wide association mapping

1.4.1.1 Genome wide linkage disequilibrium (LD) in wheat

LD refers to a non-random association between alleles at two loci. It is a pair-wise measurement between polymorphic sites. The resolution and power of association studies in a collection of cultivars depend on the extent of LD which in turn depends on population history, recombination frequency, chromosome region, sample size, mating system and mutation across the whole genome (Ersoz et al., 2009; Zhang et al., 2009; Chao et al., 2010). LD decay is a function of genetic distance. It may decay over a long or short distance based on the species and population under consideration and the region of the chromosome.

Association mapping exploits historical recombination events because LD is the net result of all the recombination events that occurred since the origin of an allele by mutation. Only closely linked loci remain associated and co-segregate for many generations (Morton et al., 2001). This provides the opportunity to dissect quantitative traits with higher resolution mapping at the gene level (Ersoz et al., 2009); hence, causative genes with modest effects can be mapped with LD-based association approaches (Hirschorn and Daly, 2005).

Several LD statistics have been used to estimate the levels of LD and to make inferences about recombination rate and mutation history. Among those, r^2 and D' are the most commonly used statistics to measure LD (Gupta 2005; Sorrells and Yu, 2009). All LD statistics measure the difference between the observed and expected haplotype frequencies (Flint-Garcia et al., 2003). If a pair of loci with alleles "A" and "a" at the 1st locus X, and "B" and "b" at the 2nd locus Y are considered,

$D= P_{AB}-(P_A)(P_B)$, where D is LD between two loci, X and Y; P_{AB} is the frequency of gamete AB ; P_A and P_B are the frequencies of alleles "A" and "B" at locus X and Y, respectively. On the other hand, the LD statistic D' (Lewontin, 1988) is calculated as:

$|D'|= (D)^2/\min(P_AP_b, P_aP_B)$ for $D <0$, where P_a and P_b are the frequencies of allele "a" and "b", respectively.

$|D'|= (D)^2/\min(P_AP_B, P_aPb)$ for $D >0$

Similarly, Hill and Robertson, 1968 defined r^2 as:

$r^2= D^2/P_AP_aP_BP_b$

The statistic, r^2 can be defined as the squared value of the Pearson's correlation coefficient (product moment) of allelic frequencies at two loci. Although the performance of both statistics are affected by small sample size and low allele frequencies, r^2 is less sensitive to

sample size and better in indicating how markers might be correlated with QTL of interest (Flint-Garcia et al. 2003; Martinez et al., 2006). While D' is useful to estimate recombination differences accurately, r^2 summarizes both recombination and mutation history. Generally, the statistic r^2 is more favored in assessing the extent and patterns of LD than D' statistics. The value of r^2 approaches one when the frequency of co-segregation of alleles at two loci is high while an r^2 value of zero shows the co-occurrence of alleles at two loci does not differ from what would be expected under random sampling (Ersoz et al., 2009). To summarize the structure and patterns of LD, r^2 for pairwise combinations of alleles are plotted against the genetic distances among alleles on a chromosome. This type of graphical display is known as a LD decay plot which allows fitting decay curve to estimate LD decay for a chromosome or for an entire genome (Gupta et al., 2005; Abdurakhmonov and Abdukarimov, 2008).

Several genome-wide association mapping studies have been reported for many crops. Most of those studies mainly focused on the determination of LD, generating information on how far the usable levels of disequilibrium extend in the genome, and how much LD pattern is affected by mating system, recombination rate, population structure, population history, genetic drift and directional selection. Different patterns of LD have been reported for crop plants such as rice (Agrama et al., 2007), maize (Wilson et al., 2004), barley (Comadran et al., 2009) and wheat (Chao et al., 2007). Broadly, the extent of LD decay over genetic distance occurs at a slower rate in self-pollinated crops such as Arabidopsis, rice, wheat, barley and sorghum than cross-pollinated crops (e.g., maize) as the number of effective recombinations is lower in self-pollinated crops compared to cross-pollinated crops.

The strength and patterns of LD in wheat vary among chromosomes and genomes. Analysis of LD for 43 U.S wheat cultivars has shown the intra-chromosome LD decay below $r^2<$

0.2 within 10 cM (Chao et al., 2007). On the contrary, significant long range LD (over 30 cM genetic distance) has been recorded for chromosomes 3DL, 4DL and 6AL. At the genome level, the B genome showed the highest proportion of significant LD despite fewer markers. In another study conducted on 96 soft winter wheats with SSR markers, LD decayed rapidly within 1 cM for chromosome 2D but extended up to 5 cM for chromosome 5A (Breseghello and Sorrells, 2006). Similarly, Yao et al. (2009) reported that LD decayed on average within 1 cM for chromosome 2D, within 0.5 cM for chromosome 3B, but extended up to 2.3 cM on chromosome 2A of hexaploid wheat implying the presence of large differences among wheat chromosomes in rate of LD decay.

The most comprehensive analysis of LD patterns has been conducted on a total of 478 spring and winter wheats genotyped with 394 SNP markers. This study revealed that LD declined to 50% of its initial value within 6-7 cM for the A, B and D genomes (Chao et al. 2010). Genome-wide LD estimation for 251 winter wheat lines with 346 DArT makers also showed on average LD declined below $r^2<0.2$ at 9.9 cM (Benson et al., 2012). Liu et al. (2010) genotyped 103 wheat accessions from China with 116 SSR markers on chromosome 4A and found extension of LD up to 3 cM with threshold level at $r^2= 0.054$. The study conducted on elite durum wheat genotypes also showed the dependence of LD on different factors. For elite durum wheat(*Triticum durum* Desf.) lines genotyped with SSR markers, LD extended up to 10 cM to reach a critical threshold of $r^2=0.06$ (Maccaferri et al., 2011). Another study on durum wheat genotyped with 58 SSR markers showed the decay of LD within 10 cM (Maccaferri et al., 2005). When both bread and durum wheats are considered together, there was no difference in LD patterns between the two. While LD in durum wheat marginally extended over larger distance, generally LD decayed within 2-3 cM for both wheat types (Somers et al., 2007). Since studies

are different in r^2 threshold levels, population sample size and marker type, it is difficult to draw an overall conclusion regarding LD extent and patterns in wheat.

In maize, LD decays in 1 kb for landraces, 2 kb for inbred lines and extends up to 100-500 kb for commercial elite inbred lines (Remington et al., 2001; Ching et al., 2002; Jung et al., 2004). However, LD extended up to 10 kb for s*hrunken* (*sh1*), an enzyme in the starch biosynthesis pathway, possibly due to its being under direct selection during domestication or breeding (Whitt et al., 2002). In rice, LD extended up to 100 kb to over 200 kb for cultivated rice (Huang et al., 2010; Mather et al., 2007) while barley had extensive LD up to 20 to 30 cM (Hamblin et al., 2010). Recently, Xu et al. (2012) determined the extent of LD for 188 tomato (*Solanum lycopersicum*) accessions with 192 SNP markers and found LD extended up to 18 cM at r^2=0.3 on average for all chromosomes. Studies on Arabidopsis indicated that LD extended 50-100 cM even if it breaks down within 10-50 kb for some genes (Tian et al., 2002). Comparison of the extent of LD across cereals showed that LD for wheat extends over a longer distance than maize and rice but decays faster than LD for barley. Within a species LD decay rate differs depending on population type and chromosome regions. Therefore, LD analysis should be done at the chromosome level for each association mapping population.

1.4.2 Population structure

The association mapping approach has been seen with skepticism by plant genetics and breeding communities until recently because of spurious associations as a consequence of the confounding effect from population structure. Population structure often leads to a genome-wide LD between unlinked loci (Sneller et al., 2009). Structured populations may show significantly different allele frequencies due to genetic drift, domestication or background selection;

consequently, genetic loci could be spuriously associated with a trait when there is no real association.

The development of a statistical model which allows accounting for population structure during association analysis has improved the application of association mapping for QTL detection in crop plants. There are two steps to account for population structure using a model-based approach; the first is to calculate the percentage of membership of each individual to population groups using unlinked random markers, and the second is to use the percentage of membership as a covariate in the model of testing associations of markers with phenotypic traits (Ersoz et al., 2009). In the unified mixed model of Yu et al. (2006), both population structure (Q) and family relatedness (K) are simultaneously considered as covariates in the model. This model accommodates both fixed and random effects.

The Q+K mixed model is represented with the following equation:

$$y = X\beta + S\alpha + Qv + Zu + e$$

where y is a vector of phenotypic observations; β is a vector of fixed effects other than marker or population structure; α is a vector of marker effects; u is a vector of random polygenic background effects; e is a vector of residuals; Q is a matrix from structure relating v to y; and X, S and Z are incidence matrices of 1s and 0s relating y to β, α and u, respectively. The variances of the random effects are assumed to be $Var(u) = 2KVg$, and $Var(e) = RV_R$ (Yu et al., 2006), where K is an n × n matrix of relative kinship coefficients that define the degree of genetic covariance between a pair of individuals; R is an n × n matrix with the off-diagonal elements being zero and the diagonal elements being the reciprocal of the number of observations for which each phenotypic data point was obtained; Vg is the genetic variance; and V_R is the

residual variance. Best Linear Unbiased Estimates (BLUEs) of β, α and v (fixed effects) and Best Linear Unbiased Predictions (BLUPs) are obtained by solving mixed model equations.

Different levels of population structure have been detected in wheat, from none to highly structured populations. Unlike rice and maize, there are no well-known structure or heterotic groups for bread wheat (Coviour et al., 2011). From population structure analysis on 96 diverse Great Plains winter wheat cultivars and advanced lines developed for genetic study of quality traits, eight subpopulations have been detected with 60 SSR loci (Zheng et al., 2009). Another study conducted on 376 bread wheat collections from Europe and East Asia using 70 SSR loci indicated the presence of only two subgroups in the population where the lines were assigned to their known gene pools (Hao et al., 2010).

1.4.3 Candidate gene association mapping

Candidate gene association studies are aimed at linking phenotypic variation with allelic variation in candidate genes and benefit from several generations of recombination in natural populations to identify causative polymorphisms (Gonzalez-Martinez et al., 2008). In plants with large genomes, the generation of molecular-linkage maps based on candidate genes (molecular-function maps) is one way to identify functional markers instead of time-consuming fine mapping.

1.4.3.1 Drought tolerance candidate genes

A large number of drought inducible genes have been identified and characterized for their function (Shinozaki and Yamaguchi-Shinozaki, 2007). There are two categories of genes based on their response to the phyto-hormone abscisic acid (ABA): ABA independent and ABA dependent. Dehydration-Responsive Element Binding (DREB) genes are ABA independent and known for their association with abiotic stress tolerance. Currently, full-length sequences of

DREB1 and *DREB2* genes have been cloned from *Triticum* a*estivum*, *Oryza sativa*, *Zea mays* and *Arabidopsis thaliana* (Wei et al., 2009). Transgenic wheat with a *DREB1A* gene from *Arabidopsis* showed more drought tolerance, more branches and better spike size than non-transgenic wheat plants (Pellegrineschi et al., 2004). However, in a recent field evaluation the transgenic *DREB1A*-wheat lines did not have a grain yield advantage over control lines under water deficit conditions (Saint Pierre et al., 2012), despite their better recovery after severe water stress and higher water use efficiency in the greenhouse. It has also been observed that the *DREB2* gene from wheat improved freezing and osmotic stress in transgenic tobacco plants (Kobayashi et al., 2008).

Fructan 1-exohydrolase (*1-FEH*) is another ABA independent gene that is implicated in cold and drought tolerance through membrane stabilization and remobilization of water soluble carbohydrates from stem to developing grain (Lothier et al., 2007; Hincha et al., 2003). The three copies of the *1-FEH* gene have been mapped to the short arms of group 6 chromosomes, i.e., 6AS, 6BS and 6DS (Zhang et al., 2008).

ABA hormone concentration rises rapidly in plant tissues in response to drought or soil water deficit, and this in turn leads to expression of ABA dependent stress-related genes (Shinozaki and Yamaguchi-Shinozaki, 2007; Wan et al., 2009). The *ERA1* (Enhanced Response to ABA) gene which has been cloned from *Arabidopsis* and hexaploid wheat is ABA dependent in its expression. It has been shown that *ERA1* mutants increased drought tolerance of Arabidopsis through stimulating stomatal closure (Ziegelhoffer et al., 2000).

1.4.3.2 Functional markers in candidate genes

A functional marker refers to a marker developed from SNPs or insertion/deletion sites within a gene (Andersen and Lubberstedt, 2003). Functional markers in molecular plant breeding

are more advantageous than linked markers because the latter may not be diagnostic due to segregation between the marker and putative causative SNPs in subsequent generations. Since functional markers are developed from SNPs within a gene, marker information can be used confidently across breeding programs to select favorable alleles for a trait of interest (Bagge and Lubberstedt, 2008). Several genes for agronomic traits (e.g., semi-dwarfism genes), quality traits (e.g., polyphenol oxidase) and drought tolerance (e.g., DREB genes) have been identified for wheat (Wei et al., 2009; Bagge and Lubberstedt, 2008), but functional markers have been developed only for a few of them. Therefore, more functional markers are needed from the genes to enhance the application of molecular markers in crop improvement as the cost of re-sequencing the genes is dramatically decreasing.

SNPs may be discovered with different methods. However, the most straightforward approach is the direct re-sequencing of amplicons of genes from different genotypes (Rafalski, 2002). Amplification of DNA segments with genome-specific primers for polyploids like hexaploid wheat is challenging due to sequence similarity among gene families. This to some extent slows down the application of functional markers in wheat breeding.

Generally, once genes that determine the genetic basis of a trait are known, developing functional markers to select for favorable alleles is an important aspect of using genetic information in practical plant breeding (Langridge and Fleury, 2011). However, for successful functional marker development, prior information about the level of DNA polymorphisms, extent of linkage disequilibrium and within gene nucleotide diversity is required.

1.4.3.3 SNP-trait associations within candidate genes

The candidate gene strategy has shown promise for bridging the gap between quantitative genetic and molecular genetic approaches to study complex traits (Cattivelli et al., 2008;

Ingvarsson and Street, 2011). Along this line, studies involved with the candidate gene approach are summarized for wheat and other crops as follows.

Vernalization requirement in wheat is controlled by four major genes, viz. *VRN1*, *VRN2*, *VRN3* and *VRN4*, with *VRN1* gene copies *VRN-A1*, *VRN-B1* and *VRN-D1* located on the long arms of chromosomes 5A, 5B and 5D, respectively (Yoshida et al., 2010). An association mapping study conducted by Rousset et al. (2011) on 235 hexaploid wheat collections revealed the effects of the flowering time candidate genes in modulating flowering time in wheat. In that study, genetic variation in *VRN-A1*, *VRN-B1* and *VRN-D1* genes has explained a large part of phenotypic variation in growth habit.

Huang and Brule-Babel (2012) studied genetic diversity, haplotype structure and association of genes involved in starch biosynthesis in wheat. Genes encoding granule-bound starch synthase (GBSSI, also known as waxy or *Wx* genes) and soluble starch synthase (SSIIa) were selected for nucleotide diversity and SNP density study. None of the SNPs within the three *SSIIa* genes and *Wx-D1* gene was associated with yield-related traits. However, both SNPs and haplotypes within the *Wx-A1* gene were associated with seed number per spike, seed weight per spike and thousand kernel weight. Another study on grain size of wheat also demonstrated the association of haplotype of a grain size gene (*TAGW2*) with larger grain size, earlier heading date and maturity in hexaploid wheat (Su et al., 2011).

Candidate gene association analysis has been used for cereal crops other than wheat. Transcription factors such as the gibberellin-regulated Myb factor (GAMYB), the barley leucine zippers 1 and 2 (*BLZ1*, *BLZ2*), and the barley prolamin box binding factor (BPBF) were evaluated for their association with agronomic traits in barley. SNPs within *BLZ1* were associated with days to flowering, and its haplotype was also associated with both days to

flowering and plant height. The haplotype of *BLZ2* was associated with thousand kernel weight while the haplotype of the *BPBF* gene was associated with both crude protein and starch in barley endosperm (Haseneyer et al., 2010). However, the candidate genes explained only a small portion of the total genetic variation. Similarly for maize, sorghum and rice, candidate genes involved in starch biosynthesis were associated with the expected traits and the results were in agreement with QTL studies (Wilson et al., 2004; Bao et al., 2006; Figueiredo et al., 2010).

The most comprehensive candidate gene association results have been recently reported for SNPs identified from 540 genes putatively involved in accumulation of carbohydrate and ABA metabolites during stress for maize (Setter et al., 2011). In that study, the SNP from a homologue of an *Arabidopsis* MADS-box gene was significantly associated with phaseic acid in ears of irrigated plants while a SNP in pyruvate dehydrogenase kinase was significantly associated with silk sugar concentrations. Similarly, a SNP from an aldehyde oxidase gene was associated with ABA levels in silk under non-irrigated conditions.

The candidate gene association mapping approach has been widely applied in forest tree genetics studies as developing a bi-parental population is practically unfeasible for most conifers. Gonzalez-Martinez et al. (2006) studied the pattern of polymorphisms of 18 drought responsive candidate genes in 32 *Pinus taeda* L. individuals. LD within the sequenced gene regions varied from low to high depending on the candidate gene locus. Thirteen genes had r^2 greater than 0.1, but they did not find tight LD among sites within the gene or sites of genes located on the same chromosomes. A total of 196 SNPS and 82 LD blocks were obtained in 18 candidate gene loci. By constructing LD blocks, 94 haplotype SNPS were identified to improve the LD values and were successfully used in detecting significant r^2 values for LD blocks study. The same authors evaluated the association of four candidate genes belonging to different functional classes with

carbon isotope discrimination (CID) at two locations. The genes were general protection factor (*dhn-1*), anti-oxidants (*sod-chl*), transcription factor (*wrky-like*) and putative cell wall protein (*lp5-like*). Anti-oxidant (*sod-ch1*) and *Cu/Zn superoxide dismutase* genes showed significant association with CID at both locations. However, none of the significant associations explained a substantial amount of phenotypic variance in CID.

1.5 Yield and yield component traits, and their genetic control

1.5.1 Grain yield

Grain yield improvement is the ultimate goal for most wheat breeding programs across the world. Although grain yield is a complex trait with low heritability and highly influenced by genotype x environment interaction, high yielding commercial varieties of many crops including wheat have been developed through direct selection for grain yield even if the relationship of yield with its component traits has already been established. The major grain yield determining traits of wheat are kernel number per unit of land area, harvest index and kernel weight.

Understating the genetic basis of yield and yield component traits is critical for crop improvement. Several studies have been reported on the genetic control of yield and its component traits. Major findings related to the genetic basis of hexaploid wheat yield and yield components are summarized in the following section.

Previous studies have shown that all 21 wheat chromosomes have been involved in controlling grain yield in wheat. Cuthbert et al. (2008) evaluated 402 doubled haploid (DH) lines derived from two spring wheat parents with contrasting yielding ability at six locations for two years in Canada. Five major QTL on chromosomes, 1A, 2D, 3B, and 5A were detected for grain yield. Out of these, a QTL on chromosome 5AL was the most significant and explained 17.4 % of the phenotypic variation in grain yield. This QTL was also detected for heading date, harvest

index, kernel number spike^{-1} and kernel weight spike^{-1}. In that study QTL detected for yield were largely consistent across environments and overlapped with QTL of at least one yield component. Among yield components, kernel weight spike^{-1} and kernel number spike^{-1} had more QTL in common with yield whereas number of spikes m^{-2} was the least coincident yield component. Huang et al. (2003) genotyped 72 lines from advanced backcross population using 210 SSR markers to identify QTL for yield and some yield component traits. They found yield QTL on chromosomes 1AL, 1BL, 3AS, 2BL, 2DL, 3BS, 4DS and 5BS.

Kumar et al. (2007) found a QTL for five traits (grain yield, harvest index, spike length, spikelet per spike and kernel number per spike) on chromosome 2DS, and another multi-trait QTL for three traits (biological yield, harvest index and spikelet per spike) on chromosome 4AL. Marza et al. (2006) found 10 yield QTL on chromosomes 1AL, 1B, 2BL, 4AL, 4B, 5A, 5B, 6B, 7A and 7D. Out of these, the QTL on 5A explained the largest grain yield variation (18.5%). El-Feki (2010) reported the most stable yield QTL on chromosome 5A from a study conducted under contrasting moisture levels in Colorado.

The significant phenotypic correlations and coincidence of QTL for grain yield and yield components have been implicated in some QTL studies (Kuchel et al., 2007b; Kumar et al., 2007). For instance, the pattern of correlations in the Cuthbert et al. (2008) study was consistent with the number of QTL shared between yield and its component traits. Positive and significant phenotypic correlation was observed for yield with thousand kernel weight, kernel weight spike^{-1}, harvest index and kernel number spike^{-1}, whereas its phenotypic correlations with number of spike m^{-2}, grain filling time, heading and maturity date were low and negative. However, Huang et al. (2003) found phenotypic correlations for yield with thousand kernel weight, plant height, ear emergence and tiller number m^{-2} to be low and inconsistent across locations. Besides QTL

results and phenotypic correlations, the Cuthbert et al. (2008) study reported the highest heritability for number of spikes m^{-2} (0.98), and the lowest heritability for yield (0.48) and days to maturity (0.48). Heritability estimates were higher for yield components such as thousand kernel weight (0.77), kernel weight $spike^{-1}$ (0.97) and kernel number $spike^{-1}$ (0.58) than for phenological traits such as grain filling duration (0.52), heading date (0.49) and days to maturity (0.48).

McIntyre et al. (2010) also found high heritability estimates (>0.70) for days to anthesis, plant height, hectoliter weight and grain weight; moderate heritability estimates (0.40-0.70) for grain per spike, grain yield, harvest index, grain number m^{-2} and spike number m^{-2}; and low heritability estimates (<0.40) for biomass at anthesis and maturity.

Kirigwi et al. (2007) detected major QTL on chromosome 4AL for grain yield, biomass, spike density, kernel number m^{-2}, grain fill rate, biomass production rate and drought susceptibility index. Li et al. (2007) evaluated 131 recombinant inbred lines (RIL) of wheat in four environments and detected five QTL for grain yield on chromosomes 2A, 2D, 3B and 6A in three environments. They also identified stable QTL for spike number on chromosome 7D which explained up to 52% of phenotypic variation, and on chromosome 1D for thousand kernel weight and spike number. Putative yield QTL have been reported also for grain yield on chromosomes 6AS, 6AL and 7AS based on 194 recombinant inbred lines evaluated in nine Australian environments (McIntyre et al., 2009).

Huang et al. (2006) reported the presence of three yield QTL on chromosomes 5A, 7A and 7B which explained from 8 to 11% of the phenotypic variation by evaluating DH lines at three locations for a total of 6 environments in Canada. McCartney et al. (2005) detected the

most significant yield QTL on chromosomes 2B and 4A from QTL analysis conducted on 185 DH lines evaluated at a total of eight site-years in Manitoba, Canada.

The genome-wide association mapping approach has been applied recently for QTL detection in wheat. Neumann et al. (2011) studied a winter wheat association mapping panel which consisted of 96 diverse lines obtained from a larger collection from 21 countries. The entries were investigated for up to eight seasons for 20 morphological and agronomic traits with 835 DArT markers. Of all morphological and agronomic traits studied, the highest number of marker-trait associations (MTAs) was recorded for number of spikelets per spike (38), whereas the lowest number of MTA was obtained for thousand kernel weight and harvest index. Similarly, the highest number of trait-specific MTA was obtained for biomass (13) followed by grain number per spike and spike length (each 12). Four grain yield-specific MTA were detected on chromosomes 3A, 3B, 4B and 5B, and another six multi-trait markers on chromosomes 1A, 3A, 4A, 6B, 7A and 7B were also associated with grain yield.

Crossa et al. (2007) conducted association analysis for yield and disease resistance using 170 spring wheat lines which were genotyped with DArT markers. They found MTA for yield on all chromosomes with the exception of chromosome 4D, indicating the power of association mapping to detect many QTL in a single population, which otherwise would be achieved only with many independent bi-parental populations.

1.5.2 Thousand kernel weight and kernel weight per spike

Thousand-kernel weight is one of the three main yield components of wheat. It has a high and consistent heritability value. Thousand-kernel weight is also phenotypically the most stable yield component (Sun et al., 2009), and the effects of most genes affecting thousand kernel weight are

additive. Hence, early generation selection for thousand-kernel weight is most likely effective (Wang et al., 2012).

Kernel weight is a function of kernel length and kernel width. The critical period of kernel weight determination starts shortly before anthesis and continues throughout the period after anthesis during grain-filling duration in which the final grain size is determined in wheat (Sinclair and Jamieson, 2006; Ji et al., 2010). Unfavorable environmental factors (e.g., high temperature and water deficit) during grain-filling duration reduce kernel weight significantly.

Kernel weight and kernel number are at least partially controlled genetically by different loci. This is mainly because environmental factors (e.g., drought stress) affect these traits in different reproductive structures and at different developmental stages (Ji et al., 2010). Kernel number is mainly determined at pre-anthesis stages whereas kernel weight is determined during the grain-filling stage, even if there is some overlap of critical periods for kernel weight and kernel number. The existence of flexibility in compensation effect between kernel number and kernel weight of wheat also hinders improvement of yield potential through simultaneously increasing both kernel number and kernel weight (Sinclair and Jamieson, 2008).

Kernel traits of wheat are generally quantitative in nature, affected by many QTL and GXE interaction (Sun et al., 2009). McCartney et al. (2005) detected two major QTL for thousand kernel weight on chromosomes 4BS and 4DS in the region of *Rht-B1b* and *Rht-D1b* with QTL on 4DS explaining 31.8% of the phenotypic variation. For both regions, the reduced plant height was correlated with reduced thousand kernel weight for the test environments. Other minor QTL were also detected on chromosomes 2A, 3D, 4A and 6D for thousand kernel weight.

Nezhad et al. (2012) evaluated 133 $F_{2:3}$ families of bread wheat under stress and fully-irrigated conditions both in the field and greenhouse for detecting QTL under post-anthesis drought stress for thousand kernel weight. They found QTL on chromosomes 7AS and 7DS which were consistently detected for both moisture stress treatments, both under the field and greenhouse conditions. From a study conducted on 402 spring wheat DH lines, Cuthbert et al. (2008) detected six QTL for thousand kernel weight on chromosome 2D, 3B, 5A and 7A, with the QTL on 5AS explaining about 11% of phenotypic variation. Similarly, seven QTL have been detected for average kernel weight spike^{-1} on chromosomes 1A, 3B, 5A, 5B, 5D and 7B with the QTL on 5AL explaining 20.9% of the phenotypic variation.

Wang et al. (2009) reported 21 QTL controlling thousand kernel weight on chromosomes 1B, 2A, 2D, 3B, 4A, 4D, 5A, 6D and 7D from 142 recombinant inbred (RIL) lines of winter wheat evaluated across four environments. Furthermore, thousand kernel weight was positively and significantly correlated with kernel weight spike^{-1}, kernel number spike^{-1}, days to maturity and grain filling duration. They also identified 10 QTL for kernel weight spike^{-1} on 1A, 2A, 3B, 4B, 4D, and 6B explaining 5.93% to 24.06%, but none of these QTL were expressed across test environments.

Wang et al. (2012) evaluated 262 wheat accessions in China in five environments and genotyped them with 531 SSR markers to detect QTL for thousand-kernel weight using the association mapping approach. The detected QTL were distributed on homoeologous groups 1, 2, 3, 5 and 7. Liu et al. (2010) detected marker-trait associations on chromosome 4A (9.9 and 70.6 cM) for thousand kernel for 103 Chinese wheat accessions with 116 SSR markers mapped on chromosome 4A. Huang et al. (2003) found QTL for thousand kernel weight on chromosomes 2DL, 4DS, 5BS, 7AS and 7B. However, in another independent experiment on 185 DH lines

evaluated in a total of six Canadian environments, Huang et al. (2006) detected thousand kernel weight QTL on chromosomes 2B, 2D, 4B, 4D and 6A, with QTL on 4D explaining 26.3 % of the phenotypic variation. Marza et al. (2006) reported QTL for kernel weight per spike on chromosomes 1B, 2BL, 2DL, 3BL, 3BS, 5A and 6B from 132 F_2–derived recombinant inbred lines. El-Feki (2010) studied 185 DH winter lines in four Colorado environments and detected kernel weight QTL on chromosomes 1A, 1B, 2B, 2D, 3B, 6A and 7D.

1.5.3 Kernel number

Kernel number is the primary determinant of yield increase in wheat. Genetic gains in wheat have been achieved due to improvement in kernel number with little or no change in individual grain weight (Gaju et al., 2009). The critical period of final kernel number determination is from the onset of stem elongation to anthesis and occurs throughout spike development. More specifically, this critical period spans 20 days before anthesis and 10 days after anthesis (Ugarte et al., 2007). Both high temperature and water deficit in this period may result in significant reduction of final kernel number and yield. Kernel number is the most susceptible yield component to abiotic stress in grain crops, accounting for greater yield loss than reduction in kernel weight (Dolferus et al., 2011). One of the direct effects of drought stress on wheat is the abortion of pollen development which leads to fewer kernels (Ji et al., 2010). The amount of nitrogen and carbon accumulated in the crop at anthesis also limits the final number of kernels and consequently grain yield (Sinclair and Jamieson, 2006). Drought stress increases the number of sterile tillers and only about half of the formed tillers of wheat survive to produce grains in semi-arid environments (Duggan et al., 2005).

Knowledge of the genetic basis of kernel number is important for wheat improvement as kernel number is the primary component of grain yield. Three putative QTL have been detected

on chromosomes 1B, 6A and 7A in 194 lines of a bi-parental spring wheat population evaluated at three locations from 2002 to 2006 in Australia (McIntyre et al., 2010). Pinto et al. (2010) identified QTL for kernel number on chromosomes 1B, 3B, 4A, 5B and 6B which explained from 4.4-12.5% of the phenotypic variation. With association analysis, kernel number QTL were detected on chromosomes 4A and 6B, with the former showing consistency across test environments (Neumann et al. 2011). Dodig et al. (2012) also detected a QTL on chromosome 2AS both under irrigated and dry conditions for kernel number using an association mapping panel of 96 diverse lines. Marza et al. (2006) detected QTL for kernel number per spike on chromosomes 1AL, 1B, 2BS, 2DL, 3BS, 4B, 6A and 7BS from an experiment conducted on 132 recombinant inbred lines evaluated at three locations for three seasons at Oklahoma. However, they found only one QTL for spike number on chromosome 3BS.

1.5.4 Harvest index (HI)

Harvest index indicates the efficiency of a crop in converting photosynthetic products or assimilates produced before and after anthesis into final grain yield. Most often it is expressed as the ratio of grain yield to above-ground dry matter. Although harvest index was not used as a selection criterion in wheat yield improvement in the past (e.g., during the Green Revolution), the achieved yield progress was actually due to an increase in the number of kernels and a genetic shift towards greater harvest index (Blum, 2005; Zhang et al., 2012).

The response of harvest index to environmental constraints (e.g., water deficits) depends on the intensity of the stresses. Harvest index, in the absence of stresses or with mild stresses, is fairly constant for several crops (Hay, 1995). However, progressive stresses which are sufficient to reduce biomass production by 30-40% can reduce harvest index, and the reduced biomass indicates the intensity of stress a crop has experienced (Fereres and Soriano, 2007). Cotton

(*Gossypium hirsutum*) and sorghum are the only two crops for which harvest index increases under moderate stresses (Fereres and Gonzalez-Dugo, 2009). Harvest index in wheat, however, is determined by the pattern of water use of the crop in the period before and after anthesis (Passioura, 1977).

The harvest index improvement in wheat has been mostly due to introduction of dwarfing gene alleles, *Rht-D1b* and *Rht-B1b*, into the background of modern cultivars. These genes reduced overall plant height and improved availability of assimilates which increased survival of growing florets to increase potential kernel number (Rebetzke et al., 2012). The harvest index of spring wheat is lower than that of winter wheat, and it rarely exceeds 45% for the former (Zhang et al., 2012). In spring wheat and winter wheat, harvest indexes of 50 and 55%, respectively, have already been realized in modern cultivars despite an estimated theoretical upper limit of 62-64% (Shearman et al., 2005). Generally, for spring wheat there is a potential of further yield improvement by increasing harvest index, as current values in breeding programs are in the range of 45 to 55% (Gaju et al., 2009).

Apart from understanding the physiological basis of the harvest index, knowledge of its QTL/genes is crucial for indirect selection for yield in wheat breeding. In the association study conducted by Neumann et al. (2011), trait-specific MTA have been detected for HI on chromosomes 1A, 3A, 7A and 7B, and multi-trait MTA have been identified on chromosomes 4A and 5A. In another association analysis with 96 diverse winter wheat lines, repeatable marker-trait associations have been detected on chromosomes 1DL and 2DS (Dodig et al., 2012). Cuthbert et al. (2008) also reported five QTL for harvest index on chromosomes 1A, 3A, 3B, 5A and 5B, and these QTL explained 4.2-11.9% of the phenotypic variation. El-Feki (2010) reported a total of eight harvest index QTL on chromosomes 1A, 1B, 2B (2), 2D (2), 3A and 6B.

1.5.5 Spike characters: spikelet number, spike length, kernel number per spike and spike number

Spikelet number affects the total number of kernels per unit area. The more spikelets per spike, the more kernels per spike, which may influence the final kernel number per land area. Neumann et al. (2011) identified trait-specific marker-trait association on chromosome 5B for spikelet number. Multi-trait markers on chromosomes 2B, 2D, 3A, 4A, 6B and 7B were also associated with spikelet number. Yao et al. (2009) detected four different QTL on chromosome 4A for spikelet number per spike using SSR markers.

Mao et al. (2007) reported a QTL on chromosome 7DS which controls both spike length and spikelet number per spike. Chromosomes 2DL and 5A also harbored QTL for spikelet number per spike. Liu et al. (2010) detected marker-trait associations for spikelet number and spike length on chromosome 4AL by conducting association analysis with 116 SSR markers mapped on chromosome 4A for 103 Chinese spring wheats. Chromosome 4DL is also involved in controlling spikelet number (Chu et al., 2008).

Long spikes with high spikelet number per spike may offer an avenue for increasing kernel number and harvest index in wheat (Gaju et al., 2009). Spike modification for increasing spikelets and kernel number per spike through breeding requires an understanding of the genetic bases underlying these traits. Many chromosome regions that affect spike length have been reported for wheat. Multi-trait marker-trait associations have been identified for spike length by Neumann et al. (2011) on chromosomes 2B, 2D, 3A, 3B, 5B, 6B and 7A, but spike length specific MTA were also located on chromosomes 3A, 4A, 5B and 7B (2). One of the MTA on chromosome 7B was significantly associated with spike length in all study years. Marza et al. (2006) also reported 10 QTL located on chromosomes 1AL, 1AS, 1B, 2BL, 2BS, 3BL, 4B, 5B,

7AS and 7BS for spike length. The QTL on chromosome 3BL was consistently detected in all test environments. Seven spike length QTL were detected by El-Feki (2010) and two QTL on chromosomes 1A and 1D were detected in all four test environments.

Yao et al. (2009) found marker-trait associations for spike length both on short and long arms of chromosome 2A, and most of the associated markers were located near QTL for multiple traits such as number of spikelets per spike and grain per spike. Ma et al. (2007) studied 136 recombinant inbred lines and detected major QTL for spike length on chromosome 7D and minor QTL on chromosomes 1A, 2D, 4A, 5A and 5B. Liu et al. (2010) detected four marker-trait associations for spike length on chromosome 4A. Dodig et al. (2012) found strong marker-trait associations for spike length on chromosomes 2DS and 6DS. However, Chu et al. (2008) reported QTL for spike length on chromosomes 3D, 4A and 5A.

Yao et al. (2009) found marker-trait associations using SSR markers on chromosome 2A on both arms for grain per spike. Cuthbert et al. (2008) also reported five QTL for kernel number spike^{-1} on chromosome 1A, 2D, 3B, 5A and 7A, and higher phenotypic variation has been explained (16%) by QTL on the long arm of chromosome 5A. Wang et al. (2009) found eight QTL which were mapped on chromosomes 1D, 3A, 4D, and 6A for kernel number spike^{-1}. Liu et al. (2010) found six marker-trait associations on chromosome 4A. McIntyre et al. (2010) detected three putative QTL which explained 5-8% of the variation on chromosomes 1D, 4D and 6B for high kernel number per spike. All three QTL were co-located with QTL for high harvest index, and two of them were also co-located with QTL for high kernel weight.

Spike number is strongly related with kernel number per unit area, the main yield component of wheat. In the study conducted by Neumann et al. (2011), five multi-trait MTA

were identified on 1A, 1D, 4A, 5B and 7A. However, spike number-specific MTA were found on chromosomes 2A, 2B and 7B with significant MTA on 2B in all years. Cuthbert et al. (2008) also found five QTL on chromosomes 3B, 5A (2), 5B and 7D with the QTL on chromosome 3B explaining about 10% of the phenotypic variation in spike number m^{-2}. Huang et al. (2003) found QTL on chromosomes 1BL, 2AL, 2DL, 3BS, 4DS, 5DL, 6DL and 7AS for tiller number per m^2.

1.5.6 Above ground dry biomass

Wheat yield genetic gain has been achieved mainly through increasing harvest index. Dry matter accumulation is the focus of future wheat yield increase as the optimum harvest index has already been achieved in modern winter wheat cultivars. Since dry biomass is a quantitatively inherited trait, it is important to understand the genetic bases for biomass production in wheat. Several studies have been conducted for biomass QTL/genes identification. A total of 12 trait-specific MTA were detected for above ground biomass by Neumann et al. (2011). Those associated markers were located on chromosomes 1D, 3B, 4B, 5B, 6A, 6B and 7B. Kirigwi et al. (2007) also reported QTL for biomass and biomass production rate on chromosome 4AL. Kadam et al. (2012) detected three QTL for shoot biomass on chromosome 4B despite inconsistency of positive allele contributions across environments. Eight QTL with intermediate effect (explained 5.6 to 8.2% phenotypic variation) were detected under different moisture levels in Colorado environments, and only one QTL detected on chromosome 2D showed consistency across environments (El-Feki, 2010).

1.5.7 Single kernel characters and test weight

The single kernel characterization system (SKCS) (Perten Instruments, Springfield, IL) is an instrument designed to measure traits such as single kernel weight, single kernel diameter and single kernel hardness by crushing the kernels (Osborne and Anderssen, 2003). Single kernel

characters such as single kernel weight and single kernel diameter affect grain yield. Grain hardness and test weight are also important quality traits in wheat. Hence, understanding the genetic basis of these traits is important to improve both grain yield and end use quality of wheat. Campbell et al. (1999) evaluated 78 RILs of wheat across six environments and detected QTL for kernel width on chromosomes 1A, 2A, 2B, 2DL and 3DL. Similarly, markers on chromosomes 1A, 1B, 3B and 7A were associated with test weight (grain volume weight). Although kernel width had many QTL in common with kernel area (obtained from digital image analysis) and test weight, it had no QTL in common with kernel length (Campbell et al., 1999). El-Feki (2010) reported QTL for kernel diameter on chromosomes 1A, 2B, 2D, 3B, 6A, 7B and 7D.

Huang et al. (2006) found QTL for test weight on chromosomes 4A, 4D, 2D, 5A and 7A with the QTL on 4D explaining the maximum 13.1% of the phenotypic variation. McCartney et al. (2005) found 10 QTL for test weight on chromosomes 1B, 1D, 2B, 2D, 3B, 3D, 4D, 5D, 6B and 7D with the most significant QTL on chromosomes 3B and 4DS. The QTL on 4DS coincided with a plant height QTL and explained 17.4% of the test weight variation. The reduced plant height at this QTL was also associated with reduced test weight. Test weight QTL were detected on chromosomes 1B, 6B, 7A and 7D in the study conducted by El-Feki (2010) in winter wheat in Colorado environments.

1.6 Phenological, morphological and drought related traits and their genetic control

1.6.1 Phenological traits: days to heading, days to maturity and grain filling duration

Heading time is an important trait for adaptation of wheat to its target environments including moisture stress areas (Lin et al., 2008). It is one of the traits effectively used in classical plant breeding programs as a mechanism of escaping terminal moisture stress and

freezing injury in early spring. Furthermore, heading date is a highly heritable trait in wheat and hence selection is usually effective.

Heading date is regulated by three well-characterized groups of loci namely, 1) vernalization requirement (*VRN*), 2) photoperiod response (*Ppd*) and 3) earliness per se (*Eps*). The requirement of exposing seeds to a period of cold temperature for flower induction in winter wheat is controlled by *VRN* genes (*VRN1* and *VRN2*). Vernalization genes, *VRN1* and *VRN2,* play an important role in vernalization pathways of both wheat and barley (Bennet et al., 2012; Fu et al., 2005). The *VRN1* genes are dominant for the spring growth habit whereas the *VRN2* genes are dominant for the winter growth habit (Yan et al., 2003, 2004a, 2004b). In wheat, the *VRN1* genes (*Vrn-A1, Vrn-B1* and *Vrn-D1*) have been mapped on the homoeologous group 5 chromosomes (5AL, 5BL and 5DL) (Lin et al., 2008). The *VRN2* locus is located on chromosome 5BL over 50 cM distal to the *Vrn-B1* gene, but it had strong epistatic interaction with the *VRN1* genes (Bennet et al., 2012). A dominant allele of *VRN3*, another vernalization gene mapped to chromosomes 7BS and 7DS, is responsible for spring growth habit (Kuchel et al., 2006; Yan et al., 2006; Chen et al., 2010). Generally, there are three types of wheat based on vernalization requirement, viz. winter, semi-winter (or facultative) and spring types.

Genes that regulate photoperiod sensitivity in wheat are located on group 2 homoeologous chromosomes. In order of their potency, these genes are *Ppd-D1*, *Ppd-B1* and *Ppd-A1*, located on chromosomes 2DS, 2BS and 2AS, respectively (Bennet et al., 2012; Beales et al., 2007; Worland et al., 1998). Wheat is naturally a long-day plant, and photoperiod-sensitive varieties need an extended period of long days to initiate floral primordia. There are some genotypes that can flower under short days (Griffiths et al., 2009). Photoperiod insensitivity in bread wheat is controlled by dominant alleles at the *Ppd-D1* and *Ppd-B1* loci, which ensure early

heading both under short-days and long-days by reducing sensitivity to photoperiod (Worland et al., 1994). Bennet et al. (2012) evaluated 368 doubled haploid lines in seven environments and genotyped the lines with 850 SSR markers. They found the most significant QTL for ear emergence time on chromosomes 2BS and 2DS which coincided with positions of *Ppd-B1* and *Ppd-D1*, respectively. They also detected significant QTL on chromosome 5BL where the *Vrn-B1* gene had been previously mapped. QTL were also detected on the homoeologous group 7 chromosomes, possibly associated with *VRN3* which was physically assigned to the chromosome arms 7AS, 7BS and 7DS (Yan et al., 2006).

Vernalization and photoperiod genes confer mainly gross adaptation to mega-environment. However, local adaptation (i.e., fine-tuning flowering time) is facilitated by the earliness *per se* (*Eps*) genes, which promote flowering independent of environmental signals (or vernalization and photoperiod response) (Bennet et al., 2012; Griffiths et al., 2009). Previous studies have shown that all chromosomes of wheat have been involved in the genetic control of earliness *per se* (Bennet et al., 2012). Chromosomes that harbored QTL for earliness *per se* include 2BL, 3A, 4B, 4D, 6B and 7B (Scarth and Law, 1983; Flood and Halloran, 1983). However, overall earliness *per se* genes are known to map to group 2 and 4 chromosomes and chromosomes 3A, 6B and 7B (Shah et al., 1999). From a study on recombinant inbred lines developed from spring wheat parents, Lin et al. (2008) observed that QTL for earliness *per se* contributed more to the flowering time variation in the population than vernalization and photoperiod genes.

Cuthbert et al. (2008) found seven QTL on chromosomes 1B, 2D, 3A, 5A, 6B, 7B and 7D, with the QTL on 5AS explaining 14.6% of the phenotypic variation in heading date. QTL for days to maturity were detected on chromosomes 1B, 3B, 5A, 5B, 6B, 7A, 7B and 7D. Out of

these, QTL on 6B explained the largest proportion of phenotypic variation (16.8%) followed by QTL on 7B (11.9%) and 5B (11.8%).

Wang et al. (2009) detected a total of eight QTL on chromosomes 1B, 2B, 3B, 5D and 6D for flowering time by evaluating 142 RIL in four environments while seven QTL were detected on chromosomes 1B, 2A, 4B, and 6D for days to maturity. Huang et al. (2006) found QTL for days to maturity on chromosomes 2D, 5D and 7D, and the maximum phenotypic variation was explained by QTL on chromosome 7D. Results from Marza et al. (2006) also indicated the presence of QTL on chromosomes 3BL, 5B and 6B for heading date, while maturity date QTL were detected on chromosomes 1B, 3AS and 6B. Kulwal et al. (2012) studied 208 elite soft white winter wheat lines for one year and detected QTL for days to heading on the short arm of chromosome 2B. Chu et al. (2008) detected QTL for heading date on chromosomes 5AL and 5BL by studying 120 doubled haploid wheat lines, with a QTL detected on 5BL explaining 39% of the phenotypic variation in heading date. In the association analysis of Neumann et al. (2011) marker-trait associations shared between heading date and flowering time were located on chromosomes 1B, 5D, 6A and 7A. El-Feki (2010) found seven QTL for heading date; a QTL detected on chromosome 7D within the interval of the *VRN3* gene was detected in all test environments and explained 32.1 to 42.6% of the phenotypic variation.

According to the study conducted by Cuthbert et al. (2008) on doubled haploid lines derived from two spring wheat parents, QTL for grain-filling duration (a period from heading date to physiological maturity) were detected on chromosomes 2D, 3A, 5A, 5B and 7D; the explained phenotypic variation was in the range of 3.9-7.2%. Hanocq et al. (2007) conducted meta-QTL analysis from 13 previous independent studies and found that photoperiod response genes on group 2 and vernalization genes on group 5 chromosomes were frequently involved in

controlling heading date in wheat with intermediate effects of other heading date QTL on chromosomes 2B, 4A, 4B and 5B.

Wang et al. (2009) identified six QTL for grain-filling duration on chromosomes 1A, 3B, 5D and 6D across two environments. Moreover, the phenotypic correlation of grain-filling duration with flowering time was negative and significant, whereas its correlation with days to maturity was negligible. Furthermore, the heritability of grain-filling duration was the lowest of all the traits evaluated such as grain filling rate, thousand kernel weight, flowering time, kernel number spike^{-1}, kernel weight spike^{-1}, days to maturity and plant height. It has been suggested that yield improvement efforts should be focused more on grain-filling rate than grain-filling duration (Wang et al., 2009). Kirigwi et al. (2007) identified QTL for grain-filling rate on chromosome 4A.

1.6.1.2 Leaf senescence

"Stay green" or delayed senescence plays an important role in grain development during stress conditions by allowing plants to retain their green leaves for a prolonged grain filling period (Kumari et al., 2012). Leaf senescence is a genetically programmed cell death that can be modified by both abiotic and biotic factors. It is a result of catabolism of chlorophyll, proteins, lipids and nutrient remobilization into developing grains (Vijayalakshimi et al., 2010; Srivalli and Khanna-Chopra, 2009). Stay green might contribute to higher yield and enhance heat tolerance if normal photosynthesis duration is prolonged and/or there is high intrinsic chlorophyll concentration. There is a stay green type which results from the lesions in the chlorophyll catabolism and a lack of photosynthesis competence (Kumar et al., 2010; Keran et al., 2007; Kumari et al., 2007).

The genetic basis of the stay green trait has been studied in crops including rice, soybean (*Glycine max* (L.), sorghum, maize and bread wheat (Kumar et al., 2010; Joshi et al., 2007a; Spano et al., 2003; Silva et al., 2000). Those studies confirmed the presence of genetic variation and the quantitative nature of the stay green trait for different crop species (Vijayalakshimi et al., 2010; Thomas and Howarth, 2000). For rice, 46 QTL distributed on all 12 chromosomes have been reported while 83 stay green QTL have been reported for sorghum in a total of seven studies (Mace and Jordan, 2011). Stay green QTL have also been reported for maize (Bertin and Gallais, 2001). The winter wheat flag leaf senescence QTL were detected on long arms of chromosomes 2D and 2B under drought stress and irrigated condition, respectively (Verma et al., 2004). Another stay green QTL mapping study in wheat identified three QTL on chromosomes 1AS, 3BS and 7DS from a recombinant inbred population developed from crossing stay green and non-stay green parents (Kumar et al., 2010). The QTL of stay green on 3BS was found in the same region with QTL for plant height. In that study, stay green was positively correlated with grain-filling duration, grain yield and biomass, indicating the importance of stay green for grain yield particularly under stress conditions. A QTL mapping study conducted in a winter wheat population of 101 RIL indicated that QTL for leaf senescence-related traits were distributed on most chromosomes under optimum conditions, but under heat stress they were mainly located on chromosomes 2A, 6A and 6B (Vijayalakshimi et al., 2010). Naruoka et al. (2011) detected stable QTL on chromosomes 2D and 5B for green leaf duration after heading by evaluating 91 RIL of a bi-parental spring wheat mapping population across several environments. These QTL co-segregated with *Ppd-D1* and *Vrn-B1* genes. El-Feki (2010) detected only two QTL on chromosomes 2B and 7D for flag leaf senescence measured in a single environment.

1.6.2 Morphological and drought related traits

1.6.2.1 Plant height

Optimum plant height is required for better yield in wheat, as tall plants are susceptible to lodging and excessively short plants are often associated with a yield penalty in resource limited areas (e.g., moisture stress environment) (Griffiths et al., 2012). Dwarfing genes *Rht-D1b* and *Rht-B1b*, which are insensitive to gibberellic acid, have increased grain yield in most resource rich environments through reducing lodging susceptibility and increased grain number (Rebetzke et al., 2012). Since the *Rht-D1b* and *Rht-B1b* alleles are also associated with reduced coleoptile length and poor seedling vigor, there is an interest in introducing alternative gibberellic acid responsive dwarfing alleles with a potential for reducing plant height without affecting coleoptile length. The *Rht8* gene on chromosome 2DS is a potential candidate in the development of semi-dwarf wheat varieties with long-coleoptiles (Rebetzke et al., 2012; Griffiths et al., 2012).

Besides dwarfing genes, photoperiod-insensitive alleles at *Ppd-D1* on chromosome 2DS and *Ppd-B1* on 2BS have pleiotropic effects on plant height (Griffiths et al., 2012). In addition to these major genes, several studies indicated the presence of QTL for plant height. In the Neumann et al. (2011) study, marker-trait associations common to plant height and peduncle length were detected on chromosomes 1A, 2B, 4A and 7B and plant height-specific markers were located on chromosome 1B, 4A, 6B and 7A.

A significant association of DArT marker wpt730772 with plant height was detected on chromosome 6AS by Kulwal et al. (2012) who studied 208 elite soft white winter lines for one season in a single environment. Haung et al. (2003) found plant height QTL on chromosomes 2BL, 4BL, 4DS, 6AL and 7BS. However, McCartney et al. (2005) detected plant height QTL next to *Rht-D1* gene on chromosome 4D; they also found another minor plant height QTL on

chromosome 2D that was not related to *Rht8* or *Ppd-D1*, but may correspond to the peduncle length QTL reported by Borner et al. (2002). Marza et al. (2006) found QTL for plant height on chromosomes 2BL, 2BS, 2DL, 3BL, 4B and 6A based on a RIL population evaluated at three locations for three seasons.

Wang et al. (2009) detected a total of six QTL for plant height on chromosomes 1D, 2D, 3D and 4D in two of the total four environments. Griffiths et al. (2012) studied four doubled haploid populations with population size ranging from 93 to 202, and found 16 QTL for plant height on chromosome 1A, 1B, 1D, 2A, 2B, 2D, 3A, 3B, 4D, 5A, 5B, 6A, 6B and 6D. With the exception of homoeologous group 7, at least one chromosome from all homoeologous groups harbors a region that affects plant height. Liu et al. (2010) detected a marker-trait association for plant height on chromosome 4A (7.6 cM). Huang et al. (2006) found four QTL for plant height on chromosomes 4B, 4D, 5D and 7B using 185 DH lines, and the QTL on chromosome 4D (closest marker *Xwmc52*) explained 29.2% of the phenotypic variation. Ten plant height QTL were detected in the study conducted in Colorado environments, and a QTL detected on chromosome 6A was stable across environments (El-Feki, 2010).

1.6.2.2 Flag leaf width, length and flag leaf area

Flag leaf photosynthesis contributes 30 to 50% of the canopy photo-assimilates during grain-filling in wheat (Lupton, 1966). Inoue et al. (2004) also found a higher photosynthetic rate for a drought tolerant wheat cultivar compared to a drought sensitive one. However, water deficit during the period of leaf expansion reduces crop leaf area, radiation interception and green leaf duration, and accelerates senescence; consequently, yield component traits such as kernel number and harvest index are directly affected (Araus et al., 2008). Although flag leaf width and length measurements can be conducted more quickly than most other yield component traits,

previous studies showed that the associations of these traits with economically important traits (e.g., grain yield) were either absent or inconsistent across environments (Hansen et al., 2005; Blake et al., 2007). Nine flag leaf QTL that individually explained 5.2 to 11.0% of the phenotypic variation were reported by El-Feki (2010), and three QTL on chromosomes 1B, 2B and 6B were detected in two environments. In the same study, eight QTL were obtained for flag leaf width, and only two QTL on chromosome 1B and 2D were detected in two environments. The remaining flag leaf width QTL were environment specific.

1.6.2.3 Normalized difference vegetation index (NDVI) and drought susceptibility index

Spectral reflectance indices have shown promise to estimate biomass production, yield, relative water content and nutrient deficiencies (Gutierrez et. al., 2010; Prasad et al., 2007). The basic idea of spectral reflectance properties is that a trait is associated with absorption of light at specific wavelengths and show unique reflectance patterns at specific wavelengths of the light spectrum (Reynolds et al., 1999). NDVI has been applied for estimating overall canopy greenness, nitrogen use efficiency and grain yield in wheat drought tolerance research. It has been hypothesized that NDVI may be used for indirectly selecting for higher biomass and yield in breeding programs. However, there have been contradictory reports regarding the relationship between NDVI measurements at different growth stages and yield (Hazratkulova et al., 2012). Some studies indicated that NDVI was well associated with yield during the grain-filling stage but not during the vegetative stage (Freeman et al., 2003; Hazratkulova et al., 2012), and other studies showed an association of NDVI with yield at the booting, heading and grain-filling stages of wheat (Babar et al., 2006; Wang et al., 2010). Identifying the chromosome regions that NDVI shares with biomass and yield might be useful to understand the genetic basis for the relationship

between NDVI with other traits and efficiently utilize the trait as an indirect selection criterion in grain yield improvement.

Association analysis of 305 diverse maize lines genotyped for 2052 SNPs indicated that 30 SNPs (23 QTL) were associated with NDVI measured at seven stages, both under dry and irrigated conditions, with three loci in common with plant height (Lu et al., 2012). In wheat, QTL have been detected for NDVI on chromosomes 2BL and 3BS based on the study conducted on a population of 249 RILs of durum wheat which were evaluated across a broad range of Mediterranean environments (Maccaferri et al., 2008). El-Feki (2010) identified three marginally significant QTL for NDVI on chromosomes 3A, 4B and 6A in a single environment for a winter wheat bi-parental doubled haploid population.

The drought susceptibility index (DSI) has been used as a criterion for distinguishing drought tolerant genotypes from susceptible ones. It is derived from the yield difference between non-stress and stressed environments (Blum et al., 1989). Kirigwi et al. (2007) studied 127 RILs of spring wheat under irrigated and moisture stress conditions, and detected QTL for DSI on chromosome 4AL at marker positions 64.4, 80.3 and 84.9 cM with the QTL detected at marker position 64.4 cM explaining 41% of the phenotypic variation. Similarly, Kadam et al. (2012) evaluated 206 spring wheat RILs derived from a cross between a high yielding but drought susceptible variety and a low yielding but drought tolerant variety under drought and control conditions for two years. A consistent QTL for DSI was detected on chromosome 4BS. This QTL was also associated with grain yield per plant, harvest index and root biomass under drought. El-Feki (2010) reported DSI QTL on chromosomes 5B and 7B for a winter doubled haploid bi-parental mapping population that was evaluated under moisture stress and irrigated conditions in Colorado.

Although many QTL have been reported for yield and yield components, and drought-related traits for wheat, the majority of the QTL have been detected with bi-parental populations with low resolution. In addition, limited information is available for effects of drought tolerance candidate genes on yield, yield components, and drought tolerance related traits.

Therefore, the objectives of the current study are:

1. To test the association of chromosome regions with yield, yield components, and drought tolerance-related traits using a genome-wide association mapping approach.

2. To associate polymorphisms in selected drought tolerance candidate genes with yield, yield components, and drought tolerance-related traits.

CHAPTER 2

Genome Wide Association Mapping for Yield and Yield Components of Spring Wheat under Contrasting Moisture Regimes

SUMMARY

Genome-wide association mapping is becoming a widespread method to identify quantitative trait loci (QTL) in crop plants including wheat (*Triticum aestivum* L.). Its benefit over traditional bi-parental mapping approaches depends on the extent of linkage disequilibrium (LD) in the mapping population and dense marker coverage across the genome.

The objectives of this study were to determine LD decay rate and population structure in a spring wheat association mapping panel and to identify markers associated with yield and yield components, morphological, phenological, and drought tolerance-related traits. The study was conducted under fully irrigated and rainfed conditions at Greeley, CO and Melkassa, Ethiopia in 2010 and 2011 (five total environments). The size of the panel varied from 285 to 294 depending on the location and year.

Genotypic correlation coefficients and heritability estimates were calculated for each trait. Grain yield was positively correlated with kernel number, harvest index, final biomass and test weight, but negatively correlated with days to heading. Grain volume weight and single kernel weight, diameter, and hardness had higher heritability estimates than yield and the remaining yield component traits.

Genotypic data was generated for 287 lines using Diversity Array Technology (DArT) markers. LD, population structure and kinship were determined for the mapping population from

the marker data. LD decay rate extended over a longer genetic distance (6.8 cM) for the D genome than for the A and B genomes (1.7 and 2.0 cM, respectively). Population structure, kinship and marker data were used in a mixed model to associate markers with phenotypic traits. A stable QTL was detected for grain yield on chromosome arm 2DS both under irrigated and rainfed conditions. A multi-trait region significant for yield and yield components was detected on chromosome 5B. A grain yield QTL on chromosome 1BS was co-localized with harvest index QTL, explaining the high positive genotypic correlation between grain yield and harvest index. A QTL for NDVI overlapped with a harvest index QTL on chromosome 1AL, while green leaf area shared a QTL region with harvest index on chromosome 5A. Clusters of QTL for flag leaf characters (leaf area, length, and width) were detected on chromosome arms 3BL and 5BL. Heading date QTL were detected on chromosomes 2B, 3AL, 3B and 7DS, while plant height QTL were detected on chromosome arms 3BL, 6AS and 7BL. Generally, in this study both stable and environment-specific QTL were detected for yield, yield components, and drought tolerance-related traits. After validation of their effects, the detected QTL may be used in breeding programs to improve performance of wheat under both irrigated and rainfed conditions.

2.0 INTRODUCTION

Wheat is the world's third most important food crop next to maize (*Zea mays* L.) and rice (*Oryza sativa* L.) (Green et al., 2012). It accounts for 19% of total production among major cereal crops and provides 55% of the carbohydrates consumed by humans around the world (Gupta et al., 1999; Bagge et al., 2007). However, its productivity is often reduced by both biotic and abiotic stresses and its potential yield is rarely achieved.

Drought is one of the most serious factors reducing crop productivity throughout the world (Peleg et al., 2009; Salekdeh et al., 2009; Ahuja et al., 2010), regularly affecting as much as 50% of the global wheat production area (Pfeiffer et al., 2005). Since water is a limiting factor in crop production, all factors that limit plant access to water aggravate the impact of drought. Furthermore, anticipated world-wide climate change will elevate temperature which accelerates evaporative-transpiration loss during the day and increases photorespiration at night (Habash et al., 2009; Mir et al., 2012). This results in reduced crop productivity and thus food insecurity on a global scale.

Plant breeding has successfully improved crop resistance to both biotic and abiotic stresses, including drought, through phenotypic selection (Araus et al., 2008; Cooper et al., 2009). However, the progress has generally been slow, and there is a large yield gap between drought prone areas and ideal production regions for most crops, including wheat. Many previous studies have shown that tolerance to drought is a complex quantitative trait that involves multiple chromosome regions (Fleury et al., 2010; Barnabas et al., 2008; Ravi et al., 2011; Mir et al., 2012). It is further complicated by the fact that the degree of drought effects on plants depends on timing, duration and intensity of drought, and different traits may be required for different patterns of drought (Passioura, 2012). Plant drought resistance can be broadly

categorized into three mechanisms (Levitt, 1972): 1) dehydration avoidance, maintaining cellular moisture through strategies of increasing soil moisture capture, water use efficiency and osmotic adjustment 2) dehydration tolerance, e.g., stem reserve mobilization and delayed senescence (i.e., stay green) 3) dehydration escape, e.g., early flowering. Suitable crop cultivars should combine drought resistance mechanisms with high and stable yield for sustainable crop production in drought prone regions (Habash et al., 2009).

Genetic improvement under drought requires identifying sources of traits associated with drought tolerance and introgressing the genes underlying the target traits to locally adapted cultivars. The challenge for implementing this strategy in breeding programs is the identification of the most suitable target traits in a time-efficient and cost-effective way for different drought scenarios (Passioura, 2012). Recent advancements in high throughput genotyping and phenotyping have improved understanding of the physiological and molecular bases underlying complex traits including drought tolerance (Collins et al., 2008; Habash et al., 2009; Mir et al., 2012; Sinclair, 2012). QTL mapping is a key approach for understanding the genetic architecture of complex traits in plants (Holland, 2007). However, QTL mapping using bi-parental populations explains only a small portion of the genetic architecture of a trait because only two alleles per locus can be evaluated at a time. Other limitations of bi-parental populations are low mapping resolution, population specificity of detected QTL, and the long time required to develop mapping populations. These limitations have partly contributed to the slow transfer of knowledge from bi-parental QTL studies to practical applications in plant breeding.

The advent of association mapping approaches has overcome some of the limitations of bi-parental mapping populations. Since association mapping utilizes diverse germplasm, QTL for many traits can be detected at high-resolution in a single study, making the method more

efficient and less expensive than bi-parental QTL mapping (Breseghello and Sorrells, 2006; Ersoz et al., 2009; Sorrells and Yu, 2009; Waugh et al., 2009). The resolution and power of association studies, however, depend on the extent of LD across the genome. LD needs to be determined in each study as it is affected by several factors such as population history, recombination frequency and mating system.

The correlation of allele frequency (r^2) among the markers is the common statistic used to measure LD (Gupta, 2005; Sorrells and Yu, 2009). LD is expected to decay as a function of the nucleotide or linkage distance, as recombination reduces LD. This guides decisions on the number of markers required to conduct association mapping in a crop species (Waugh et al., 2009). To visualize LD patterns and the rate of LD decay for a chromosome, r^2 values are usually plotted against nucleotide or linkage distance (Abdurakhmonov and Abdukarimov, 2008).

Previous studies have demonstrated unique LD patterns for different crop species and populations within a species, with rapid levels of LD decay observed in cross-pollinated species (e.g., maize) compared to self-pollinated species (e.g., wheat) (Wilson et al., 2004; Chao et al., 2007; Comadran et al., 2009). Although association mapping has advantages over bi-parental populations, QTL identification could be confounded by population subgroups and plant phenology. Another limitation of this method is that markers with low allele frequencies are often not considered in association analysis. However, statistical models have been developed to account for population structure and familial relationship among the genotypes in the mapping panel (Yu et al., 2006).

Association mapping has been used successfully to detect QTL in wheat for disease resistance (Crossa et al., 2007; Maccaferri et al., 2010; Yu et al., 2011; Yu et al., 2012; Adhikari

et al., 2012); end-use quality traits (Breseghello and Sorrells, 2006; Zheng et al., 2009), Russian wheat aphid (*Diuraphis noxia*) resistance (Peng et al., 2009), and yield and yield component traits (Maccaferri et al., 2011; Neumann et al., 2011). The suitability of DArT markers for association studies has been proved particularly for species lacking cost effective single nucleotide polymorphism (SNP) markers (Benson et al., 2012). However, the majority of previous studies have been conducted either with low marker density or a small number of lines in the mapping population. Therefore, the objectives of the present study were to (1) determine LD decay rate in a spring wheat association mapping panel, (2) analyze population structure in the panel, and (3) identify markers associated with yield and yield components, morphological, phenological and drought tolerance-related traits.

.

2.1 MATERIALS AND METHODS

2.1.1 Mapping population

The spring wheat association mapping panel used in this study (WAMII, wheat association mapping II) was originally developed by the International Maize and Wheat Improvement Center (CIMMYT) with the intention of identifying QTL/genes for drought and heat tolerance. The panel comprised a total of 294 diverse lines which were assembled from the Elite Spring Wheat Yield Trial (26^{th}, 27^{th} and 28^{th} ESWYT), Semiarid Wheat Yield Trial (1^{st} to 16^{th} SAWYT) and High Temperature Wheat Yield Trial (HTWYT) (Lopes et al., 2012). Many synthetic hexaploid-derived wheat lines were included in the panel (Lopes and Reynolds, 2012). A complete list of the association mapping panel (WAMII) is presented in Table A.1. In the study reported here, 283 to 294 lines were evaluated depending on the location and year due to limitations in seed quantity.

2.1.2 Experimental design and phenotypic trait evaluation

In 2010, a total 285 lines (including two local check cultivars, Reeder and Butte 86 (Mergoum et al., 2006)) were evaluated under fully irrigated conditions at the USDA-Agricultural Research Service Limited Irrigation Research Farm in Greeley, CO (latitude 40º 27' N; longitude 104º 38' W; elevation 1427 m). The trial was planted on April 5, 2010. The soil at the site is well-drained with fine sandy loam to clay loam texture and a pH range of 7.4-8.4. The site received a total of 271 mm of rainfall from January through July in 2010 (Table A.2), and the experimental plots were supplemented with 93.8 mm from three irrigations (twice during the vegetative stage and once after heading).

In 2011, we evaluated 288 lines (including two local checks, SD3870, a breeding line form South Dakota and Granger (Glover et al., 2006)) at Greeley under both fully irrigated and

rainfed conditions. The irrigated treatment was supplemented three times with drip irrigation, (twice before flowering and once during the grain filling stage), while the rainfed treatment was irrigated only once at flowering to avoid complete failure of the experiment. Therefore, the irrigated treatment received a total of 313 mm water (rain plus irrigation) whereas the rainfed treatment received only 192 mm water (rain plus irrigation) from January to July (Table A.3).

In the 2010 and 2011 experiments at Greeley, the lines were planted in four-row plots 1.52 m wide and 0.92 m long with 0.20 m spacing between rows and a seeding rate of approximately 173 seeds m^{-2}. Each entry was replicated twice in a Latinized incomplete block row-column design with CycDesign 3.0 software (www.cycdesign.co.nz). The experimental field was maintained free of weeds by manually removing weeds as required. In both seasons, the plants experienced heat stress mainly from heading through physiological maturity, as maximum temperatures were over 30^{o}C for a total of 13 days in June and 22 days in July 2010; temperatures exceeded 30^{o}C for 15 days in June and 27 days in July 2011.

A total of 294 lines was planted at the Melkassa Agricultural Research Center of the Ethiopian Institute of Agricultural Research, Melkassa, Ethiopia (latitude 8^{o} 24' N; longitude of 39^{o} 21'E, elevation 1550 m), on 17 July 2011 on wet soil from rain in the previous few days. The same set of lines was planted on drier soil on 19 July 2011. The dominant soil type at Melkassa is sandy loam (Andosol of volcanic origin) with pH ranging from 7.0 to 8.2 (). The experiment was laid out as an alpha lattice design with 14 entries per incomplete block and two replications. A two-row plot of length 2.5 m, width 0.4 m and between row spacing of 0.20 m was used. Seeding rate was based on the local recommendation of 150 kg ha^{-1}. Nitrogen fertilizer was applied in split doses at planting and tillering at a rate of 50 kg ha^{-1} for each dose. Phosphorus fertilizer was applied as diammonium phosphate at planting at the rate of 100 kg ha^{-1}. The site

received a total of 533 mm rainfall during the growing season (July-September, 2011). The average maximum and minimum temperatures for the four month period (July-October, 2011) at Melkassa were 27.3 ^{o}C and 8.5 ^{o}C, respectively (Table A.4). The temperature was below 30 ^{o}C for all days from emergence time through physiological maturity.

The phenotypic traits evaluated in this study are defined as follows. Plant height was recorded as the average of three values for each plot measured in cm from the soil surface to the tip of the spike excluding awns. Days to heading was recorded as the number of days from planting until 50% of the spikes in each plot had completely emerged above the flag leaves. Days to maturity was recorded as the number of days from planting until 50% of the peduncles in each plot had turned yellow. Grain filling duration was calculated as the difference between the days to heading and days to maturity. Normalized vegetation index (NDVI) was obtained by scanning plants in each plot at the grain filling stage with a GreenSeeker instrument model 3541 (NTech Industries Inc., Boulder, CO). Green leaf area was obtained from a photo taken at a height of approximately 0.5 m directly above each plot with a digital camera (Coolpix S8100, Nikon Corp., Japan) during vegetative stage , and pictures were processed with Breedpix software (Casadesus et al., 2007). Leaf senescence was scored on a scale from 0 to10, where 0 indicates completely green leaves and 10 indicates that all leaves in a plot had changed completely to yellow. Flag leaf length (measured from the leaf collar to the tip) and width (measured at the widest part of the flag leaf) were recorded as the average measurement of three flag leaves per plot. Flag leaf area (cm^2) was calculated as flag leaf length x flag leaf width x 0.75.

Single kernel diameter (mm), kernel hardness and single kernel weight (mg) were determined from 100 seeds (sampled from grain yield of biomass sample) in a single kernel characterization system instrument Model 4100 (Perten Instruments, Springfield, IL). Spike

length, spikelet number per spike, kernel number and weight (g) per spike, and kernel number per spikelet were recorded as the average of five spikes per plot. Thousand kernel weight was determined by extrapolation after counting seeds of five spikes with a seed counter (International Marketing and Design Corp Model 900-2; San Antonio, TX) and obtaining the weight of the seeds. Number of spikes m^{-2} was calculated by dividing the number of kernels m^{-2} by kernel number per spike. The number of kernels m^{-2} was obtained from the ratio of grain weight m^{-2} to thousand kernel weight, multiplied by 1000. Final dry biomass was determined by weighing samples after 48 hours in a 40 ^{o}C drier. These samples were threshed to obtain grain weight, and harvest index was recorded as the ratio of grain weight to oven-dried biomass of the 1-m strip. Grain yield was the total weight of seed yield in each plot divided by the plot area and expressed as kg ha^{-1}.

Test weight (kg hL^{-1}) was determined using standard procedures from a small sample of the grain collected at harvest. Drought susceptibility index (DSI) was calculated using grain yield and kernel number m^{-2} under irrigated and rainfed conditions as described by Fisher and Maurer (1977). DSI= $(1-Y_d/Y_i)/DII$, where Y_d=yield of each line in the dry treatment, Y_i=yield of each line under fully irrigated conditions and DII =$1-(Y_{dm}/Y_{im})$ where Y_{dm} is the average yield of the dry treatment and Y_{im} is the average yield of the irrigated treatment. Traits measured in this study are presented in Table 2.1. Some of the traits were not measured in all environments.

Table 2.1. Lists of Traits evaluated in the WAMII spring wheat association mapping panel in five environments.

		Environments†				
Traits	Abbreviations	GRW10	GRW11	GRD11	MLKW11	MLKD11
Grain yield	GYLD	X	X	X	X	X
Thousand kernel weight	TKW	X	X	X	X	X
Harvest index	HI	X	X	X	X	X
Kernel number m^{-2}	KN	X	X	X	X	X
Spike length	SL	X	X	X	X	X
Kernel number $spike^{-1}$	KNS	X	X	X		
Spike number	SN	X	X	X		
Single kernel weight	SKW	X	X	X		
Single kernel diameter	SKD	X	X	X		
Single kernel hardness	SKH	X	X	X		
Kernel weight $spike^{-1}$	KWS	X	X	X		
Kernel number $spikelet^{-1}$	KNL	X	X	X		
Spikelet number $spike^{-1}$	SPN	X	X	X	X	X
Biomass	BM	X	X	X	X	X

†GRW10, Greeley irrigated treatment in 2010; GRW11; Greeley irrigated treatment in 2011; GRD11, Greeley rainfed in 2011; MLKW11; Melkassa non-stressed treatment in 2011; MLKD11, Melkassa stressed treatment 2011.

Table 2.1. Continued

		Environments†				
Traits	Abbreviations	GRW10	GRW11	GRD11	MLKW11	MLKD11
Test weight	TW	X	X	X		
Day to heading	DH	X	X	X	X	X
Days to maturity	DM	X	X	X	X	X
Grain filling duration	GFD	X	X	X	X	X
Flag leaf length	LL	X	X	X	X	X
Flag leaf width	LW	X	X	X	X	X
Flag leaf area	LA	X	X	X	X	X
Leaf senescence	LS	X		X		
Plant height	PHT	X	X	X	X	X
Normalized difference vegetation index	NDVI	X	X	X		
Green leaf area	GA	X	X	X	X	X
Kernel number- based drought susceptibility index	DSI_KN			X	X	
Grain yield-based drought susceptibility index	DSI-YLD			X	X	

†GRW10, Greeley irrigated treatment in 2010; GRW11; Greeley irrigated treatment in 2011; GRD11, Greeley rainfed in 2011; MLKW11; Melkassa non-stressed treatment in 2011; MLKD11, Melkassa stressed treatment 2011.

2.2 Statistical analysis

2.2.1 Phenotypic data analysis

The phenotypic data analyses were conducted with SAS v. 9.3 software (SAS Institute Inc., Cary, NC). First, the general linear model (GLM) procedure was used get best linear unbiased estimates, considering genotype, replications, rows and columns as fixed in the model for each environment. Normality of the data for each trait was checked using a Q-Q plot of residuals in the SAS GLIMMIX procedure. The presence of statistically significant differences among the genotypes for each trait was also checked with the GLM procedure. Then, best linear unbiased predictions (BLUPs) and variance components were obtained for all traits using a Mixed model procedure, considering genotypes as random and all other factors in the model as fixed. Environment was considered fixed in the combined data analysis. To account for spatial variations in the experimental field, four spatial variability adjustment models (spatial power, anisotropic spatial power, Matérn spatial and autoregressive) were tested for each trait. The correlation values due to spatial variability in each model were found to be very low for all data sets except for Greeley in 2010. Thus, the autoregressive spatial adjustment model was applied for the data set in 2010, but no adjustment was made for the remaining environments.

Broad sense heritability (h^2) for all traits in each environment and the combined dataset was calculated from variance components (obtained from SAS PROC VARCOMP) as: h^2 = genotypic variance / (genotypic variance + error variance/r) where r = number of replications for a single environment. For combined data, heritability estimates were calculated as genotypic variance/ ((genotypic variance + (GxE variance/n) + (error variance/nr)) where, n =number of environments. Genotypic correlations among traits were estimated using the restricted maximum

likelihood estimation (REML) of genotypic variance and covariance components as described by Holland (2006).

2.1.3 Genotypic data analysis

DArT marker genotypes were obtained following the procedures of Akbari et al. (2006) at Triticarte Pty. Ltd. (Canberra, Australia; http://www.triticarte.com.au), a whole genome profiling service laboratory. A total of 1863 DArT markers were used in the analyses, after markers with < 5% allele frequency and those with a high percentage of missing data points (> 6%) were removed. Genome-wise distribution of the markers was 558 on genome A, 617 on genome B, and 290 on genome D (http://www.triticarte.com.au). Chromosome map positions were not known for 398 markers. A DArT marker physical map (based on Chinese spring wheat deletion lines) (http://www.cerealdb.uk.net/) was used to assign trait-associated markers to chromosome arms.

2.1.3.1 Population structure and linkage disequilibrium analyses

Seventy-eight markers (3-4 markers spaced > 10 cM per chromosome) were selected from all chromosomes (except for chromosome 4D and 5D) from a total of 1863 markers for analysis of population structure. To determine population structure, an admixture model with correlated allele frequency in STRUCTURE software was applied (Pritchard et al., 2000). A burn-in of 20, 000 iterations followed by 20,000 Monte Carlo Markov Chain (MCMC) replicates was conducted to test k values (number of subpopulations) in the range of three to 12. Each k was replicated five times and the run that assigned the most lines with probability >0.5 in all clusters was used. The likely number of subpopulations was determined using the approach of Evanno et al. (2005) and the likelihood distribution of k was examined. Genetic distance-based cluster analysis was conducted using hclust script in the R package (www.cran.r-project.org)

using the same 78 markers to compare results with STRUCTURE software output. Multiple regression analysis was also done for all phenotypic traits using population subgroups in the model to determine the extent of the confounding effect of population structure on the phenotypic traits.

LD among markers was calculated using observed versus expected allele frequencies of the markers in TASSEL v.3.0 (Bradbury et al., 2007). Only mapped markers were used for LD calculation both for the panel and for model-based subgroups. The critical r^2 value beyond which LD is due to true physical linkage was determined by taking the 95^{th} percentile of the square root transformed r^2 data of unlinked markers (Breseghello and Sorrells, 2006). The percentage of marker pairs significant at different critical r^2 values (0.2 and 0.2641) and $P<0.001$ was determined for each chromosome to compare the degree of LD among chromosomes. Locally weighted polynomial regression (LOESS) based curves were fitted on scatter plots of r^2 versus distance among markers. LOESS is a non-parametric method of estimating local regression surfaces, and it is a robust fitting method particularly when there are outliers in the data (Cleveland, 1979). The LOESS model is written as:

$y_i = g(x_i) + \varepsilon_i$, where y_i is i^{th} measurement for a response variable y, x_i is the corresponding measurement of a predictor variable x, ε_i is a random error and g is the regression function.

Analysis of molecular variance was conducted using the seven groups with Arlequin software (http://cmpg.unibe.ch/software/arlequin3).

2.1.3.2 Marker-trait association (MTA) analysis

A total of 1863 high quality DArT markers (missing data <6%) was used in this study. The GAPIT (Genomic Association and Prediction Integrated Tool) R package (Lipka et al., 2012) was used to determine the association between markers and phenotypic traits. GAPIT uses all algorithms implemented in TASSEL software, but association analysis can be done much faster in GAPIT than TASSEL. MTA analysis was conducted for each environment separately and combined data using BLUPs for each trait. A mixed linear model was employed by including BLUPs, markers, kinship matrix (K) and probability of membership of each line (Q) in the model for each trait (Yu et al., 2006). The kinship matrix was calculated as implemented in TASSEL software. Kinship is calculated in TASSEL as the proportion of alleles shared between each pair of lines. Once this matrix is calculated, the numbers are rescaled between 0 and 2 (Bradbury et al., 2007).

Model comparison was made among K (kinship) using the GLM model, Q+K (population structure and kinship) using the mixed model, and P+K (principal component and kinship) using the mixed model. Mean square of the difference (MSD) based on observed *P*-values and expected *P*-values, and a Q-Q plot were used to compare the models; MTA *P*-values for yield of five environments plus the combined data set were used for model comparisons. Among the three models, the model taking into account population structure and genotype relationship showed the least deviation from the nominal alpha level in most cases, and it was found to be better in controlling false positives. For multiple comparison adjustment, false discovery rate (FDR) adjusted *P*-values were calculated for each trait (Benjamini and Hochberg, 1995), and FDR=0.05 was taken as the threshold for significance of marker-trait associations.

However, since many markers within 10 cM distance were in LD, FDR adjustment is still too stringent as it assumes independent testing.

2.2 RESULTS

Wide differences among the study materials were visually observed for many traits during field evaluations. Water deficit reduced full expression of those traits in rainfed treatments despite increased expression for some drought related traits (e.g., leaf waxiness). The rainfed treatments at Melkassa experienced water deficit only at the emergence stage, whereas at Greeley the rainfed treatment was exposed to water deficit starting from vegetative stage through grain filling. Data were collected for a total of 26 traits, but this number varied depending on the year and location.

2.2.1 Agronomic trait means

Analysis of variance showed significant differences ($P<0.05$) among genotypes for most traits in all environments and for the combined analysis across environments. The mean grain yield of individual lines in the five environments were within the range of 1087 kg/ha (recorded at Greeley in 2011 under rainfed conditions) to 5377 kg/ha (obtained at Melkassa under non-stressed conditions). The mean grain yield (2156 kg/ha) recorded under fully irrigated conditions in 2010 was the highest of the three trials grown in Greeley (Table 2.2). The vegetative stage of the lines (calculated as the number of days from planting to heading) was longer in the Greeley environments (mean 68 days) than in the Melkassa environments (mean 55 days). However, the grain filling duration at Melkassa was longer than that of the Greeley environments (39 vs 34 days). On average, the genotypes headed 13 days earlier at Melkassa than at Greeley. Heading date ranged from 47 to 69 days at Melkassa and from 63 to 72 days at Greeley (Tables 2.2 and 2.3). In Melkassa environments plants grew taller than in the Greeley environments.

Table 2.2. Mean values of the WAMII spring wheat association mapping panel for traits measured under rainfed and well-watered conditions at Greeley, CO in 2010 and 2011.

	Environments†									
	GRW10 (285)			GRW11 (288)			GRD11(288)			
Trait‡	Mean	Standard deviation	Range	Mean (W)	Standard deviation	Range	Mean (D)	Standard deviation	Range	% reduction (D/WX100)
YLD	2156.00	229.58	1510.00-2791.00	1524.00	116.63	1179.00-1943.00	1304.00	74.18	1087.00-1511.00	14.38
TKW	35.65	2.33	31.00-43.20	21.46	0.86	18.90-24.50	25.15	1.64	20.80-31.60	-17.21
DH	67.18	1.37	62.50-71.60	70.03	1.01	67.70-72.30	68.37	1.04	65.60-71.10	2.38
LL	16.07	1.49	13.00-21.40	15.49	0.68	13.30-19.00	12.27	0.67	10.40-14.20	20.79
LW	1.51	0.10	1.30-2.00	1.31	0.03	1.20-1.40	1.24	0.03	1.20-1.30	5.50
DM	103.26	1.37	99.60-107.20	104.14	0.86	102.10-106.80	99.68	1.24	97.30-106.80	4.28
HI	0.25	0.02	0.20-0.30	0.29	0.03	0.20-0.40	0.36	0.01	0.30-0.50	-26.87
PHT	62.63	6.21	47.60-78.80	63.50	4.11	47.60-73.60	49.49	3.35	40.80-58.80	22.09
SPN	16.38	0.47	14.70-17.70	16.02	0.40	15.10-17.30	15.77	0.49	14.40-17.40	1.54
BM	7818	318.32	6297.00-8793.00	4311	55.34	4165.00-4458.00	3866	168.77	3279.00-4290.00	10.30
TW	77.95	1.37	72.30-82.00	65.61	1.38	60.30-69.70	69.31	1.80	62.50-78.20	-5.64
NDVI	0.67	0.01	0.60-0.70	0.40	0.02	0.36-0.45	0.27	0.01	0.25-0.31	33.07
GA	0.61	0.02	0.56-0.66	0.28	0.03	0.20-0.38	0.12	0.01	0.08-0.18	56.53
KN	6276.00	677.54	4372.00-8120.00	7272.00	449.17	6102.00-9027.00	5315.00	321.83	459.00-6641.00	26.90
SL	9.25	0.70	7.20-11.30	8.82	0.75	6.80-11.00	8.65	0.71	6.40-10.60	1.88
KNS	35.22	1.91	29.20-41.00	38.79	0.66	37.10-40.50	38.63	1.98	32.90-45.80	0.40

†GRW10 (285), Greeley wet 10; GRW11 (288), Greeley wet 11; GRD11(288), Greeley dry 11; W, grain yield under wet; D, grain yield under dry; numbers in parenthesis stand for number of lines evaluated at each environment. ‡ Trait description is as given in Table 2.1.

Table 2.2. Continued.

	Environments†									
	GRW10 (285)			GRW11 (288)			GRD11(288)			
Trait‡	Mean	Standard deviation	Range	Mean (W)	Standard deviation	Range	Mean (D)	Standard deviation	Range	% reduction (D/WX100)
SKW	35.60	2.29	30.90-41.40	25.74	1.79	21.50-30.90	27.10	2.05	22.40-32.30	-5.33
SKD	2.87	0.07	2.70-3.10	2.53	0.07	2.40-2.80	2.56	0.08	2.34-2.80	-1.18
SKH	72.17	9.79	30.90-88.20	67.65	11.84	17.20-83.00	68.15	11.75	21.54	-0.69
KWS	1.25	0.07	1.10-1.40	0.84	0.02	0.80-0.90	0.97	0.03	0.89	-15.38
GFD	36.06	0.98	32.50-40.20	34.10	0.47	32.90-36.00	31.35	0.68	29.72	8.11
LA	18.25	2.43	12.80-27.60	15.27	0.99	12.80-19.10	11.47	0.86	8.97	24.90
KNL	2.14	0.11	1.77-2.46	2.42	0.05	2.30-2.60	2.45	0.08	2.18	-1.12
LS	5.40	0.83	3.30-8.20	NA	NA	NA	7.71	0.42	6.49	NA

†GRW10, Greeley wet 10; GRW11, Greeley wet 11; GRD11, Greeley dry 11; W, grain yield under wet; D, grain yield under dry; Numbers in Parenthesis stand for number of lines evaluated at each environment.
NA=data not available
‡ Trait description is as given in Table 2.1.

Table 2.3. Mean values of the WAMII spring wheat association mapping panel for traits measured under rainfed (D) and non-stressed (W) conditions at Melkassa, Ethiopia in 2011.

	Environments†						
	MLKW11 (W)			MLKD11 (D)			Drought reduction (D/Wx100)
Trait‡	Mean	Standard deviation	Range	Mean	Standard deviation	Range	
YLD	4378.00	546.98	2592.00-5507.00	3862.00	454.33	2444.00-5377.00	11.79
TKW	25.52	3.37	16.28-34.73	22.87	2.14	16.66-30.39	10.38
DH	55.05	2.86	49.78-68.03	55.83	3.70	47.91-69.27	-1.43
LL	20.21	0.60	18.58-21.83	22.17	0.87	19.92-24.89	-9.70
LW	1.29	0.03	1.15-1.38	1.34	0.04	1.25-1.66	-3.88
DM	93.02	2.75	87.74-99.13	95.09	1.69	89.86-99.7	-2.23
HI	0.27	0.03	0.15-0.38	0.23	0.02	0.15-0.30	14.81
PHT	81.84	2.44	74.74-87.98	NA	NA	NA	NA
SPN	16.55	0.97	13.88-20.09	NA	NA	NA	NA
BM	16231.00	642.36	14478.00-18639.00	17160.00	600.22	15511.00-18967.00	-5.72
KN	17442.00	2179	10329.00-21940.00	16993.00	1460.00	13541.00-21869.00	-2.57
GA	0.77	0.05	0.58-0.88	0.84	0.023	0.75-0.89	-9.09
GFD	37.97	1.38	32.85-42.47	39.26	1.74	33.63-44.91	-3.40
LA (cm^2)	19.65	1.07	16.50-22.65	22.42	1.37	19.54-31.42	-14.10

†MLKW11 (W), Melkassa wet 11; Melkassa dry 11 (D); W, grain yield under wet; D, grain yield.
NA=data not available or analysis not possible.
‡ Trait description is as given in Table 2.1.

The effect of water deficit on the phenotypic traits was assessed based on the ratio of mean performance of the genotypes under rainfed conditions to their performance under full irrigated conditions, expressed in percentage units (Table 2.2). Of all traits evaluated at Greeley in 2011, green leaf area (57%) and NDVI (33%) were affected the most by moisture stress, followed by spike number m^{-2} (29%), kernel number m^{-2} (27%) and plant height (22%). Grain yield was reduced by 14%, which is about 200 kg ha^{-1}. However, traits such as thousand kernel weight, harvest index, test weight, kernel weight $spike^{-1}$ and single kernel weight showed from 5% to 26% increase under moisture stress conditions at Greeley. Harvest index was the trait most affected by moisture stress conditions during the early growth stage at Melkassa (Table 2.3)

2.2.2 Genotypic correlations

The genotypic correlation coefficients of grain yield and NDVI with other measured traits are presented in Tables 2.4 and 2.5, respectively. Genotypic correlation coefficients among all phenotypic traits measured in five environments are given in Table A.5 through Table A.9. Grain yield showed consistently high and positive genotypic correlation with kernel number m^{-2}, test weight and final biomass regardless of the moisture level. Grain yield had high and positive genotypic correlations (r_g=0.73 to 0.91) with harvest index except at Greeley under moisture stress condition (r_g=0.118). The genotypic correlation for grain yield and thousand kernel weight was positive but weak in all Greeley environments (r_g=0.11 to 0.22), but much larger in the two Melkassa environments (r_g=0.61 and 0.77). Grain yield was negatively associated with days to heading in all environments. The genotypic correlation of yield with NDVI was weak under fully irrigated conditions (r_g=0.13 and 0.16 for GRW10 and GRW11, respectively), but considerably higher (r_g=0.42) under moisture stress conditions (GRD11).

Genotypic correlation values of NDVI with leaf area, days to heading and days to maturity were positive and significant in all environments where NDVI data were collected (Table 2.5). Plant height also had a strong positive genotypic correlation with NDVI in two of the three Greeley environments. The genotypic correlation of NDVI with leaf green area (GA) was consistently positive and high under all moisture conditions. On the other hand, NDVI was negatively correlated with leaf senescence both under irrigated and moisture stress conditions. The genotypic correlation values of NDVI with final biomass was low (0.15 to 0.36) in this study.

2.2.3 Heritability estimates of agronomic traits

The estimated heritability values for each measured phenotypic trait in each environment and combined across environments are presented in Table 2.6. High (>75%) heritability estimates were obtained for days to heading, single kernel weight, test weight, single kernel diameter and kernel hardness. Most yield component traits showed low and inconsistent heritability estimates in different environments. Moderate (>50%) heritability estimates were obtained for yield, green leaf area, grain filling duration and flag leaf area in four out of six environments. Above ground biomass, NDVI, kernel weight spike^{-1}, kernel number spikelet^{-1}, spike number m^{-2}, and kernel number spike^{-1} had low heritability estimates in individual environments despite heritability values over 50% obtained for above ground biomass, number of spikes m^{-2} and number of kernel m^{-2} for combined data across environments.

Table 2.4. Genotypic correlation coefficients between grain yield and other measured traits in the WAMII spring wheat association mapping panel grown in five environments.

	Environments†				
Trait‡	GRW10	GRW11	GRD11	MLKW11	MLKD11
Kernel weight	0.22**	0.11ns	0.15**	0.77**	0.61**
Harvest index	0.75**	0.73**	0.12*	0.91**	0.90**
Kernel number	0.77**	0.99**	0.71**	0.99**	0.74**
Spike length	-0.11ns	0.01ns	0.20**	0.12*	0.37**
Spikelet number	-0.15*	-0.45**	-0.12*	-0.24**	NA
Biomass	0.47**	0.84**	0.63**	0.50**	0.39**
Kernel number spike^{-1}	0.47**	0.23**	0.10ns	NA	NA
Kernel weight spike^{-1}	0.58**	0.25**	0.35**	NA	NA
Spike number	0.19**	0.99**	0.75**	NA	NA
Kernel spikelet^{-1}	0.55**	0.56**	0.27**	NA	NA
Test weight	0.37**	0.48**	0.38**	NA	NA
Single kernel weight	0.23**	0.29**	0.12*	NA	NA
Hardness index	0.01ns	-0.21**	-0.12*	NA	NA
Days to heading	-0.32**	-0.26**	-0.04 ns	-0.54**	-0.68**
Days to maturity	-0.21**	-0.30**	-0.06ns	-0.44**	-0.67**
Grain filling duration	0.12ns	0.01ns	-0.03ns	0.13*	0.55**
Leaf length	-0.11ns	-0.14*	0.10ns	-0.48**	0.06ns
Leaf width	-0.18ns	-0.07ns	-0.06ns	0.01ns	0.03ns
Leaf area	-0.20**	-0.09ns	0.06ns	-0.27**	0.03 ns
Plant height	-0.05ns	-0.41**	0.45**	0.35**	NA
NDVI	-0.13ns	0.16**	0.42**	NA	NA

†GRW10, Greeley wet 11; GRW11, Greeley wet 11; GRD11, Greeley dry 11; MLKW11, Melkassa wet 11; MLKD11,Melkassa wet 11;
‡Traits description is as given in Table 2.1
ns= non-significant at p<0.05; *=significant at P<0.05; **=significant at P<0.01; NA= data not available or analysis not possible.

Table 2.4 Continued.

	Environments†				
Trait‡	GRW10	GRW11	GRD11	MLKW11	MLKD11
Green Leaf area (GA)	-0.08ns	0.267**	0.34**	0.24**	-0.22**
Leaf senescence (LS)	0.20**	NA	-0.21**	NA	NA

†GRW10, Greeley wet 11; GRW11, Greeley wet 11; GRD11, Greeley dry 11; MLKW11, Melkassa wet 11; MLKD11,Melkassa wet 11;
‡Traits description is as given in Table 2.1
ns= non-significant at p<0.05; *=significant at P<0.05; **=significant at P<0.01; NA= data not available or analysis not possible.

Table 2.5. Genotypic correlation coefficients between NDVI measured after heading and phenological and morphological traits of the WAMII spring wheat association mapping panel.

	Environments†		
Trait‡	GRW10	GRW11	GRD11
Days to heading	0.27**	0.50**	0.49**
Day to maturity	0.37**	0.39**	0.71**
Grain filling duration	0.14*	-0.28**	0.42**
Leaf senescence	-0.84**	NA	-0.74**
Green leaf area	0.99**	0.57**	0.62**
Flag leaf length	0.65**	0.31**	0.39**
Flag leaf width	-0.09ns	0.14*	0.02ns
Flag leaf area	0.39**	0.27**	0.27**
Plant height	0.70**	-0.18**	0.96**
Biomass	0.36**	0.16*	0.15*

†GRW10, Greeley wet 10; GRW11, Greeley wet 11; GRD11, Greeley dry 11.
ns, non-significant; *, significant at P<0.05; **, significant at P<0.01.

Table 2.6. Heritability estimates of agronomic and morphological traits in the WAMII spring wheat association mapping panel grown in five environments.

	Environments†					
Trait‡	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
Yield	0.60	0.43	0.40	0.68	0.61	0.61
Thousand kernel weight	0.74	0.31	0.51	0.80	0.65	0.81
Harvest index	0.38	0.58	0.16	0.68	0.57	0.45
Kernel number	0.55	0.31	0.32	0.68	0.48	0.45
Spike length	NA	0.83	0.60	0.74	0.04	NA
Spikelet number	0.45	0.44	0.54	0.73	NA	0.57
Biomass	0.32	0.07	0.23	0.40	0.32	0.48
Kernel weight per spike	0.28	0.11	0.19	NA	NA	0.54
Kernel number per spikelet	0.26	0.14	0.28	NA	NA	0.40
Spike number	0.15	0.01	0.28	NA	NA	0.31
Kernel number per spike	0.26	0.11	0.38	NA	NA	0.45
Test weight	0.77	0.65	0.84	NA	NA	0.75
Single kernel diameter	0.12	0.79	0.78	NA	NA	0.86
Single kernel weight	0.75	0.77	0.75	NA	NA	0.85
Single kernel hardness	0.95	0.93	0.96	NA	NA	0.97
Days to heading	0.83	0.72	0.69	0.89	0.93	0.75
Days to maturity	0.82	0.66	0.56	0.83	0.67	0.69

† GRW10, Greeley wet 10; GRW11, Greeley wet 11; GRD11, Greeley Dry 11; MLKW11, Melkassa wet 11; MLKD11, Melkassa dry 11; Combined, combined data across five data sets.
NA=data not available or analysis not possible.
‡Trait description is as given in Table 2.1.

Table 2.6 Continued.

	Environments†					
Trait‡	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
Grain filling duration	0.67	0.37	0.38	0.58	0.63	0.55
Leaf length	0.66	0.45	0.61	0.28	0.46	0.72
Leaf width	0.75	0.42	0.38	0.27	0.23	0.62
Leaf area (LA)	0.70	0.47	0.55	NA	NA	0.67
Plant height (PHT)	0.83	0.78	0.76	0.43	NA	0.83
Normalized vegetation index (NDVI)	0.28	0.49	0.32	NA	NA	0.62
Green leaf area (GA)	0.31	0.63	0.52	0.63	0.41	0.55
Leaf senescence (LS)	0.76	0.44	NA	NA	NA	0.73

† GRW10, Greeley wet 10; GRW11, Greeley wet 11; GRD11, Greeley Dry 11; MLKW11, Melkassa wet 11; MLKD11, Melkassa dry 11; Combined, combined data across five data sets.
NA=data not available or analysis not possible.
‡Trait description is as given in Table 2.1.

2.2.4 Model-based population structure and linkage disequilibrium

Population structure analysis of 287 spring wheat lines conducted with the STRUCTURE program indicated the likely number of subpopulations was seven based on change of k (Figure 2.1). Of these, subpopulations II, IV, V and VI were dominated by the lines with Kauz, Pastor, TUI and WBLL1 background, respectively (Figure 2.2). Lines with different backgrounds were grouped together for subpopulations I, III and VII. There was evidence for the presence of population structure from the cluster analysis based on genetic distance using the Ward method (Figure A. 1). Molecular variance analysis for the seven model-based populations indicated that 78.5% of the total variation is explained by within-population variation, whereas 21.5% of the variation is due to among-population variation (Table 2.7). Population differentiation (Fst)

values ranged from 0.14 to 0.73 and were highly significant (P<0.0001) for all pairs, supporting the presence of population structure. There were highly significant (P<0.0001) differences among the populations.

Linkage disequilibrium among markers was calculated for all chromosomes (except chromosomes 4D and 5D that were represented by only a single marker each). A critical value of $r^2 > 0.264$ was determined to be the appropriate threshold for LD due to physical linkage. Chromosomes 4A (62%) and 1B (55%) showed a higher percentage of significant (P<0.01) marker pairs in LD whereas chromosomes 5A (20%), 2B (23%) and 7A (23%) had the least number of significant (P<0.01) marker pairs (Figure 2.3; Table A.10). The percentage of marker pairs due to physical linkage was high for chromosome 3D (24%) followed by chromosomes 2D (17%) and 1B (16%). The percentage of LD due to physical linkage mimics the percentage of LD at $r^2 > 0.2$ for all chromosomes, but had no similarity with that of LD at P <0.01. Marker pairs at $r^2 > 0.2$ and $r^2 > 0.264$ were significant at $P < 0.001$ for all 19 chromosomes considered in this study.

Linkage disequilibrium decay rate evaluation was conducted at the genome and individual chromosome level. The genome level LD decayed below r^2=0.2 at about 1.7 cM for the A genome (Figure 2.4), while the smoothing curve crossed the r^2=0.2 line at approximately 2 cM for the B genome (Figure 2.5). For the D genome, the curve crossed the r^2=0.2 line near 6.8 cM genetic distance (Figure 2.6). For all 19 chromosomes, the LD decay curve crossed the r^2=0.2 line at about 3.4 cM (Figure 2.7). We were able to determine the genetic distance at the baseline r^2=0.2 for four out of the seven model-based subgroups for all chromosomes together; LD decayed below r^2=0.2 within 8-9 cM for three of them and within 6 cM for the fourth sub-population.

Multiple regression analysis with population structure in the model showed that plant height (29.5%) was the trait most affected by the genotype groupings, followed by single kernel weight (25.9%), thousand kernel weight (21.0%), single kernel diameter (20.5%) and NDVI (18.9%) (Table 2. 8). On the other hand, the variations explained due to population structure were non-significant for kernel number m^{-2}, drought susceptibility index, kernel hardness, harvest index and days to maturity. Moreover, population structure explained only about 5% of the total variation in days to heading, test weight and kernel weight per spike. The variation explained due to population structure in grain yield (7.5%) was also low.

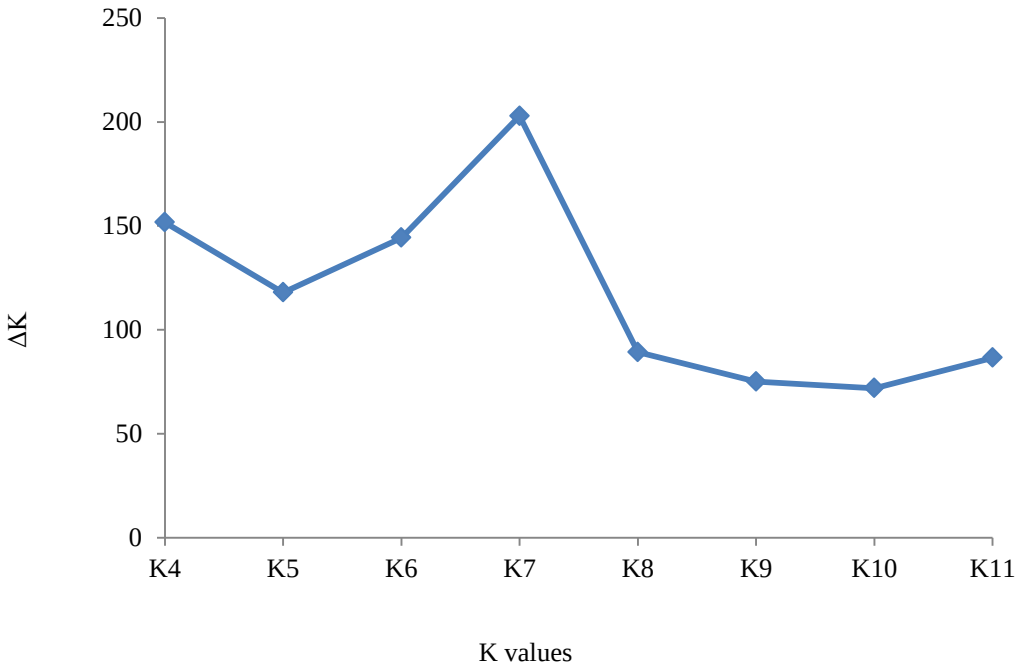

Figure 2.1. Change of k values between k=3 and k=12 for 287 spring wheat lines of the association mapping panel based on STRUCTURE software.

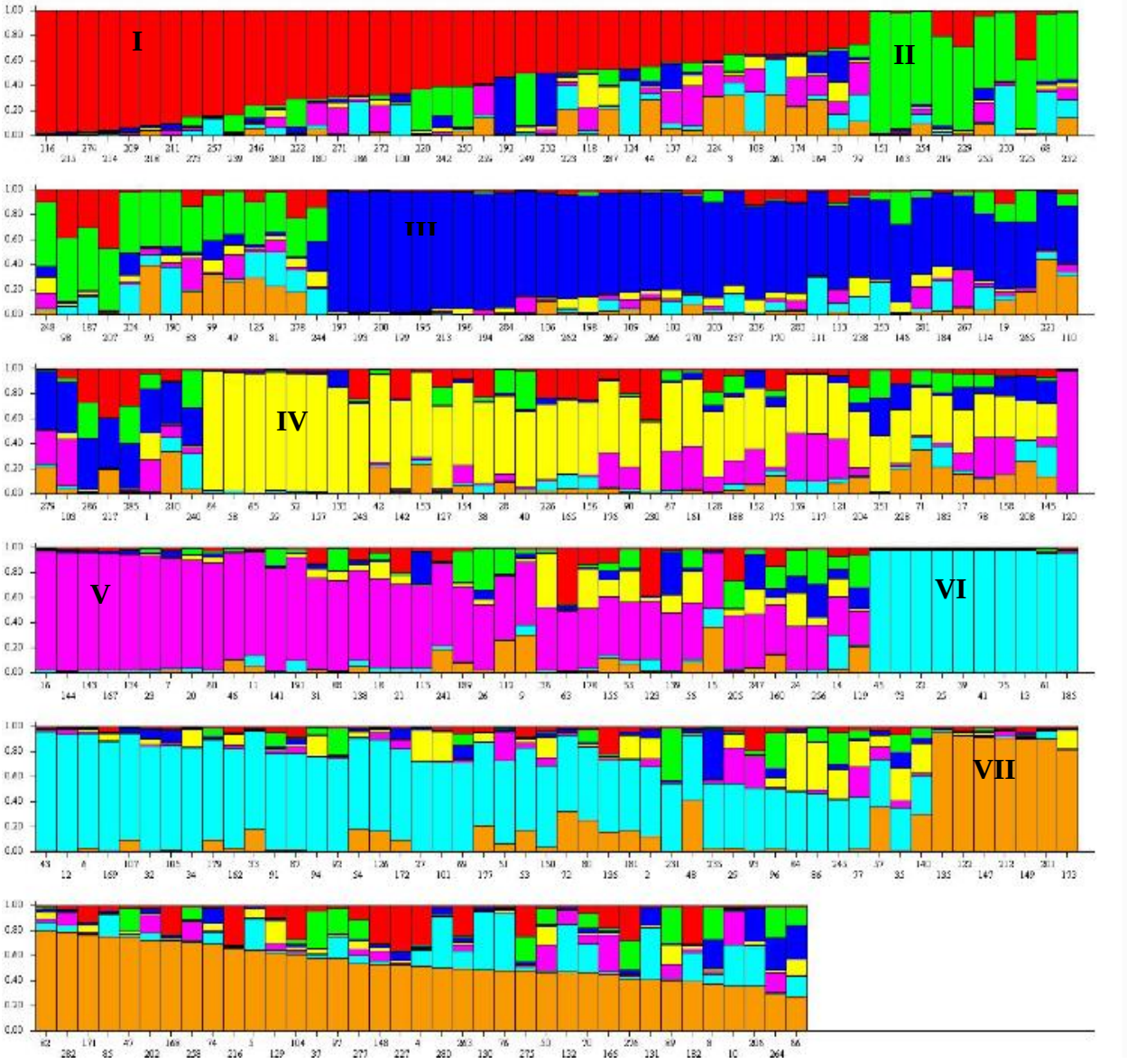

Figure 2.2. Population structure for 287 genotypes in a spring wheat association mapping panel based on 78 DArT markers.

Table 2.7. Variability among and within seven clusters of the spring wheat association mapping panel based on 78 DArT markers.

Source of variation	DF	Sums of squares	Variance components	Percentage of variation
Among populations	6	620.60	2.36	21.51
Within populations	280	2408.81	8.60	78.49
Total	286	3029.41	10.96	

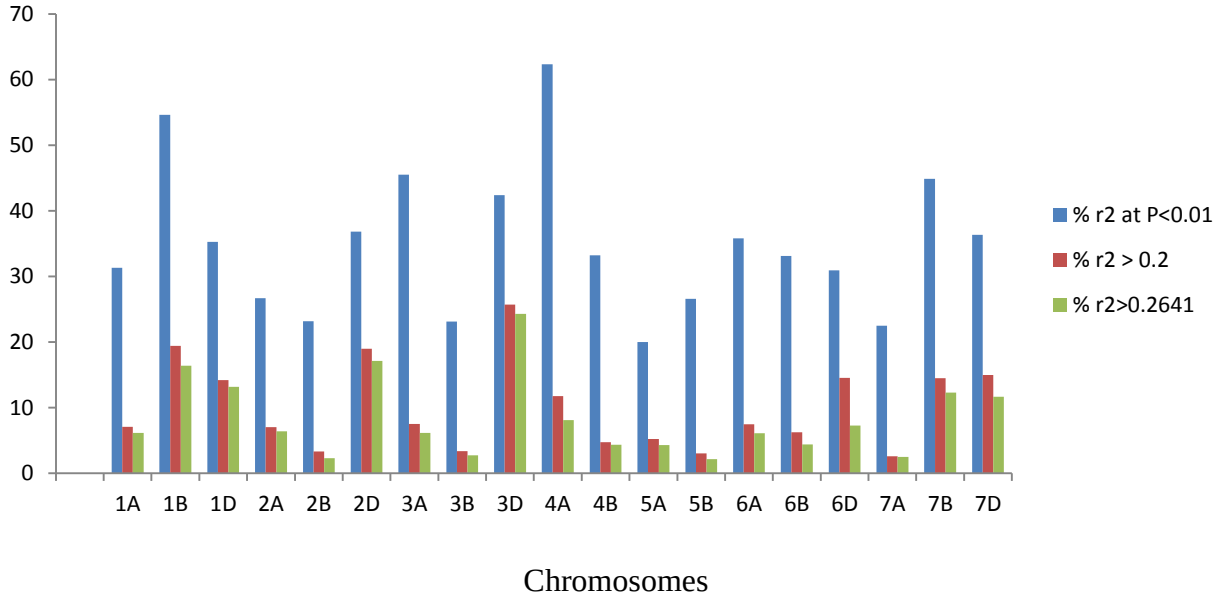

Figure 2.3. Percentage of significant linkage disequilibrium at r^2 >0.2641, r^2> 0.2 and r^2 at P<0.01 for 19 hexaploid wheat chromosomes in 287 lines of the spring wheat association mapping panel.

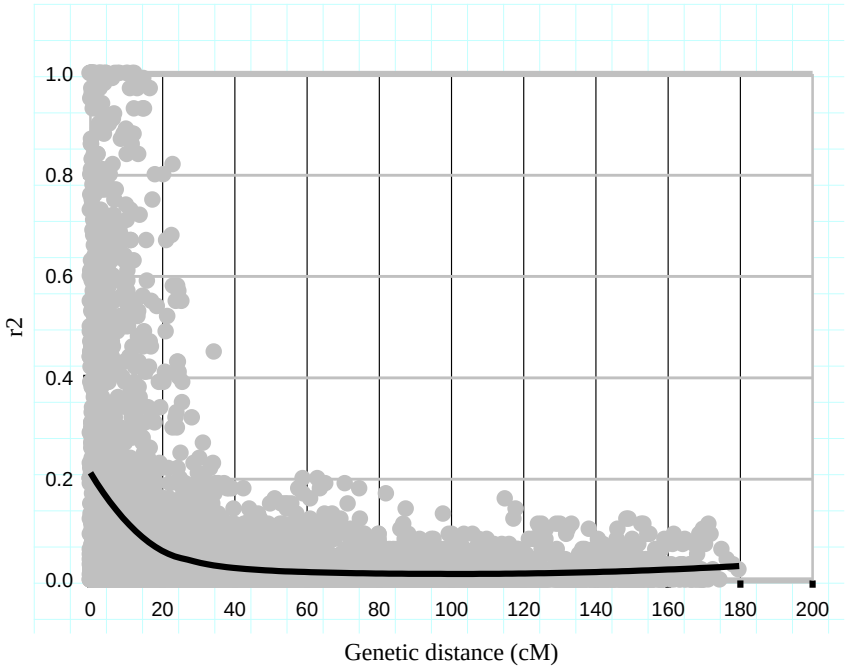

Figure 2.4. Linkage disequilibrium (r^2) plot of all chromosomes of the A genome in 287 lines of a spring wheat association mapping panel.

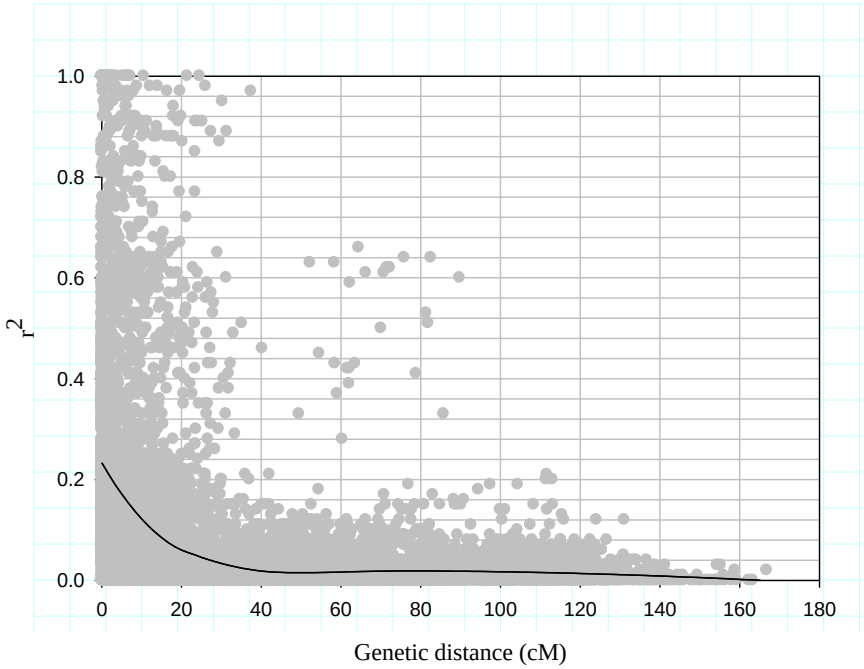

Figure 2.5. Linkage disequilibrium (r^2) plot of all chromosomes of the B genome in 287 lines of a spring wheat association mapping panel.

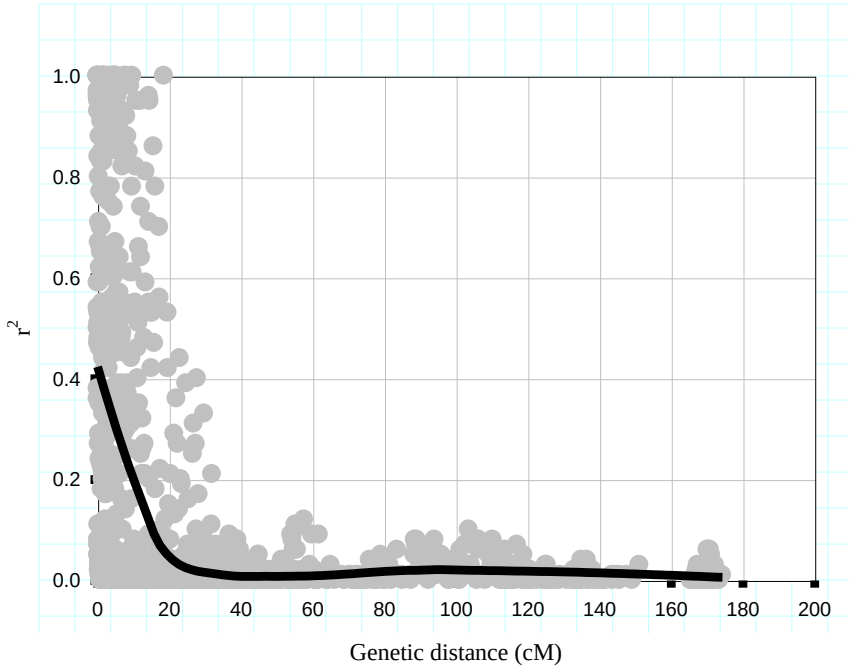

Figure 2.6. Linkage disequilibrium (r^2) plot of all chromosomes on the D genome in 287 lines of a spring wheat association mapping panel.

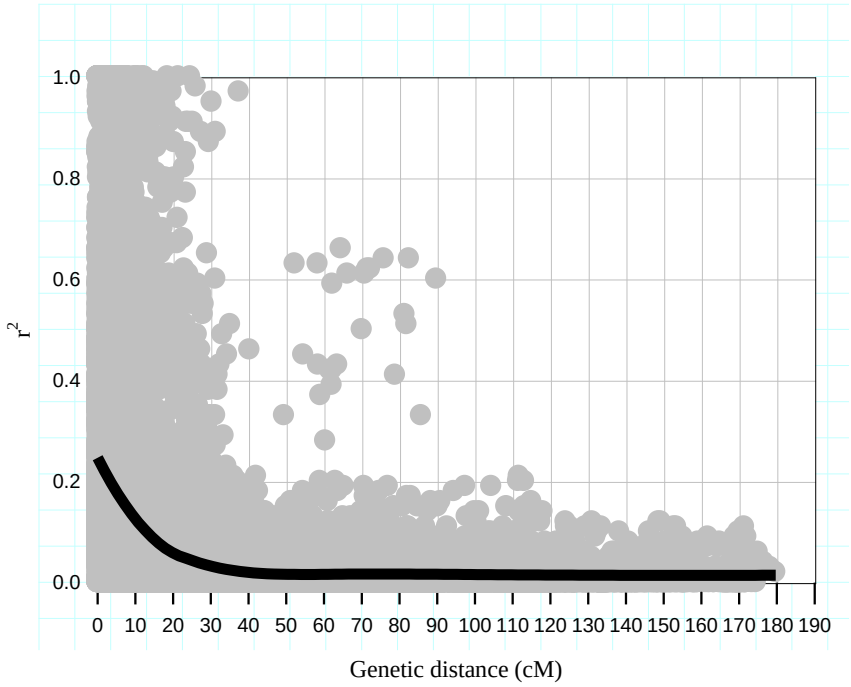

Figure 2.7. Linkage disequilibrium (r^2) plot for 19 chromosomes of 287 lines of a spring wheat association mapping panel.

Table 2.8. Percent of phenotypic variation explained (R^2) by population structure based on combined data across environments.

Trait	Environments †	R^2 (%)	*P*-value
Thousand kernel weight	5	21.0	0.0001
Days to heading	5	4.7	0.0362
Flag leaf length	5	11.4	0.0001
Flag leaf width	5	9.4	0.0001
Days to maturity	5	3.0	0.1970
Harvest index	5	2.2	0.3950
Plant height	4	29.5	0.0001
Grain yield	5	7.4	0.0015
Kernel number	5	3.5	0.1182
Spikelet number	3	13.0	0.0001
Biomass	5	12.4	0.0001
NDVI	3	18.9	0.0001
Leaf green area	5	14.4	0.0001
Test weight	3	5.0	0.0212
Kernel weight per spike	3	5.3	0.0200
Grain filling duration	5	9.1	0.0001
Flag leaf area	5	13.2	0.0001
Kernel hardness	3	1.9	0.5060
Single kernel weight	3	25.9	0.0001
Kernel number per spikelet	3	9.0	0.0002
Spike number per m^2	3	5.6	0.0140
Single kernel diameter	3	20.5	0.0001
Kernel number per spike	3	15.7	0.0001

† Number of environments used in the combined data analysis.

2.2.5 Marker-trait associations (MTA)

Although MTA were detected at $P<0.05$ for all traits, we are reporting only strong MTA ($P<0.001$) for single environments and moderate MTA significant ($P<0.01$) in at least half of the test environments. Consistency across environments was used as an additional criterion for MTA significant at $P<0.01$ in order to reduce the risk of including false MTA. A summary of MTA in different environments for each phenotypic trait is given in Table 2.9.

Considering both criteria together ($P<0.001$ and $P<0.01$ in half or more of the environments), a total of 565 MTA was detected in one or more environments for 26 measured or calculated phenotypic traits in five environments plus combined data across environments. Out of these, about 20% of the MTA were detected only in a single environment and the remaining 80% were observed in two or more environments. A total of 130 (22.9%) of MTA involved unmapped markers. The numbers of MTA detected for grain yield both under irrigated and rainfed treatments at Greeley were similar, while at Melkassa the number of MTA detected for grain yield under the stressed treatment was lower than the number of MTA detected under the non-stressed treatment for grain yield (Table 2. 9). The highest number of MTA was recorded for kernel hardness (113) followed by test weight (44) and flag leaf length (39) while the fewest MTA were obtained for drought susceptibility index, flag leaf senescence, kernel number per spikelet, kernel number per spike and spikelet number per spike. Moreover, kernel hardness had the largest number of stable MTA (15) followed by test weight (9). Chromosome-wise, the highest number of MTA was detected on chromosomes 5B, 3B, 7A and 1B, while chromosomes 1A, 2A, 2D, 3A and 5A harbored the fewest MTA in this study. No MTA were detected for chromosomes 4D and 6D (Figure 2.9).

Grain yield MTA were detected on chromosomes 1BS, 2DS, 5B (73 and 76.4 cM) and 7B. Unmapped marker *wpt0419* was also associated with grain yield at Melkassa under non-stressed conditions. The marker *wpt6531* on chromosome 2DS was associated with yield in four out of six environments including both irrigated and rainfed conditions, and can be considered a stable marker for grain yield. However, *wpt3457* (5B) showed the strongest association with yield under rainfed condition at Greeley in 2011 (Figure 2.8; Table 2.10).

Stable MTA were also detected for the major yield component traits kernel number per m^2 on chromosome 7AS and harvest index on chromosomes 5AL, 5B (72.4 cM) and *wpt0286* (unmapped). Regions of chromosomes 1BL, 3BS, 4A, 5B (72.4 cM) and 5BL were also consistently associated with thousand kernel weight. Regions of chromosome 4B, 5B, 6B, 7AS and 7AL were associated with spike number per m^2 at two environments. However, all MTA obtained for final biomass on chromosomes 1AS, 5B (72.4 cM), 7BL, 7D, 7DL and 7DS were environment-specific (Table 2.10). Similarly, all MTA detected for number of spikelets per spike (2B and 7B), kernel number per spikelet (1DL, 7A and 7BL) and kernel number per spike (1AS, 3BS and 7A) were detected only at single environments despite the presence of very strong associations for some MTA (Table 2.10).

Single kernel traits such as single kernel weight and diameter and kernel hardness had more stable MTA than most of the yield component traits. The MTA for single kernel weight were distributed on chromosomes 1BL, 1D, 4A, 2AL, 4BL and 5BL while MTA of single kernel diameter were detected on chromosomes 1BL, 2D, 3AS, 3B, 3D, 4AL, 6BS, 7BL and 7DL. Several MTA were obtained for kernel hardness and the most stable ones (those detected in all environments) were found on chromosomes 1BL, 1D, 3AS, 3D, 4AL and 7A. Similarly, many

stable MTA were obtained for test weight, with chromosomes 2DL, 3BS, 4A, 4BL and 7BL comprising the location of MTA detected in three out of the total four environments.

The most stable MTA for days to heading was detected on chromosome 1DS (four out of six environments), followed by MTA on chromosomes 2B, 3AL, 3B and 4BL (three out of six environments each). The most significant MTA ($P<0.001$) was detected on chromosome 2AL for marker *wpt9277* at GRW10. This same marker was consistently associated with days to maturity. Grain filling duration had stable MTA on chromosomes 1BL, 3BS and 7AL (each showing up in three of six environments).

Marker-trait associations were found for plant height on chromosomes 3BL, 5BS, 6AS, 7AS and 7BL, of which the MTA on chromosomes 6AS and 7BL were the most consistent. QTL regions for flag leaf length were noted on chromosomes 1BS, 1BL, 2BL, 3BL, 3AL and 5B. Most of these associations were consistent, particularly marker *wpt5072* on chromosome 3BL which was detected in five out of six environments. For flag leaf width, however, only *wpt667461* (unmapped) was consistently associated with the trait despite the presence of flag leaf width associated markers on chromosomes 2DL, 3BL, 5BS, 6A and 7AS. Both stable and environment-specific MTA were detected for flag leaf area; the chromosomes 3BL and 5BL harbored stable QTL for this trait. Unmapped markers *wpt0605* and wpt1370 were also consistently associated with flag leaf area.

Significant MTA were also obtained for drought tolerance-related traits and vegetation indices. Regions of chromosome 4AL, 7A and 7BL comprised QTL for drought susceptibility index. Leaf senescence QTL were found in three regions of chromosome 6B (36.1, 50.6 and 84.6 cM) and another five unmapped markers also showed associations with leaf senescence. Regions

of chromosomes 1AL, 1BS, 2AS and 6BL harbored QTL for NDVI, and unmapped marker wpt0694 was also associated with NDVI in two environments.

Some of the MTA were significant at FDR=0.05 after correcting for multiple comparisons. These significant MTA at FDR=0.05 were obtained for spikelet number spike^{-1} on chromosome 2BS, plant height on chromosome 6AS, grain filling duration on chromosome 3BS and green leaf area on chromosome 1BL. Associations of unmapped markers with flag leaf width (*wpt730263*), spike number per spike (*wpt666595* and *wpt667101*) and drought susceptibility index (*wpt0419*) were also significant at FDR 0.05 (Table 2.11). None of the MTA obtained for the remaining traits survived the FDR adjustment for multiple testing. At a relaxed FDR of 0.25, however, MTA were identified for test weight, biomass, leaf green area, harvest index, leaf length, leaf width, single kernel diameter, kernel hardness, flag leaf area, kernel per spikelet and kernel number- based drought susceptibility index (data not shown).

Multi-trait MTA were detected in many chromosome regions. Their chromosome positions are shown with other trait-specific QTL in Figure 2.10. Clusters of QTL were detected for kernel size-related traits on chromosomes 1BL, 4AL and 7DL. Kernel quality traits (SKH and TW) had QTL in common with one or more kernel size-related traits on chromosomes 1D, 2DL, 3BS, 3D, 4AL, 5B and 7AS. Markers near the centromeric region of chromosome 5B (67.7-76.4 cM) were associated with yield, spike number per m^2, biomass, plant height, harvest index, thousand kernel weight and test weight. A region of chromosome 1AL was associated with both harvest index and NDVI, and QTL for green leaf area was detected close to the region of a harvest index QTL on chromosome 5AL. The QTL on 1BL for leaf green area was in the same region with a QTL detected for SKD, while green leaf area QTL on 3BL was close to the QTL region for TKW, TW and GFD. Similarly, yield and harvest index had QTL in common on

chromosome 1BS. Flag leaf area and flag leaf length had QTL in common on chromosomes 3BL and 5BL as well.

Multi-trait QTL were also detected among drought tolerance-related traits and vegetation indices. Marker-trait associations were obtained on chromosome 6BL for NDVI, leaf senescence and leaf green area index. However, only MTA for leaf senescence and leaf green area were roughly in the same region (within 3 cM). Among QTL detected for drought susceptibility index, QTL on chromosome 4AL was in the same region with the QTL for single kernel diameter, single kernel weight and thousand kernel weight. Similarly, a QTL on 7A was detected in the same region with QTL detected for kernel number, kernel number per spikelet and kernel number per spike. Drought susceptibility QTL on 7BL was detected at a distance of 3.4 and 3.9 cM away from plant height and spike number QTL, respectively.

Table 2.9. Summary of marker-trait associations detected for agronomic traits and drought related indices detected in five environments.

	Environments†						
Trait‡	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined	Total§
BM	0	1	1	4	0	0	6
DH	4	3	3	3	3	4	20
DM	1	1	2	1	0	1	6
DSI_KN	NA	NA	NA	NA	NA	NA	2
DS_YLD	NA	NA	NA	NA	NA	nA	1
GA	8	7	8	0	0	9	32
GFD	9	3	2	3	2	4	23
HI	0	5	1	5	4	2	17
SKH	23	30	21	5	0	34	113
KN	1	4	2	1	3	4	15
KNL	1	0	4	0	0	0	5
KNS	3	0	2	0	0	0	5
KWS	5	1	0	0	0	5	11
LA	6	3	6	0	3	7	25
LL	8	5	4	5	6	11	39
LS	1	2	1	0	0	0	4
LW	3	1	2	1	6	2	15
NDVI	2	2	1	0	0	5	10
PHT	7	5	3	8	0	9	32
SKD	6	8	9	0	0	15	38
SKW	8	8	6	0	0	16	38
SL	1	4	3	0	0	0	8
SN	4	1	6	0	0	7	18
SPN	0	-	4	1	0	0	5
TKW	1	5	3	3	4	5	21
TW	6	8	15	0	0	15	44
YLD	0	3	3	4	2	3	15

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa non-stressed 2011; MLKD11, Melkassa stressed 2011; Combined, combined data across environments;
‡ Trait description is as given in Table 2.1.
§ Total number of marker-trait associations detected for traits across environments.

Table 2.10. Marker-trait associations detected in five environments and combined across environments for agronomic traits.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
BM	*wPt1405*	7DL	170.7				0.000166		
	wPt1548	5B	72.4			0.000704			
	wPt4177	1AS	19.8				6.99E-05		
	wPt4300	7BL	210.9		0.000677				
	wPt745106	7DS	1.1				0.000132		
	wPt8422	7DL	170.6				0.000904		
DH	*wPt10142*	3B	75.1	0.000437					
	wPt10991	3B	30.2				0.00314	0.002824	0.004188
	wPt4199	2B	81.6				0.009216	0.006652	0.007951
	wPt5996	4BL	104.9	0.004347	0.007476	0.000269			
	wPt6979	1DS	19.5	0.002156	0.002914	0.000806			0.009662

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; MLKD11; Melkassa dry 2011; Combined, combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt732035	Unknown			0.000938				
	wPt733104	7DS	1.4			0.000669			
	wPt9277	2AL	109.4	9.99E-05					
	wPt9422	3AL	166.7				0.006001	0.006033	0.003054
DM	*wPt3728*	2D	90.5			0.00056			
	wPt6013	7AL	145.2		0.000875				
	wPt669696	Unknown					0.000616		
	wPt9277	2AL	109.4	0.003722		0.005515			0.003878
DSI_KN	*tPt1755*	7A	79			0.000353			
	wPt5069	7BL	224.9			0.00077			

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
DSI_yld	*wPt5003*	4AL	87.9					0.000528	
GA	*rPt9074*	1BL	88.3	0.002971	0.000569	0.00514			7.44E-06
	wPt0944	Unknown		0.006707	0.007166				3.32E-05
	wPt0950	Unknown		0.001249		0.004379			0.006217
	wPt3728	2D	90.5		0.001648				0.007637
	wPt5270	6BL	87.7	0.002306					0.000539
	wPt5374	2BS	37.9	0.008236		0.006542			0.005234
	wPt664378	Unknown		0.000355		0.005287			
	wPt665030	Unknown			0.001154				0.009245
	wPt667089	Unknown		0.000355		0.005287			

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt669484	1AS	2.7			0.00049			
	wPt6945	3B	56.5		0.006373				0.007626
	wPt7225	3BS	12.9		0.00736	0.009469			
	wPt734145	5AL	122.9	0.005019		0.008303			
	wPt9094	5AL	46.2		0.006475				0.00646
GFD	*wPt0959*	Unknown					0.002797	0.001897	0.008037
	wPt3226	7AL	158.4	0.004059			0.009257	0.005011	
	wPt5836	3BS	39.1	3.41E-05		0.007319			0.000988
	wPt8168	1BL	41.1	0.000307	0.000871				0.000284
	tPt8942	5BS	35.1			0.000218			

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	tPt9267	3BS	24.6	2.05E-05					
	wPt10006	3BS	20.8	0.000843					
	wPt3566	1BL	45	0.000649					
	wPt665725	1AL	76.6				0.000613		
	wPt740564	3D	39.4		0.000115				
	wPt741750	3BS	25.1	5.68E-05					
	wPt742337	3BS	14.2	0.000682					
HI	*rPt3825*	Unknown						0.000148	
	wPt0286	Unknown					0.000836	0.001344	0.000208
	wPt0419	Unknown					0.00072		

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt1911	1BS	21.6				0.000778		
	wPt3509	5AL	42.3				0.00191	0.002286	0.000159
	wPt6105	5B	72.4			0.008134	0.008043	0.009944	
	wPt742925	5A	40.6		0.000156				
	wPt744567	5A	40.9		0.000141				
	wPt7769	5AL	41.2		0.0002				
	wPt8347	1AL	63.6		0.000732				
	wPt9641	4AL	98.4		0.000575				
SKH	*wPt0137*	Unknown		0.004302	0.003792	0.001374			0.001254
	wPt0551	Unknown		0.000558	0.001258	0.001846			0.000916

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt0929	Unknown		0.001147	0.001112	0.000541			0.000494
	wPt1400	4B	38.2				0.00582		0.00476
	wPt1862	1AS	35.7				0.00315		0.007219
	wPt665999	Unknown			0.001002	0.004767			0.002459
	wPt2424	6B	58.1		0.008014				0.007526
	wPt2523	7AS	10.4	0.000162	0.000597	0.003398			0.000771
	wPt3373	7A	77.8	0.001994	0.000726	0.000435			0.000362
	wPt3572	7AS	11.2	0.000638	0.002796				0.003775
	wPt4166	Unknown			0.003184				0.003311
	wPt5167	1AL	125.4	0.003516					0.007256

† GRW10, Greeley wet 2010; Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt5333	6BS	37.6	0.008965					0.002357
	wPt5590	7AS	9.8	0.000558	0.001258	0.001846			0.000916
	wPt5604	5BL	97.4	0.009916	0.004339				0.005253
	wPt5987	7AL	107.1	0.001111	0.00376				0.002277
	wPt6477	2BS	70		0.006261		0.002491		
	wPt664824	1D	51.6	0.003589	0.000615	0.003054			0.001745
	wPt666111	Unknown			0.00059	0.003221			0.001625
	wPt666162	Unknown		0.004302	0.003792	0.001374			0.001254
	wPt6667	6BS	14.3		0.002741	0.009731			0.00533
	wPt669314	Unknown			0.001107	0.005713			0.002613

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt6979	1DS	19.5				0.006422		0.003008
	wPt7160	1BL	73.5	0.000318	0.000583	0.000899			0.000588
	wPt7280	4AL	71.3	0.004626	0.006512	0.006557			0.004044
	wPt732636	Unknown		0.004881	0.000902	0.008391			0.0034
	wPt732908	3DL	160.2	0.007523	0.001379				0.004839
	wPt740691	3D	49.9	0.008374	0.005753	0.004555			0.008061
	wPt741961	3D	66.1	0.000324	0.001177	0.000278			0.000637
	wPt742360	Unknown		0.006869		0.007036			
	wPt745076	3AS	71	0.001777	0.000434	0.001627			0.00052
	wPt7662	6BS	6.1		0.009413	0.005648			

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt8183	6B	54.3		0.001226				0.005346
	wPt8492	2BS	65.7				3.14E-06		0.003729
	wPt8796	4B	65		0.002809				0.001263
	wPt9467	5B	44.8	0.008418	0.006841				0.008685
	wPt9913	Unknown		0.00229	0.00036	0.003524			0.001369
KN	*rPt4199*	7AS	13.1		0.006548			0.006048	0.0008
	tPt1755	7A	79			0.000957			
	wPt0065	Unknown							0.000636
	wPt0419	Unknown					0.000563		
	wPt0866	Unknown						0.000776	

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt4936	5B	59.8	0.000145					
	wPt5265	4BL	108.3						0.000792
	wPt5896	5B	93.7					0.000264	
	wPt671560	3D	46.4		0.000257				
	wPt740903	3D	51.6		0.000286				
	wPt8279	1BL	50.2		0.000492				
	wPt8292	4BL	110.8						0.000792
	wPt8473	7AS	13.7			0.000502			
KNL	*tPt1755*	7A	79			9.17E-05			
	wPt1445	1DL	88.5	0.000682					

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt2994	7BL	158			0.000821			
	wPt5987	7AL	107.1			0.000235			
	wPt8981	7BL	149.5			0.000785			
KNS	*tPt1755*	7A	79			0.000733			
	wPt0065	Unknown		0.000445					
	wPt664939	Unknown		0.000923					
	wPt665174	1AS	11.4			0.000938			
	wPt733544	3BS	39.1	0.000946					
KWS	*wPt0065*	Unknown		0.000489					
	wPt1272	4B	16.6	0.003124					0.004445

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt1445	1DL	88.5	0.007878					0.003755
	wPt5680	2BL	88.3	0.002362					0.00975
	wPt7011	2A	86.1		0.001235				0.005965
	wPt8399	7AL	85	0.001662					0.004129
LA	*wPt0605*	Unknown		0.001629		0.004014			0.001502
	wPt1370	Unknown		0.001693		0.005471			0.001807
	wPt3183	3DL	155.2					3.53E-05	
	wPt3833	5BS	25.6		0.000923				
	wPt4091	5BL	150.6	0.001786		0.004978			0.003525
	wPt4996	5BS	41.1			0.000257			0.006799

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt5072	3BL	131.5	0.009134				0.001175	0.000385
	wPt6971	5BL	154.4	0.006511	0.004622	0.00583			0.006384
	wPt7160	1BL	73.5			0.000312			
	wPt729877	6AS	8.9	0.000903					
	wPt7350	2BL	101.6		0.000635				
	wPt8168	1BL	41.1					0.000106	
	wPt8513	3BL	128.8						0.000905
LL	*wPt0049*	Unknown			0.000996				
	wPt0471	Unknown		0.003603				0.006753	0.002156
	wPt0605	Unknown		0.000771					

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt0605	Unknown		0.000771		0.004913			0.002063
	wPt0837	Unknown					0.005756	0.006833	0.000636
	wPt0896	Unknown		0.002415	0.004702			0.006959	0.001813
	wPt1370	Unknown		0.001509		0.009164			0.004149
	wPt3109	2BL	85.9	0.003603				0.006753	0.002156
	wPt4366	1BS	17.7						0.000919
	wPt4628	5B	69.4				9.47E-05		0.0006
	wPt5072	3BL	131.5	0.000686	0.008348		0.007765	0.001414	0.000304
	wPt6135	5B	76.4				0.00077		
	wPt6971	5BL	154.4		0.002935	0.009873			0.001155

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt7160	1BL	73.5			0.000518			
	wPt730156	3AL	183				0.000499		
	wPt8513	3BL	128.8	0.001293				0.002402	0.000271
	wPt9422	3AL	166.7		0.00043				
LS	*wPt666826*	6BL	84.8	0.000194		0.007055			
	wPt666829	6B	36.1		0.000231				
	wPt666839	6B	50.6		0.000513				
LW	*wPt3833*	5BS	25.6		0.000139				
	wPt4329	2DL	103.6	0.000273					
	wPt4996	5BS	41.1						0.000553

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt666266	6AS	41			0.000463			
	wPt667461	Unknown		0.009363				0.00198	0.003527
	wPt667618	6AS	41.8					0.000353	
	wPt7063	6AS	43.2					0.000459	
	wPt730263	Unknown						7.58E-06	
	wPt740561	7AS	13.1				0.000144		
	wPt742357	Unknown						0.000628	
	wPt742493	Unknown						0.000549	
	wPt8845	3BL	125.5			0.000284			
	wPt9256	6BL	115.2	0.000683					

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
NDVI	*wPt0694*	Unknown		0.005054					0.009837
	wPt3107	1AL	63.7		0.000426				0.003346
	wPt3168	6BL	65.7			0.008064			0.000914
	wPt667155	1BS	13.5		0.003696				0.007153
	wPt9320	2AS	71.9	0.004456					0.004062
PHT	*wPt0934*	Unknown					0.000809		
	wPt2810	5BS	37						0.000888
	wPt3226	7A	158.4						
	wPt3457	5B	73			0.000639			
	wPt5261	3BL	122.1				0.000668		0.000817

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt5704	3BL	101.2	0.008339	0.006755				0.004711
	wPt5816	7BL	221.5	0.000697	0.000154	0.001798			0.001349
	wPt6105	5B	72.4			0.000646			
	wPt667618	6AS	41.8	0.005742			0.000258		0.000107
	wPt667746	3BL	97.5				0.000217		0.000624
	wPt7037	3BL	115.5				0.000166		
	wPt7063	6AS	43.2	0.003			0.000198		5.65E-05
	wPt729839	6AS	45.4	0.000473	0.005548		0.000383		3.83E-06
	wPt731499	Unknown		0.001982			0.001655		0.004002

† GRW10, Greeley wet 2010; GRW2011, Greeley wet 11; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt742680	6AS	8.8		0.000284				
	wPt8418	7AS	9.8	0.000615					
	wPt9925	5BS	17.3		0.000552				
SKD	*rPt9074*	1BL	88.3			0.005021			0.001799
	wPt0408	Unknown			0.008762	0.008971			0.008124
	wPt0944	Unknown				0.007321			0.006843
	wPt1770	1BL	106.5		0.001188	0.000912			0.000993
	wPt2994	7BL	158		0.005241	0.000109			0.00141
	wPt3150	4AL	88.9		0.004905				0.004385
	wPt3349	4AL	84.8		0.009168				0.006299

† GRW10, Greeley wet 2010; GRW2011, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt4209	3B	57.1	0.000658					
	wPt4364	3B	68.6		0.000373				0.006929
	wPt667054	2D	101.2	0.000576					0.005703
	wPt671560	3D	46.4		0.000639	0.00617			0.003204
	wPt731910	3B	70.8		0.000373				0.006929
	wPt744556	Unknown				0.003057			0.007688
	wPt8034	7DL	170.7	0.007909		0.008938			
	wPt9601	6BS	10.7	0.003538					0.008678
	wPt9833	4AL	77	0.007617		0.009332			0.005204
	wPt9928	3AS	52.2	0.006851					0.004044

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
SKW	*wPt0551*	Unknown			0.008219				0.007478
	wPt1387	1D	50.4	0.004681					0.003976
	wPt1480	2AL	110.2		0.004778	0.009378			0.003465
	wPt1492	7DL	172.1	0.001149					0.005852
	wPt1770	1BL	106.5		0.000323	0.00075			0.000466
	wPt3150	4AL	88.9		0.004341				0.008195
	wPt4091	5BL	150.6		0.000227				
	wPt5338	4BL	114.6		0.001216	0.008546			0.002367
	wPt5590	7AS	9.8		0.008219				0.007478
	wPt663755	7DL	172	0.001646					0.005145

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 11; MLKW11, Melkassa wet 11; Melkassa dry 11; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt664012	7DL	171.7	0.002189					0.007041
	wPt667506	7DL	172.2	0.001427					0.007078
	wPt7280	4AL	71.3	0.002435					0.002021
	wPt731740	5BL	163.3		0.000413				
	wPt744556	Unknown				0.004886			0.007662
	wPt7924	4A	64.3			0.003966			0.007241
	wPt8034	7DL	170.7	0.004708					0.006441
	wPt9833	4AL	77	0.004957		0.008352			0.005533
SL	*wPt1272*	4B	16.6		0.009716	0.007468			
	wPt1912	1BS	7.4	0.009036	0.002644				

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt2810	5BS	37		0.003275	0.005618			
	wPt5506	Unknown			0.002878	0.003924			
SN	*wPt1409*	5B	67.7	0.000746					0.002499
	wPt2424	6B	58.1			0.00882			0.004164
	wPt4140	7BL	228.8		0.000392				
	wPt4172	7AS	10.6			0.000426			
	wPt6273	7AS	10.5			0.000426			
	wPt666595	Unknown		2.93E-06					0.00084
	wPt667101	Unknown		2.78E-05					0.005074
	wPt730835	Unknown				0.000609			

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt732448	4B	16.6			0.004362			0.002591
	wPt8399	7AL	85	0.008202					0.001663
	wPt8473	7AS	13.7			0.00023			0.003432
SPN	*wPt1294*	2BL	78.9			7.58E-05			
	wPt3132	2B	77.7			0.000863			
	wPt4230	7B	134.5			8.53E-05			
	wPt742806	Unknown				0.000325			
	wPt8492	2BS	65.7				3.14E-06		
TKW	*wPt0419*	Unknown						0.000793	
	wPt0965	Unknown			0.000446				

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt1770	1BL	106.5		0.001409	0.004722			0.002258
	wPt4091	5BL	150.6		0.003613	0.000396			0.005111
	wPt6043	3BS	15.6				0.005306	0.005267	0.006119
	wPt6105	5B	72.4				0.002784	0.001019	0.000626
	wPt666266	6AS	41					0.00017	
	wPt7024	2AL	81.2		0.000874				
	wPt7280	4A	71.3	0.0004			0.003639		0.003943
	wPt9645	4AL	94.6		0.000159				
	wPt9833	4AL	77			0.00044			
TW	*tPt3719*	5B	72.4			0.006423			0.002304

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	tPt4614	7DL	172	0.005728					0.003427
	tPt6105	2BL	126.8	0.007203		0.004691			
	wPt0817	Unknown			0.004906	0.009478			0.005752
	wPt1264	6B	114.6		0.001573				0.002612
	wPt1733	5BL	93.6		0.001592	0.001985			0.002077
	wPt2994	7BL	158			0.000896			
	wPt5040	Unknown			0.002979				0.006773
	wPt5338	4BL	114.6		0.001539	0.000349			0.000865
	wPt5892	7BL	192.4		0.002273	0.006053			0.001465
	wPt6064	2DL	103.6	0.000183		0.000789			0.000229

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt6132	3BS	24.6	0.001485		8.85E-05			0.0011
	wPt666459	Unknown		0.009043		0.005173			0.004961
	wPt742337	3BS	14.2	0.008509		0.002457			
	wPt744897	7AL	153.4			0.002139			0.007029
	wPt7946	1D	45.6			0.000873			0.006468
	wPt9299	7BL	185.2		0.000866	0.007053			0.000794
	wPt9738	4A	7		0.001025	0.00829			0.003347
YLD	*tPt7183*	1B	27.4			0.000374			
	wPt0419	Unknown					0.000563		
	wPt3457	5B	73		0.000558	0.000137			

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.
‡ Trait description is as given in Table 2.1.

Table 2.10. Continued.

				Marker-trait association *P*-values at each environment†					
Trait‡	Name	Chromosome arm	Position (cM)	GRW10	GRW11	GRD11	MLKW11	MLKD11	Combined
	wPt6135	5B	76.4		0.00055				
	wPt1911	1BS	21.6				0.007398	0.005631	0.008713
	wPt6531	2DS	67.3		0.000961	0.008817	0.002462		0.003398
	wPt8211	7B	69.6				0.002938	0.008011	0.004128

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; Melkassa dry 2011; Combined, Combined data across environments.

‡ Trait description is as given in Table 2.1.

Table 2.11. Marker-trait associations significant at FDR=0.05 for phenotypic traits measured in the WAMII spring wheat association mapping panel in five environments.

Trait	Environment	Marker	Chromosome	Position (cM)	R^2(%)	FDR_*P*-values
Green leaf area	Combined	wPt4532	1BL	88.3	5.6	0.0097
Green leaf area	Combined	wPt0944	unknown		4.9	0.0206
Flag leaf width	MLKD11	wPt730263	unknown	9.0	7.3	0.014
Plant height	Combined	wPt729839	6AS	45.4	5.3	0.0071
Spikelet number	MLKW11	wPt8492	2BS	65.7	7.0	0.0058
Spike number	GRW10	wPt666595	unknown	NA	8.2	0.0054
Spike number	GRW10	wPt667101	unknown	NA	6.6	0.026
Grain filling duration	GRW10	tPt9267	3BS	24.6	6.3	0.0317
Grain filling duration	GRW10	wPt5836	3BS	39.1	6.0	0.0317
Grain filling duration	GRW10	wPt798970	3BS	25.1	5.6	0.0353
Kernel number-based drought susceptibility index	Melkassa	wPt0419	unknown	NA	6.0	0.0226

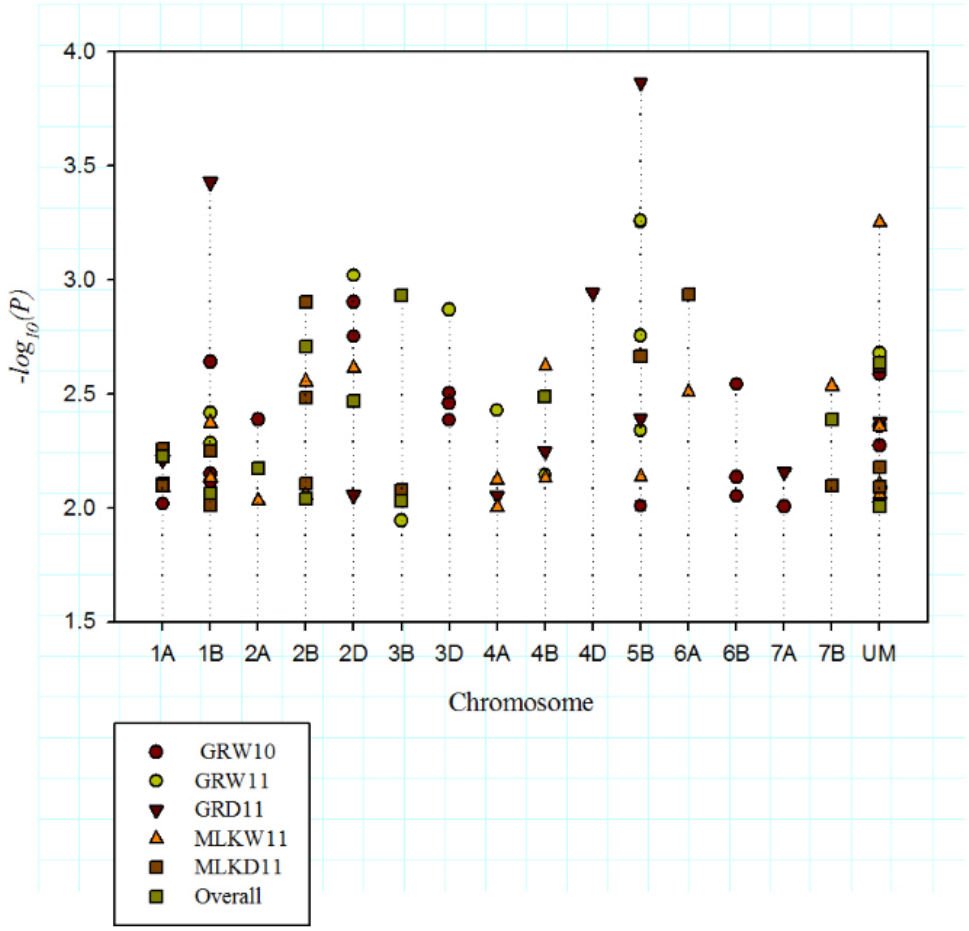

Figure 2.8. Graphical display of marker-trait associations for grain yield at $P<0.01$.

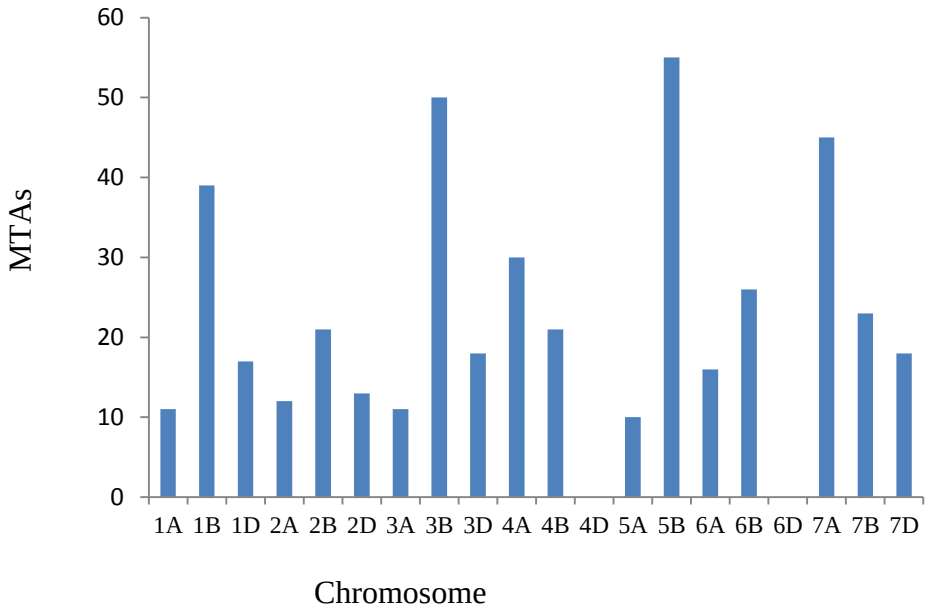

Figure 2.9. Chromosome-wise distribution of marker-trait associations for 26 phenotypic traits significant at $P<0.001$ for single environments or $P<0.01$ for two or more environments.

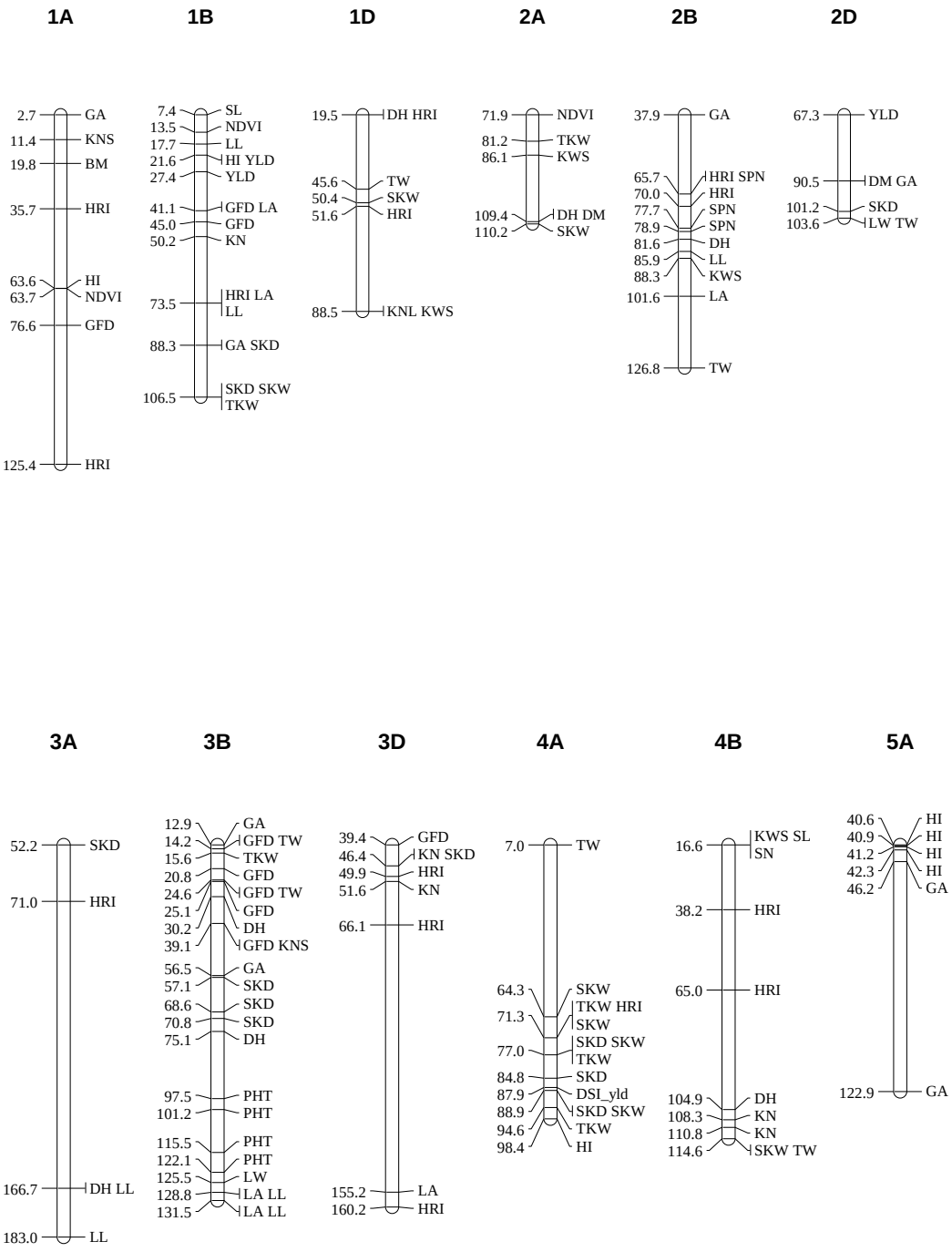

Figure 2.10. Chromosomal regions of QTL identified for phenotypic traits measured in this study.

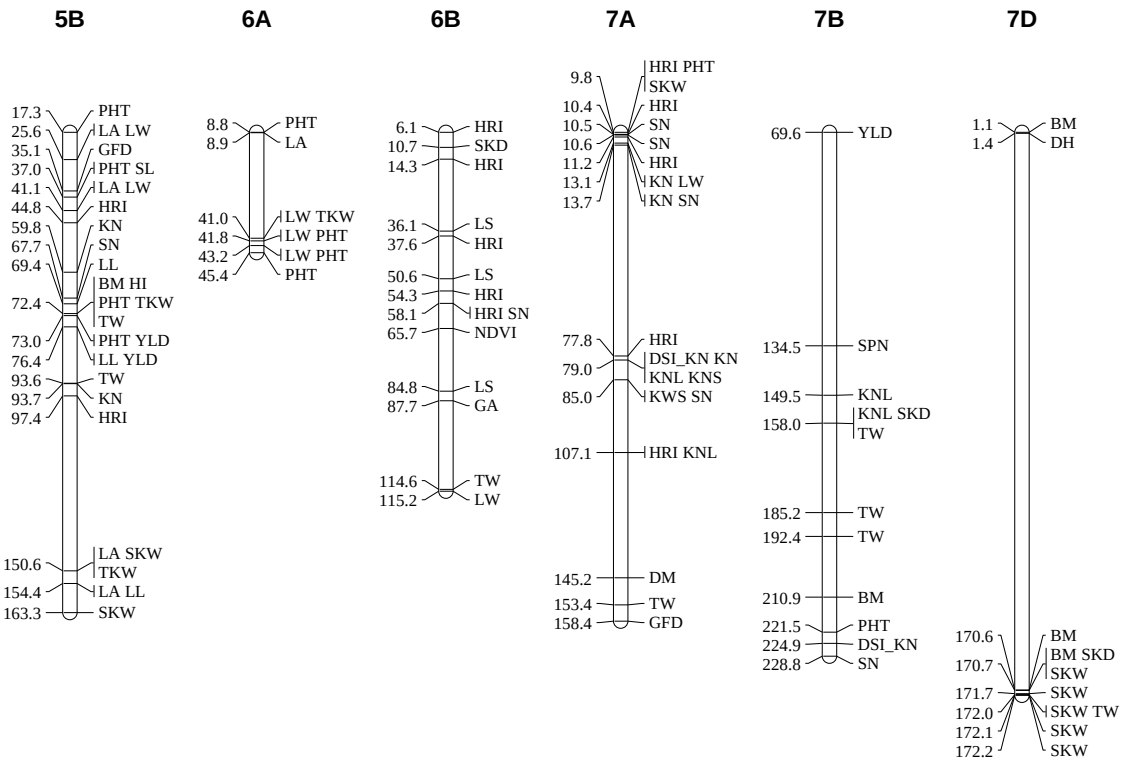

Figure 2.20. Continued.

2.3 DISCUSSION

Trait means, correlation, and heritability

The spring wheat association mapping population panel (WAMII) used in this study was developed by CIMMYT with the intention of identifying QTL/genes underlying drought and heat tolerance related traits. The accessions in the panel had wide differences in morphological characters and agronomic traits. These allowed us to apply a genome-wide association mapping approach for studying the genetic basis of phenotypic variation for traits evaluated under a wide range of environmental factors. The accessions have been exposed to water stress (as low as 192 mm for the entire growing season under rainfed conditions at Greeley in 2011) and heat stress (maximum temperature > 30^{o}C for a majority of the days after heading throughout the grain filling period at Greeley). For wheat anthesis and grain filling, the optimum temperature ranges from 12 to 22 ^{o}C (Farooq et al., 2011). Temperatures above 30 ^{o}C during floret formation in wheat may lead to complete sterility (Saini and Aspinal, 1982).

The measured phenotypic traits responded differently to water deficit and high temperature stresses. The effect was more severe on vegetation indices (NDVI and GA) followed by spike number, kernel number and plant height (Table 2.2). The large effect of water deficit on traits measured at the vegetative stage was reflected mainly on kernel number followed by grain yield due to the severe effect of water deficit on tiller production. Spike number m^{-2} had strong positive genotypic correlation coefficients with both kernel number m^{-2} and grain yield in most of the environments. Generally, kernel-size related traits were increased under drought conditions in this study, and El-Feki (2010) also found similar results for single kernel traits such as kernel diameter and kernel weight for a winter wheat bi-parental population evaluated at the same location (Greeley, CO) as the current study.

Grain yield had significantly negative genotypic correlations with heading date (r_g=-0.26 to -0.68) in four out of five environments, including where water was not a limiting factor. This may be due to high temperature at the end of the growing season, implying that the escape mechanism of stress avoidance is not only useful for terminal drought stress but also for high temperature stress.

Selection for earliness has been effectively used in wheat breeding programs to avoid drought stress. The phenology was restricted in a narrow range during the assembly of this panel to minimize the confounding effect of phenology on QTL detection (Reynolds et al., 2009; Pinto et al., 2010). The presence of significant genotypic correlations between yield and heading date indicates that a small range of heading date may be sufficient to modulate plant adaptions to growing conditions as previously observed by Dodig et al. (2012) and Maccaferri et al. (2011) for Mediterranean environments. The genotypic correlations of grain yield with plant height under irrigated conditions at Greeley (GRW11) were negative while the genotypic correlation under water deficit conditions at Greeley (GRD11) was significant and positive. This is in agreement with the conceptual model of drought tolerance that taller plants provide higher yield under drought stress than shorter plants, which can adapt better to resource-rich environments (Dodig et al., 2012; Reynolds et al., 2005).

Correlation coefficients among phenotypic traits varied depending on the environment. Lopes et al. (2012) reported weak or absence of phenotypic correlations of yield with yield components and other phenotypic traits for the same panel used in the current study but evaluated in different environments. In their report (based on combined means across 12 environments), grain yield was not correlated with thousand kernel weight, days to heading, days to maturity and plant height, but it had a moderate correlation coefficient with kernel number m^{-2}

(r_p=0.45). However, we found significant phenotypic correlation coefficients for yield with TKW (r_p=0.475), PHT (r_p=0.285), DH (r_p=-0.526), DM (r_p=-0.452) and KN (r_p=0.857) for the combined means. In individual environments, however, except for the correlation of grain yield with kernel number, there was inconsistency in magnitude of phenotypic correlation coefficients of grain yield with the remaining traits. For example, unlike Ethiopian environments where the phenotypic correlation coefficients of yield with thousand kernel weight, days to flowering and days to maturity were significantly positive, the phenotypic correlations of yield with these traits in the Colorado environments were weak and non-significant in most cases (data not shown). Again, we found consistently strong positive genotypic and phenotypic correlation coefficients for yield with kernel number m^{-2}, harvest index, final biomass and test weight for individual environments. These discrepancies of our correlation results with the same panel tested in different environments confirms the importance of environmental factors in changing the magnitude and direction of correlations among traits.

Among drought related indices, NDVI recorded at heading date showed relatively higher genotypic correlation (r_g=0.419) with grain yield under rainfed conditions (GRD11) than in more favorable environments (GRW10 and GRW11). Low positive genotypic correlations of NDVI with final biomass were recorded except in one case (GRW10). However, the genotypic correlations between biomass and grain yield were significant and positive in all environments. Prasad et al. (2007a) reported high genotypic correlation of yield with three spectral reflectance indices such as red normalized difference vegetation index (RNDVI), green normalized vegetation index (GNDVI) and simple ratio (SR) taken at booting, heading and grain-filling stage under rainfed conditions with few exceptions. However, Hazratkulova et al. (2012) recently reported the absence of phenotypic correlation between grain yield and NDVI at two

locations for two years both at booting and flowering stages, while there was strong positive correlation for NDVI measurements at milk and dough stages. In our study, we found that NDVI is a good indicator of leaf senescence, as it was consistently negatively correlated with leaf senescence both under irrigated and rainfed conditions. Therefore, stay-green differences among wheat cultivars during grain filling can be quantitatively assessed using NDVI, as previously reported by Lopes and Reynolds et al. (2012).

In the current study, none of the NDVI measurements collected during the vegetative stage showed genotypic differences among the entries. Previous studies have reported similar non-significant variation for NDVI measurements before heading (Babar et al., 2006; Prasad et al., 2007a) and it has been suggested that NDVI measurements taken at heading and grain-filling stages are the best estimates of yield and biomass (Marti et al., 2007; Prasad et al., 2007b). Since NDVI also had a high correlation with leaf green area index recorded with a digital camera, the latter may be used as an alternative for assessing canopy greenness during the vegetative stage because of its ability to discriminate among genotypes before booting stage as observed in our study (data not shown). In the current study, heritability estimates of NDVI, which ranged from 0.28 to 0.62, were comparable with grain yield heritability values (ranging from 0.40 to 0.61). However, Prasad et al. (2007a) reported greater repeatability of NDVI heritability estimates compared to those for yield.

In the current study, heritability estimates of yield were higher than the heritability estimates recorded for the major yield components including harvest index, kernel number and final biomass in most cases. However, traits such as thousand kernel weight, single kernel weight, kernel diameter, test weight and kernel hardness showed consistently higher heritability than yield and yield components. The heritability estimate for yield based on combined data

across environments in this study is roughly similar to that reported by Lopes et al. (2012) (0.61 vs 0.64). However, our heritability estimates (ranging from 0.45 to 0.83) for kernel number m^{-2}, thousand kernel weight, days to heading, days to maturity and plant height are lower than that in the Lopes et al. (2012) report (ranging from 0.83 to 0.97), indicating more environmental influence in the current study. This may be due to artificial environmental stresses imposed, in our case, by withholding water during the vegetative stage in one environment.

Population structure and linkage disequilibrium

Population structure can lead to false associations between markers and traits if not taken into account during association analysis (Zhao et al., 2007). A model-based approach was used to detect subgroups for 287 spring wheat lines in the association mapping panel, and we were able to detect seven subpopulations. In the molecular variance analysis of our study, the significant (P<0.001) population differentiation (Fst ranged from 0.14 to 0.73) for the seven groups reaffirms the presence of population structure. Genetic distance-based cluster analysis also provided evidence for the presence of subpopulations despite the lack of similarity between its clusters and the subgroups of model-based analysis in STRUCTURE. The majority of the variation was explained by within-population variation (78.5%), and among-population variation accounted for 21.5% of the variation. The higher within-population variation demonstrates the impact of selection in maintaining allele diversity in the breeding populations. The magnitude of among- population variation in this study is comparable to variation explained due to differences between European and Asian wheat germplasm (Hao et al., 2010), and even higher than the variation explained due to differences among geographical groups of wheat populations in Europe (Roussel et al., 2005). Chao et al. (2010) reported a higher among-subpopulation genetic variation in spring wheat (17.2%) than in winter wheat (10.5%) from the United States and

CIMMYT breeding programs. Although lines from CIMMYT were used in the present study, the subpopulations are more genetically differentiated than the U.S. spring wheat subpopulations in the study by Chao et al. (2010). In our mapping panel, a substantial number of lines shared one or more parents. Therefore, some of the groups were dominated by lines that trace back to a common parent (data not shown). Because a few elite lines are routinely used as parents of crosses in many breeding programs, this can be expected to lead to some sort of population structure as observed in the current study.

LD information is critical in association studies because LD values can be affected by many factors such as population type, chromosome region and mating system. The number of markers needed for association studies depends on the extent of LD under consideration. In the current population, chromosomes showed large differences in the proportion of marker pairs in significant LD ($P<0.01$) from the maximum 62% for chromosome 4A to the minimum 20% for chromosome 5A. Although chromosome 4A contained more markers in LD at $P<0.01$, markers on chromosome 3D are more physically linked ($r^2=0.264$). However, the proportions of marker pairs at $r^2=0.2$ and $r^2=0.264$ are comparable for all chromosomes implying the importance of choosing an appropriate r^2 value as a threshold in addition to statistical significance. In the current analysis, $r^2=0.2$ was used only for chromosomes with weak LD which do not allow evaluating LD decay rate at the threshold level of physically linked markers ($r^2=0.264$).

The magnitude of LD across a genome or chromosome is a function of nucleotide or linkage distance. LD decay rate was determined both at the genome and individual chromosome level. LD decayed within 2 cM for both A and B genomes, while it extended up to ~ 6.8 cM of genetic distance for the D genome. Chao et al. (2010) reported a similar finding using 394 genetically mapped SNP markers on 478 spring and winter wheat cultivars. The reason for more

extended LD in the D-genome than in the A- and B-genomes could be the introduction of new haplotypes, which can increase the extent of LD, from *Aegilops tauschii* (D-genome donor) into the D-genome of hexaploid wheat germplasm through synthetic wheats. Many lines with synthetic background were included during assembly of this association mapping panel (Lopes et al., 2012). Another potential explanation for extended LD is the genetic bottleneck that occurred with the D-genome as a result of hybridization of tetraploid wheat with few plants of *Ae. tauschii* to form hexaploid wheat (Warburton et al., 2006; Chao et al., 2010). On the other hand, the greater genetic diversity of the A and B genomes is most likely due to early gene flow occurring between hexaploid *T. aestivum* and its tetraploid progenitor *T. turgidum*, but with no similar gene flow occurring between the hexaploid and *Ae. tauschi* (DD) (Berkman et al., 2013).

The LD decay rate was also determined for subpopulations. Generally, LD extended over longer nucleotide or linkage distances (6-9 cM) for subpopulations than the whole panel, which is expected because grouping of genetically similar genotypes reduces within sub-population genetic diversity; consequently, large blocks of a chromosome region could be in LD. We were able to fit LOESS curves only for four out of the seven model-based subpopulations; three of them had similar LD decay rates (within 8-9 cM), and for one group LD decayed relatively faster (within 6 cM) at r^2=0.2. When all subpopulations were considered together, on average LD decayed below the base line r^2=0.2 at ~3.4 cM which is about 50% of the genetic distance within subpopulations. In other words, this translated to the doubling of genetic distance over which LD extended just by grouping similar genotypes together or using closely related genotypes for assessing LD levels.

The effect of subpopulations on phenotypic traits was assessed with multiple regression analysis. Among the traits confounded by population structure, plant height, kernel traits (single

kernel weight, thousand kernel weight and single kernel diameter) and NDVI showed the greatest percentage of phenotypic variation explained by population structure. Interestingly grain yield, kernel number m^{-2}, spike number m^{-2}, harvest index and phenological traits (DH, DM and GFP) were among the group of traits least affected by subpopulations. Except for TKW and HI, these results are in agreement with Dodig et al. (2012) who reported large effects of population structure in winter wheat on stem related traits (stem height, peduncle length and peduncle extrusion); a moderate influence on sterile spikelets $spike^{-1}$ and biomass per plant; and a low effect on yield and yield components (KN, TKW and SN). The greatest effect of population structure on kernel-size related traits in our panel may be due to intensive selection for kernel size in CIMMYT's breeding program (Ravi Singh, personal comm.). Elite lines are most likely larger in kernel size than the remaining lines included in the panel for the purpose of maintaining genetic diversity during assembly of the mapping panel. The low effect of population structure on heading date indicates the minimum confounding effect of phenology on population structure, unlike plant height and kernel size.

Marker-trait associations

Although grain yield QTL were detected on all wheat chromosomes in previous studies, relatively consistent MTA in our study were detected on chromosomes 1BS, 2DS, 5B and 7B. Broad comparison of MTA results from the current study with previous studies were made using chromosome arms because of differences in marker type and marker positions on different genetic maps. The DArT marker *wpt6531* on the short arm of chromosome 2D, which was associated with yield in the current study, is about 8 cM away from the *wpt4144* marker, which was associated with yield in the Crossa et al. (2007) study. Kumar et al. (2007) detected QTL for yield in this region linked to SSR marker *gwm261* which is 14.4 cM distal to *Ppd-D1* and 0.6 cM

distal to the height-reducing semi-dwarfing *Rht8* locus (Korzun et al., 1998; Ellis et al., 2007). Dodig et al. (2012) also detected QTL on chromosome 2DS (near *gwm484*) that explained about 22% of the phenotypic variation for grain yield. Therefore, the stable and highly significant grain yield MTA on 2DS in the current study is probably due to a grain yield QTL in proximity to the *Ppd-D1* locus, which is known for its influence on wheat yield through optimization of flowering time (Worland, 1996). Significant MTA for yield were detected on the short arm of chromosome 1B in the Crossa et al. (2007) study. Quarrie et al. (2005) found major QTL which explained up to 35% of the phenotypic variation and were expressed in 11 out of 24 trials on 7BL.

In the current study, *wpt8211* on chromosome 7B (69.6 cM) was associated with yield in three environments. This marker (*wpt8211*) had sequence similarity with transposable element-related sequences in the Triticeae repeat-sequence (TREP) database based on the DArT characterization study of Marone et al. (2012). The marker *wpt3457* on chromosome 5B (73 cM) was associated with yield both under irrigated and rainfed conditions, and marker *wpt6135*, which was physically in LD and 3.4 cM away from *wpt3457*, was strongly associated with yield under irrigated conditions. Moreover, many other markers consistently associated with traits such as thousand kernel weight, final biomass, harvest index, plant height and flag leaf length also resided close to either side of the QTL position for yield on 5B, indicating the importance of this region in influencing yield and yield components. This region may explain a portion of the genotypic correlations of yield with yield component traits.

In previous studies, yield QTL have been detected on both long and short arms of chromosome 5B (Neumann et al., 2011; Crossa et al., 2007; Marza et al, 2006; Groos et al, 2003; Huang et al., 2003) and some of their QTL may coincide with the QTL detected here on chromosome 5B. However, to our knowledge, there are no reports on the presence of multi-trait

QTL near the centromeric region of chromosome 5B. In fact, chromosome 5B comprised the highest number of MTA in this study. Kumar et al. (2007) reported multi-trait QTL for yield and yield components on chromosomes 2DS and 4AL. No multi-trait regions were observed on 2DS for the yield component traits in this study, but a region of chromosome 4AL was identified as a multi-trait QTL region for kernel size and quality traits.

Grain yield and harvest index shared an association region on chromosome 1BS, implying that there is a genetic basis for the high and consistent genotypic correlation observed between grain yield and harvest index. In addition, NDVI has QTL in common only with harvest index on chromosome 1AL of all yield component traits, while green leaf area shared QTL with harvest index on chromosomes 5A (42.3 cM) and with single kernel diameter on chromosome 3B (56.5 cM). The benefits of assessing yielding ability of wheat with these vegetation indices may be dictated by the expression of genes in the chromosome regions that harvest index and single kernel diameter shared with the indices.

Trait-specific stable MTA were detected for main yield component traits such as kernel number per m^2, harvest index and thousand kernel weight. Unlike harvest index and thousand kernel weight, only one marker on chromosome 7AS showed consistency across environments for kernel number. Among environment-specific MTA for kernel number, the unmapped marker *wpt0866* had sequence similarity with 1, 3-beta glucan synthase (Marone et al., 2012). Similarly, all MTA of final biomass, kernel number per spikelet and kernel number per spike were environment-specific, showing the presence of higher genotype by environment interaction for these yield component traits than yield itself. Among yield and yield component traits, however, very strong (FDR=0.05) MTA were obtained for spikelet number per spike on chromosome 2BS (*wpt8492*), and for spike number per m^2 for two unmapped markers *wpt666595* and *wpt667101*.

None of these MTA are in agreement with previously detected QTL in bi-parental populations with the exception of MTA noted for spikelet number on chromosomes 2B and 7B by Neumann et al. (2011) which may be comparable with our current findings.

Stacking QTL that control traits of interest from different chromosome regions into one background is a challenging and time consuming task in plant breeding. Using multi-trait markers in marker-assisted selection may increase QTL pyramiding efficiency. With the exception of chromosomes 4D, 5D and 6D, two or more traits shared the same region or reside within 5 cM in all chromosomes. Kernel size-related traits (single kernel weight, single kernel diameter and thousand kernel weight) had QTL in common on chromosomes 1BL, 4AL (SKW, SKD and TKW) and 7DL (SKW and SKD). Test weight also shared the same regions with one or more kernel size-related traits on chromosomes 1B, 2DL, 4BL, 7BL and 7DL. These traits could be under the same genetic control and markers in those multi-trait regions could be used in future for improvement of kernel size-related traits through marker-assisted selection.

Similarly, clusters of QTL for flag leaf characters (LA, LL and LW) were found on chromosomes 3BL and 5BL. Moreover, there is a pattern of coincidence between leaf character QTL and kernel size-related traits TKW or SKW. This may be related with the translocation of flag leaf photosynthetic product to kernels during grain filling period (Lupton, 1966).

Although a wide range of mean phenotypic values were recorded for plant height from 41 cm (GRD11) to 88 cm (MLKW11), the major plant height reducing loci *Rht-D1* and *Rht-B1* were not detected in this panel. This could be due to low marker coverage in the region of semi-dwarfing genes (e.g., only two markers on chromosome 4D). However, we detected MTA in the regions of previously reported plant height QTL on chromosomes 3BL (Maccaferri et al., 2011),

5B (Cadalen et al., 1997; McIntyre et al., 2010; McCartney et al., 2005), 6AS (Spiel Meyer et al., 2005) and 7BL (McCartney et al., 2005). Similarly, regions of group 5 chromosomes where *VRN-1* genes reside were not detected for heading date in this panel. Nonetheless, in agreement with the results in this study, QTL that affect flowering time in wheat have been reported on chromosomes 2B, 3AL, 3B and 7DS (Borner et al., 2002; Marza et al., 2006; Cuthbert et al., 2008; Wang et al., 2009).

Photoperiod genes which have been mapped on short arms of homoeologous group 2 chromosomes were not detected for heading date in this study. However, the QTL detected on 3AL may indicate variation in an earliness *per se* gene as this gene has been mapped on chromosome 3AL (Borner et al., 2002). During assembly of association mapping panel, a wide range of variation both in heading date and plant height is not desired. In this particular population, most likely the number of spurious MTA due to confounding effects from major plant height reducing genes, vernalization genes and photoperiod response genes are minimal, implying the validity of the panel to conduct association studies for traits of interest with dense SNP markers.

In conclusion, we have shown that LD decay varied both at the genome and chromosome levels. Genome-wide association mapping effectively detected both stable and environment-specific QTL for yield, yield components, and drought-related traits. Multi-trait chromosome regions have been detected and particularly the region on chromosome 5B associated with yield and yield component traits may be useful in MAS following proper validation. In the context of drought tolerance, QTL regions that control both drought tolerance-related traits and yield component traits were detected on chromosomes 1AL (NDVI and harvest index), 5AL (green

leaf area and harvest index) and 3B (green leaf area and single kernel diameter), implying the possibility of using vegetation indices for indirect assessment for certain yield component traits.

CHAPTER 3

Association Mapping and Nucleotide Sequence Variation in Five Drought Tolerance Candidate Genes in Spring Wheat

SUMMARY

Drought tolerance is an integrative trait that involves the expression of many genes. Functional markers are needed for key genes to facilitate the application of marker-assisted selection (MAS) for improvement of drought stress tolerance.

The objectives of this study were to (1) characterize five drought tolerance candidate genes in wheat (*Triticum aestivum* L.) for nucleotide and haplotype diversity, Tajima's D value, and linkage disequilibrium (LD), and (2) determine the association between within-gene single nucleotide polymorphisms (SNPs) and phenotypic traits in a spring wheat association mapping panel.

Five candidate genes, namely dehydration responsive binding 1A (*DREB1A*), enhanced response to abscisic acid (*ERA1-B* and *ERA1-D*), and fructan 1-exohydrolase (*1-FEH-A* and *1-FEH-B*), were amplified and sequenced from 126 spring wheat lines. The lines were evaluated in field experiments under contrasting moisture regimes at Greeley, CO, USA and Melkassa, Ethiopia. Polymorphic sites were identified within DNA sequences of each gene. Differences were observed among the candidate genes for nucleotide and haplotype diversity, Tajima's D test, and patterns of LD. The genes were associated ($_{P}$FDR=0.1) with yield, yield components, and morphological and phenological traits in one or more environments. If validated in relevant genetic backgrounds, the identified marker-trait associations may be applied to functional MAS.

3.0 INTRODUCTION

Drought tolerance is a complex trait that involves the expression of many genes. A better understanding of the roles and relative importance of those genes would aid the development of drought tolerant crop cultivars. A drought tolerance candidate gene is a DNA sequence that co-maps with a drought tolerance quantitative trait locus (QTL) and encodes a protein that can be functionally associated with the stress response/adaptation process (Cattivelli et al., 2008). In plants, the construction of molecular linkage maps based on candidate genes is one way of identifying the genes underlying QTL instead of time-consuming fine mapping. This candidate gene strategy shows promise for bridging the gap between quantitative genetic and molecular genetic approaches to study complex traits like drought tolerance. Candidate gene association mapping is aimed at linking phenotypic variation with polymorphic sites in candidate genes to identify causative polymorphisms (Gonzalez-Martinez et al., 2008).

Drought stress induces a large number of genes that have been identified and characterized for their function (Shinozaki and Yamaguchi-Shinozaki, 2007). There are two categories of genes in terms of response to the phyto-hormone abscisic acid (ABA): ABA-independent and ABA-dependent. For example, ABA-independent dehydration responsive element binding (DREB) genes are known for their association with abiotic stress tolerance (Latini et al., 2007). Full-length sequences of *DREB1* and *DREB2* genes have been cloned from rice (*Oryza sativa*), maize (*Zea mays*), *Arabidopsis thaliana*, and wheat, and the *DREB1* gene sequences from the three genomes of wheat have been mapped to chromosomes 3A, 3B and 3D (Wei et al., 2009). Transgenic wheat with the *DREB1A* gene from *Arabidopsis* controlled by the stress-inducible *rd29a* promoter showed greater root branching, increased drought tolerance, and larger spike size than non-transgenic wheat plants in a greenhouse study (Pellegrineschi et al.,

2004). However, in a recent field evaluation the transgenic *DREB1A*-wheat lines did not have a grain yield advantage over control lines under water deficit conditions (Saint Pierre et al., 2012), despite their better recovery after severe water stress and higher water use efficiency in the greenhouse. In transgenic groundnut (*Arachis hypogea*) plants, *DREB1A* improved transpiration efficiency (Bhatnagar-Mathurwater et al., 2007), increased root/shoot ratio (Vadez et al., 2007), and increased root length density in deeper soil layers under water deficit conditions, thereby enhancing water uptake of transgenic plants (Vadez et al., 2013).The *DREB2* gene from wheat improved freezing and osmotic stress tolerance when expressed in tobacco (*Nicotiana tabacum*) plants (Kobayashi et al., 2008). Fructan 1-exohydrolase (*1-FEH*) is another ABA-independent gene that is implicated in cold and drought tolerance through membrane stabilization and remobilization of water-soluble carbohydrates from stem to developing grain (Lothier et al., 2007; Hincha et al., 2003). The three copies of the *1-FEH* gene were mapped to the short arms of group 6 chromosomes, i.e., 6AS, 6BS, and 6DS (Zhang et al., 2008).

Increased ABA production under drought conditions activates expression of ABA-dependent drought tolerance-related genes (Shinozaki and Yamaguchi-Shinozaki, 2007). Expression of the *ERA1* (Enhanced Response to ABA) gene, which has been cloned from *Arabidopsis* (Cutler et al., 1996) and hexaploid wheat (Manmathan et al., 2013), is ABA-dependent in its expression. It has been shown that *ERA1* mutants increased drought tolerance of Arabidopsis through ABA stimulated stomatal closure, thereby effectively reducing water loss through transpiration (Pei et al., 1998; Ziegelhoffer et al., 2000).

Marker-assisted selection has increased the precision of the variety development process in classical plant breeding for genes of relatively large effect. Single nucleotide polymorphisms are becoming the markers of choice in plant breeding programs for construction of high-

resolution genetic maps and genomic selection. SNPs are generally more abundant, stable, amenable to automation, efficient, and cost-effective than other forms of genetic markers (Rafalski, 2002; Akhunov et al., 2009). SNPs can be individually responsible for phenotypic variation of a trait or linked to causative SNPs (Langridge and Fleury, 2011). However, selecting the most suitable set of SNPs (either causative or linked) in a cost-effective manner is a key step toward application of molecular markers for crop improvement (McCouch et al., 2010).

SNPs may be discovered with different methods. However, the most straightforward approach is direct re-sequencing of amplicons of genes from different genotypes (Rafalski, 2002). Amplification of DNA segments with genome-specific primers for polyploids like hexaploid wheat is challenging due to sequence similarity among gene copies on homoeologous chromosomes, and among genes within a gene family. This may slow down to some extent the application of functional markers in wheat breeding.

A functional marker is a marker developed from a SNP or insertion/deletion (indel) within a gene that is responsible for variation in the trait of interest (Andersen and Lubberstedt, 2003). The use of functional markers in molecular plant breeding is more advantageous than linked markers because the latter are not diagnostic across breeding populations due to recombination between the marker and the putative causative SNP region in subsequent generations. Since functional markers are developed from SNPs within a gene, marker information can be used confidently across breeding programs to select favorable alleles for a trait of interest (Bagge and Lubberstedt, 2008). Several genes for agronomic traits (e.g., semi-dwarfism genes) and quality traits (e.g., polyphenol oxidase) have been identified for wheat (Wei et al., 2009; Bagge and Lubberstedt, 2008), but functional markers have been developed for only

a few of them. Therefore, more functional markers are needed to enhance the application of molecular markers in crop improvement.

Generally, once genes that determine the genetic basis of a trait are known, developing functional markers to select for favorable alleles is an important aspect of using genetic information in practical plant breeding (Langridge and Fleury, 2011). However, for successful functional marker development, prior information about the level of DNA polymorphism, extent of linkage disequilibrium, and within gene nucleotide diversity is required. This information is rare for drought tolerance genes in hexaploid wheat. Therefore, the objectives of this study were to (1) characterize five drought tolerance candidate genes in wheat for nucleotide and haplotype diversity, Tajima's D value, and LD, and (2) determine the association between within-gene SNPs and phenotypic traits in a spring wheat association mapping panel.

3.1 MATERIALS AND METHODS

Plant materials

A total of 126 lines was selected based on their phenotypic diversity from a set of 294 spring wheat lines of an association mapping panel (WAMII), which was developed by the International Maize and Wheat Improvement Center (CIMMYT) from entries in the Elite Spring Wheat Yield Trial (26^{th}, 27^{th} and 28^{th} ESWYT), Semiarid Wheat Yield Trial (1^{st} to 16^{th} SAWYT) and High Temperature Wheat Yield Trial (HTWYT) (Lopes et al., 2012). The panel was developed to identify QTL or genes for drought and heat tolerance, and included many synthetic hexaploid-derived wheat lines (Lopes and Reynolds, 2012). The entire mapping panel was grown and self-pollinated for one generation before sub-sampling. The lines in the sub-sample were chosen based on their diversity in morphology and agronomic characters in field evaluations under rainfed and irrigated conditions in 2011 at Greeley, CO. The pedigree of each line was also taken into account to minimize parental relatedness.

Phenotypic evaluation

The field trial was conducted at Greeley, CO (latitude 40.45N, longitude 104.64E, elevation 1427 m) in 2010 and 2011, and at Melkassa, Ethiopia (latitude 8.40 N, longitude 39.33 E, elevation 1550 m) in 2011. The soil at the Greeley site is well-drained with fine sandy loam to clay loam texture and a pH of 7.4-8.4. The dominant soil type at Melkassa is sandy loam (Andosol of volcanic origin) with pH ranging from 7.0 to 8.2 (Ethiopian Institute of Agricultural Research, Melkassa Agricultural Research Center).

On 5 April 2010, we planted 285 lines for evaluation under fully irrigated conditions in Greeley. The site received a total of 271 mm of rainfall from January through July, and the plots were supplemented with 94 mm from three irrigations (twice during the vegetative stage and

once after heading). In 2011 we evaluated 288 lines at Greeley under both fully irrigated ("wet") and rainfed ("dry") conditions. Both treatments were irrigated similarly at planting (15 April), but later the wet treatment was supplemented three times with drip irrigation during the vegetative and grain filling stages, while the dry treatment received supplemental irrigation only once at heading to avoid complete failure of the experiment. The wet treatment received a total of 313 mm water (rainfall plus irrigation), whereas the dry treatment received 192 mm of water (rainfall plus irrigation) during the growing season and the preceding three months (January through July).

In both years each entry was replicated twice in a Latinized row-column design prepared with CycDesign 3.0 software (www.cycdesign.co.nz). Each line was planted in four-row plots 1.53 m long and 0.92 m wide with 0.20 m spacing between rows. The seeding rate was approximately 173 seeds m^{-2}. Weeds were controlled manually as required.

At Melkassa Ethiopia, 294 lines were planted on 17 July 2011 on wet soil and on 19 July 2011 drier soil in an adjacent field. The experiment was laid out as an alpha lattice design with 14 plots per block and two replications. Plots were two rows, 2.5 m long, with 0.2 m spacing between rows and 0.4 m spacing between plots. Seeding rate was based on local recommendation of 150 kg ha^{-1}. Nitrogen fertilizer was applied in split applications at planting and tillering at a total rate of 50 kg ha^{-1}. Phosphorus fertilizer was applied at planting as diammonium phosphate at a rate of 100 kg ha^{-1}. The site received a total of 533 mm rainfall during the growing season (July-September, 2011).

Morphological and phenological traits

Flag leaf maximum length (measured in cm from leaf collar to the tip) and maximum width (measured in cm on the widest part of the leaf) were recorded as the average measurement of three flag leaves per plot, and flag leaf area (cm^2) was calculated as flag leaf length x flag leaf width x 0.75.

Plant height was recorded as the average of three values measured in cm from the soil surface to the tip of the spike excluding awns. Days to heading was recorded as the number of days from planting until 50% of the spikes in each plot had completely emerged above the flag leaves. Days to maturity was recorded as the number of days from planting until 50% of the peduncles in each plot had turned yellow. Grain filling duration was calculated as the difference between the days to heading and days to maturity.

Vegetation indices and leaf senescence

Normalized vegetation index (NDVI) was obtained by scanning plants in each plot during the grain filling stage with a GreenSeeker instrument model 3541 (NTech Industries Inc., Boulder, CO). A green leaf area index was obtained from a photo taken at a height of approximately 0.50 m directly above each plot with a digital camera (Coolpix S8100, Nikon Corp., Japan), and processed with Breedpix software (Casadesus et al., 2007). Leaf senescence was scored during grain filling stage a week before physiological maturity time on a scale from 0 to10, where 0 indicates completely green leaves and 10 indicates that all leaves in a plot had changed completely to yellow.

Kernel and grain yield-related traits

Biomass samples were taken by cutting all the plants at ground level in one row of each plot at maturity. Final dry biomass was determined by weighing samples after 48 hours in a 40

^{o}C drier. Those samples were threshed and the grain weight was used to calculate harvest index as the ratio of grain weight to total biomass weight. The remaining plants in each plot were harvested by a combine. Grain yield was the total weight of seed in each plot (combine harvest + biomass grain weight) divided by the plot area and expressed as kg ha^{-1}. Spike length, spikelets number per spike, kernel number and weight (g) per spike, and kernel number per spikelet were recorded as the average of five spikes per plot. Thousand kernel weight was determined by extrapolation after counting seeds of five spikes with a seed counter (International Marketing and Design Corp. Model 900-2, San Antonio, TX) and obtaining the weight of the seeds. Number of spikes m^{-2} was calculated by dividing the number of kernels m^{-2} by kernel number per spike. The number of kernels m^{-2} was obtained from the ratio of grain weight m^{-2} to thousand kernel weight, multiplied by 1000 (= (YLD (g/m^{-2}/TKW (g)) x 1000). Single kernel diameter (mm), kernel hardness and single kernel weight (mg) were determined from 100 seeds using a Single Kernel Characterization System Instrument model 4100 (Perten Instruments, Springfield, IL). Test weight (kg hL^{-1}) was determined using standard procedures from a small sample of the grain collected at harvest.

Phenotypic data analysis

Analysis of variance for the phenotypic data was conducted first using the GLM procedure of SAS v. 9.3 (SAS, Institute, 2011), considering genotype as a fixed effect. Normality of the data for each trait was checked using a Q-Q plot of residuals in the SAS GLIMMIX procedure, and all traits were consistent with a normal distribution. Best linear unbiased predictions (BLUPs) and variance components were obtained for all traits using the Mixed procedure in SAS, considering genotype as a random variable. In the combined data analysis, environment was considered a fixed variable. To account for spatial variation in the experimental

field, four spatial models (spatial power, anisotropic spatial power, Matérn spatial, and Autoregressive models (AR1 x AR1)) were compared using minimum Akaike Information Criterion and Bayesian Information Criterion for each trait (SAS Institute, 2006). Since the correlation value due to spatial variability in each model was very low for each data set (except at Greeley in 2010), there was little benefit from spatial adjustment in this study.

Candidate gene selection and analysis

Three drought tolerance candidate genes (*ERA1, DREB1A* and *1-FEH*) were selected for SNP identification, nucleotide diversity, and association analyses. Reference DNA sequences of two genes, *DREB1A* and, *1-FEH* were obtained from National Center for Biotechnology Information (NCBI) GenBank database (www.ncbi.nlm.nih.gov). The third candidate gene, *ERA1*, was recently cloned from wheat (Manmathan et al., 2013). Its cloning involved designing primers from conserved regions of previously identified homologous genes of related species, amplifying the gene region from hexaploid wheat, cloning the PCR products into plasmids, and sequencing plasmid clones to identify sequences of the gene on the A, B, and D genomes. A primer pair reported by Wei et al. (2009) was used to amplify the *DREB1A* gene. Both coding and non-coding regions were amplified with this primer pair. The primers used to amplify *ERA1* and *1-FEH*, however, were designed for genome-specific amplification with primers designed from unique regions of the genes using primer3 software (frodo.wi.mit.edu/). The gene structure for *ERA1* was predicted using GeneMark software (http://exon.gatech.edu/). All primer pairs used to amplify *1-FEH* were from non-coding regions of the gene copies. Genome specificity of the primers was verified by PCR amplification of the corresponding gene in the hexaploid wheat progenitors *T. urartu* (AA, 2n=2x=14), *Aegilops speltoides* (BB, 2n=2x=14) and *Ae. tauschii* (DD, 2n=2x=14). A complete list of the genome-specific primers is given in Table 3.1.

Table 3.1. Primer sequences used to amplify drought tolerance candidate genes.

Target gene	Primer	Sequence
DREB1A	P21F	5'-CGGAACCACTCCCTCCATCTC-3'
	P21R	5'-CGGTTGCCCCATTAGACGTAA-3'
ERA1-B	ERA1BF	5'-GATGTGACAATACATTACATATGCAGCT-3'
	ERA1BR	5'-GGTGGGTACGTTTCTAAGGATGG-3'
ERA1-D	ERA1DF	5'-CAACTCTGAACTATTGCAAAAGTGAACTTTC-3'
	ERADR	5'-CTGCAATATCGGTGAGTTTCTTGTAGTTAA-3'
1-FEH-A	W12F	5'-TATGCCACTTCCATGCTGGTA-3'
	W12R	5'-CGATGCTGCTGCCAAGAATATAC-3'
1-FEH-B	W32F	5'-CAAGAACTGGATGAACGGTACAT-3'
	W32R	5'-CAATGGCTACTTGTGTTTAGCC-3'

To extract DNA, leaf tissues were sampled from 2-week-old seedlings of the 126 lines grown in a greenhouse. The leaf samples were immediately transferred to 2-mL tubes and stored at -80 ^{o}C. DNA was extracted following a standard Cetyl Trimethyl Ammonium Bromide (CTAB) extraction method with minor modification (Wei et al., 2009).

The following PCR protocol was used for the *ERA1* and *1-FEH* genes. A total volume of 25 μL containing 100 ng of genomic DNA, 1x PCR reaction buffer, 0.20 μM of each primer, 0.20 mM dNTPs, 1.5 mM $MgCl_2$,0.5 U of VELOCITY DNA polymerase (www.bioline.com), and 3% dimethyl sulfoxide (DMSO) was used. The PCR was carried out on a MJ PTC-200 programmable thermal controller (MJ Research, Bio-Rad, Hercules, CA) as follows: initial denaturation at 98 ^{o}C for 2 min; 30 cycles of 98 ^{o}C for 1 min, an annealing step at 68 ^{o}C for *ERA1* and 64 ^{o}C for *1-FEH* for 1 min, and 72 ^{o}C for 1.5 min; and final extension at 72 ^{o}C for 10 min. Amplification of *DREB1A* was conducted using a total volume of 25 μL containing 100 ng of genomic DNA, 1x PCR reaction buffer, 0.25 μM of each primer, 0.45 mM dNTPs, 4.0 mM $MgCl_2$, and 1.6 U of *Taq* DNA polymerase (Promega, Madison, WI). The PCR amplification was done on a MJ PTC-200 programmable thermal controller at an initial denaturation temperature of 94 ^{o}C for 3 min, followed by 34 amplification cycles at 94 ^{o}C for 1 min, annealing temperature of 63 ^{o}C for 1 min, and 72 ^{o}C for 1.5 min, and final extension at 72 ^{o}C for 10 min. For each candidate gene, a primer optimization step was done on two genotypes from the mapping panel.

The expected size of each PCR product was confirmed by separation on 1.5-2% agarose gels, stained with ethidium bromide, and visualized under UV light. The amplified PCR products were purified and sequenced on an ABI sequencing instrument at Beckman Coulter Genomics (Beckman Coulter Genomics, Danvers, MA). Sequences were initially obtained from 32

representative diverse genotypes. After confirming the presence of SNPs within the genes, PCR products from the remaining 94 lines were sequenced. The overall sequence data quality was high, with a pass rate of 96.35% and 97.60% for the 32 diverse genotypes and the remaining 94 lines, respectively.

To identify SNPs, consensus sequences were first obtained by aligning reverse and forward sequences with the reference sequences of each gene using SeqMan software (www.dnastar.com/t-nextgen-seqman-ngen.aspx). Those sequences that showed less than 80% sequence identity with the references were excluded from subsequent analyses. Aligned sequences of each gene were analyzed for sequence diversity by characterizing nucleotide diversity, haplotype diversity, and linkage disequilibrium using DnaSP version 5 software (Rozas et al., 1999). Genome-specific sequences of the *ERA1* gene were mapped to the long arms of chromosomes 3A, 3B, and 3D through a BLAST search against the survey sequences of all individual chromosomes of bread wheat in the International Wheat Sequencing Consortium database (www.wheatgenome.org).

Population structure, LD, and marker-trait association analysis

Diversity Array Technology (DArT) markers (Triticarte Pty. Ltd., Canberra, Australia; www.triticarte.com.au; Akbari et al., 2006) were used to account for population structure and genetic relationship of the evaluated lines. A total 78 DArT markers (spaced > 10 cM) was selected from all chromosomes to determine the population structure. An admixture model with correlated allele frequency model in STRUCTURE software (Pritchard et al., 2000) was applied with a burn-in of 20,000 iterations and 20,000 Markov Chain Monte Carlo (MCMC) duration to test a k value in the range of 3 to 12. Each K was replicated five times and the run that assigned

more lines with probability of >0.5 in all clusters was used. The likely number of subpopulations was determined using the approach of Evanno et al. (2005).

Single nucleotide polymorphisms within each gene were used to determine pair-wise LD with GGT2 computer software (www.plantbreeding.wur.nl/uk). Fisher's exact test was used to decide the significance of the LD among SNPs. Linkage disequilibrium was calculated across chromosomes 3A and 6A to compare the extent of LD decay around *DREB1A*, *ERA1*, and *1-FEH* genes.

Phenotypic data collected from five environments were used to determine the effects of SNPs within each gene on the phenotypic traits. Since the selected lines were highly homozygous breeding lines developed via several generations of self-pollination, only a few sites were found to be heterozygous and these sites were considered as missing values in association analysis. DArT markers were used to calculate kinship matrices among the lines as suggested by Bernardo (1993). A false discovery rate adjusted probability value of 0.1 was used as the threshold for significance of SNP-trait associations (Benjamini and Hochberg, 1995). A mixed linear model (Yu et al., 2006) with population structure and kinship in the model, as implemented in the TASSEL software version 3.0 (Bradbury et al., 2007), was applied for association analysis. This model showed least deviation of observed *P*-values from expected *P*-values in Q-Q plot when compared with that of Q (population structure) or K (kinship) model only. For SNPs that explain larger portions of phenotypic variation, phenotypic means for genotypic classes were graphically displayed to compare the effects of common alleles against that of rare allele of each SNP. Haplotype-trait associations were also conducted using within-gene SNP combinations in the TASSEL software.

3.2 RESULTS

Phenotypic Evaluation

Analysis of variance revealed significant differences ($P<0.05$) among the genotypes for most traits in all environments. A total of 26 traits were measured or scored in this study, but this number varied depending on the year and location. In the combined data analysis, variation due to genotype by environment interaction was about 20% of the total variation while about 13% of the total variation was due to genotypic differences. The mean yield of individual lines ranged from a low of 1087 kg ha^{-1} at Greeley under dry conditions in 2011 to 5513 kg ha^{-1} at Melkassa under non-stressed conditions in 2011 (Table 3.2). The mean yield performance of genotypes in the non-stressed treatment at Melkassa was about three-fold higher than that of the irrigated treatment at Greeley in the same year (Table 3.2). Furthermore, the genotypes had longer flag leaves, greater plant height, longer grain filling duration, more final biomass production, and consequently, higher grain yield at Melkassa compared to Greeley. Although days to heading occurred within a range of four to seven days in the Greeley environments for the sub-sampled population, the range at Melkassa was 15 to 18 days (Table 3.2).

Table 3.2. Phenotypic mean and range of selected spring wheat association mapping panel entries evaluated at five environments.

	Environments†									
	GRW10		GRW11		GRD11		MLKW11		MLKD11	
Trait‡	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
YLD	2179	1510-2791	1528	1241-1865	1301	1087-1511	4420	2615-5513	3904	2444-4893
TKW	35.45	31.00-41.48	21.37	18.93-23.90	25.11	20.81-31.64	25.53	16.28-34.73	22.6	16.66-27.74
TW	77.84	72.30-81.97	65.46	62.43-68.97	69.07	62.54-72.83				
DH	67.24	62.88-69.67	69.96	67.87-72.30	68.29	65.56-70.80	54.92	50.75-66.10	55.68	50.61-68.34
DM	103.22	99.64-105.65	104.04	102.06-106.01	99.58	97.30-106.76	92.78	88.09-98.89	95.16	92.05-98.52
GFD	35.98	31.67-41.40	34.00	32.05-37.32	30.59	28.59-38.25	37.86	32.79-43.68	39.47	30.18-44.97
KN	6368	4701-7684	7319	6268-9027	5304	4610-6641	17610	10419-21964	17296	13577-20097
HI	0.25	0.20-0.29	0.29	0.21-0.39	0.36	0.33-0.41	0.28	0.16-0.35	0.23	0.15-0.28
PHT	62.79	47.61-76.23	63.38	53.50-71.85	49.53	40.84-58.79	81.91	75.33-87.51		
NDVI	0.67	0.64-0.70	0.4	0.356-0.449	0.2716	0.25-0.29				
BM	7798	7092-8501	4315	4187-4437	3863	3435-4236	16246	14494-18621	17237	15893-18848

†GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11,Melkassa wet 2011; MLKD11, Melkassa dry 2011.
‡SKD, Single kernel Diameter; SKW, Single Kernel Weight; KNL, Kernel number per spikelet; KWS, Kernel weight per spike; TW, Test Weight; NDVI, Normalized difference Vegetation index; BM, Biomass; SPN, Spikelet number; SL, spike length; KN, kernel number; PHT, Plant Height; DH, Days to heading; TKW, Thousand Kernel weight; HI, Harvest index; DM, Days to maturity; LL, Flag leaf length; LW, Flag leaf with; SN, Spikes number m^{-2} ; LA, Leaf area; KNS, Kernel number spike^{-1}.

Table 3.2. Continued.

	Environments†									
	GRW10		GRW11		GRD11		MLKW11		MLKD11	
Trait‡	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
LL	15.84	13.02-21.42	15.41	13.66-17.77	12.17	10.35-13.95	20.179	18.70-21.63	22.22	19.92-24.52
LW	1.51	1.26-2.02	1.31	1.17-1.39	1.23	1.16-1.32	1.29	1.16-1.37	1.34	1.28-1.58
LS	5.43	3.89-7.62			7.75	6.49-8.53				
KNS			38.69	37.08-40.45	38.71	34.12-45.78				
SPN	16.4	14.78-17.52	16	15.05-17.33	15.8	14.43-17.06	16.58	14.02-18.88		
KWS	1.24	1.05-1.41			0.97	0.90-1.04				
SN			198.94	198.38-199.83	140.62	122.55-178.16				
SL	9.33	7.59-11.33	8.89	7.20-10.89	8.73	6.64-10.62				
KNL	2.12	1.77-2.37	2.41	2.28-2.58	2.38	2.27-2.69				
SKD	2.87	2.69-3.03	2.53	2.35-2.73	2.56	2.36-2.80				
SKW	35.46	30.92-41.44	25.71	21.54-30.63	27.08	22.53-32.10				

†GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11,Melkassa wet 2011; MLKD11, Melkassa dry 2011.
‡SKD, Single kernel Diameter; SKW, Single Kernel Weight; KNL, Kernel number per spikelet; KWS, Kernel weight per spike; TW, Test Weight; NDVI, Normalized difference Vegetation index; BM, Biomass; SPN, Spikelet number; SL, spike length; KN, kernel number; PHT, Plant Height; DH, Days to heading; TKW, Thousand Kernel weight; HI, Harvest index; DM, Days to maturity; LL, Flag leaf length; LW, Flag leaf with; SN, Spikes number m^{-2} ; LA, Leaf area; KNS, Kernel number $spike^{-1}$.

Sequence diversity and SNP-trait association analyses

The amount of genetic variation at the DNA level can be assessed by the average of pairwise nucleotide differences among sequences from different individuals or by the number of segregating sites along the length of DNA sequences (Tajima, 1989). Therefore, in this study nucleotide diversity (π), i.e., the average number of pairwise nucleotide differences per site (Nei, 1987), per site estimates of diversity (θ), and haplotype diversity for each gene were determined (Table 3.3). A total of 37 SNPs with minor allele frequency greater than 5% was detected in the 126 genotypes that were sequenced over a total length of 5038 bp. This is roughly one SNP per 136 bp. Large differences were found among the candidate genes both in number of SNPs and nucleotide diversity parameters. The number of SNPs varied from one in *1-FEH-B* to 16 in *DREB1A.* Similarly, the nucleotide diversity ranged from 0.00078 to 0.18 for *1-FEH-B* and *DREB1A*, respectively (Table 3.3).

Table 3.3. Summary of measures of nucleotide variability in drought tolerance candidate gene sequences.

Gene	Sample n	Length (bp)	SNPs (MAF>0.05)†	Nucleotide diversity (π)	Theta (θ)/site	Haplotype diversity (hd)	Tajima's D	Fu and Li' F* test
DREB1A	126	971	16	0.180	0.392	0.948	-1.809*	-0.61NS
ERA1-B	122	1410	8 (5 INDELS)	0.00094	0.0065	0.508	-2.649***	-6.95**
ERA1-D	121	1388	7	0.0023	0.011	0.826	-2.457***	-5.91**
1-FEH-A	126	601	5	0.00224	0.0035	0.45	-0.896NS	-1.94NS
1-FEH-B	124	668	1	0.00078	0.0049	0.153	-2.307***	-2.52*

*, *** Significant at the 0.05 and 0.001 probability levels, respectively. NS=non-significant.
† MAF, minor allele frequency > 0.05.

With the exception of *DREB1A*, the nucleotide diversity values obtained for the remaining four genes are within the range of nucleotide diversity values (0-0.003) reported for cultivated wheat by Haudry et al. (2007). Except for *DREB1A*, all SNPs for the remaining genes were detected in non-coding regions of the genes (Table 3.4). On average, the number of transversions (nucleotides changes from purine to pyrimidine or vise-versa) was higher than average number of transitions (nucleotide changes from purine to purine or pyrimidine to pyrimidine). However, for some genes (e.g., *DREB1A*) number of transition SNPs was higher than number of transversion SNPs.

Table 3.4. Summary of SNP properties for five drought tolerance candidate genes

Gene	Non-coding SNPs	Coding SNPs		Transitions SNPs	Transversion SNPs
		Synonymous	Nonsynonymous		
DREB1A	10	3	3	9	7
ERA1-B	3	0	0	0	3
ERA1-D	7	0	0	2	5
1-FEH-A	5	0	0	0	5
1-FEH-B	1	0	0	1	0

The effect of selection on the candidate genes was assessed using Tajima's D statistics. The Tajima's D test showed that there was significant difference between π and θ for *ERA1-B, ERA1-D, 1-FEH-B,* and *DREB1A,* indicating that those genes are under selection (Table 3.3). The negative sign for all candidate genes shows that selection has resulted in the accumulation of many low frequency SNPs with respect to predictions of the neutral theory (Fusari et al., 2007; Giordani et al., 2011). However, the estimate of Tajima's D, and Fu and Li's F* test were non-significant for the *1-FEH-A* gene, indicating the absence of a selection footprint for this gene. Except for *DREB1A* gene, the significance test for Tajima's D estimate, and Fu and Li's F* test statistic agreed for the remaining genes (Table 3.3).

The candidate genes also differed in the extent of LD among SNPs. Although large numbers of SNPs were observed for *ERA1-B* and *DREB1A,* the percentages of significant pairwise comparisons among SNPs were higher for *1-FEH-A* (40%) followed by *ERA1-D* (24%) as shown in Table 3.5.

Table 3.5. Linkage disequilibrium (LD) analysis of five drought tolerance candidate genes.

Gene	Length (bp)	Number of informative SNPs (MAF>0.05)†	Number of pairwise comparisons	Number of significant pairwise comparisons (Fisher exact test, $P<0.01$)	% of significant pairwise comparisons	SNP pairs in complete LD
DREB1A	971	16	121	17	14	None
ERA1-B	1410	8 (5 indels)	28	2	7	None
ERA1-D	1388	7	21	5	24	None
1-FEH-A	601	5	10	4	40	3 (30%)
1-FEH-B	668	1	-	-	-	-

†MAF, minor allele frequency

When the recombination rate is low, LD is extended over a large genetic distance and more SNPs in that range would show significant pairwise associations. Except for SNPs within *1-FEH-A*, SNP pairs for other genes varied from weak LD to strong LD (Figure 3.1- 3.4) indicating the inconsistency of LD within a gene region.

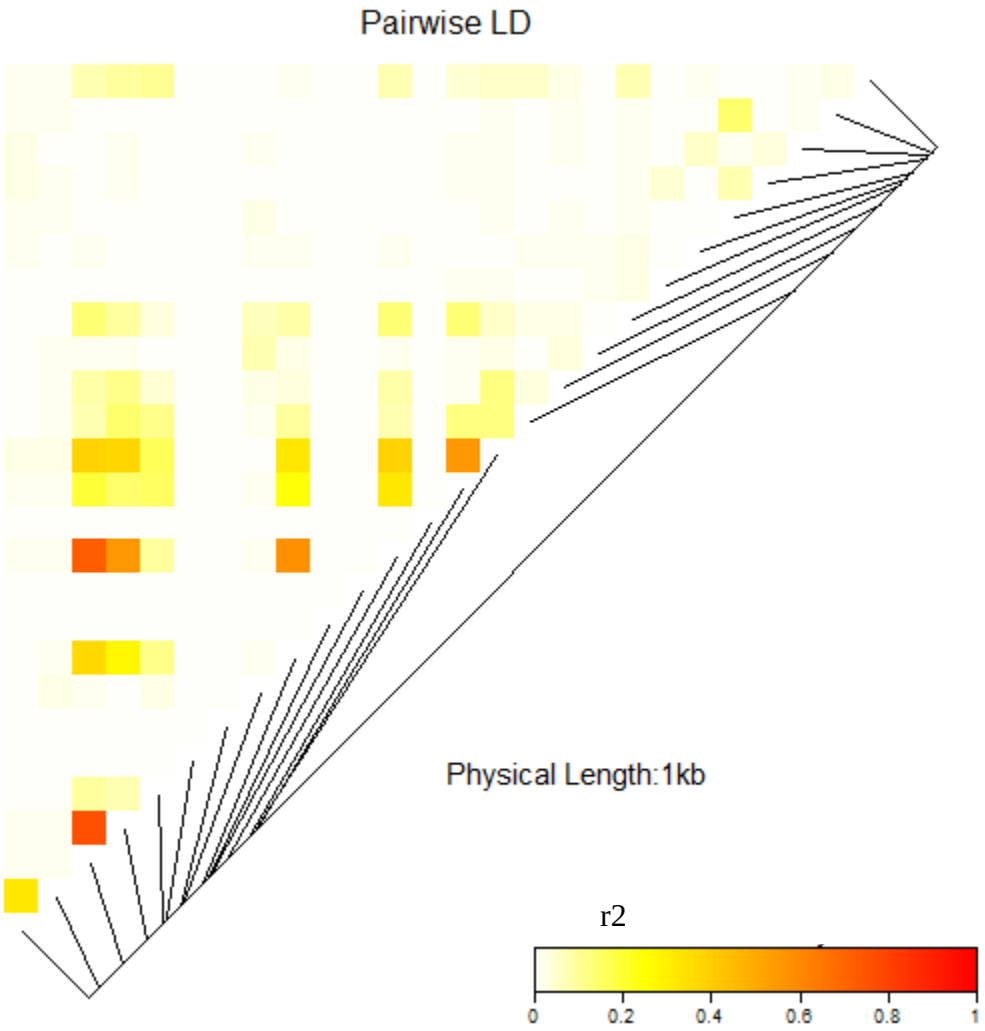

Figure 3.1. Graphical representation of linkage disequilibrium in the *DREB1A* gene.

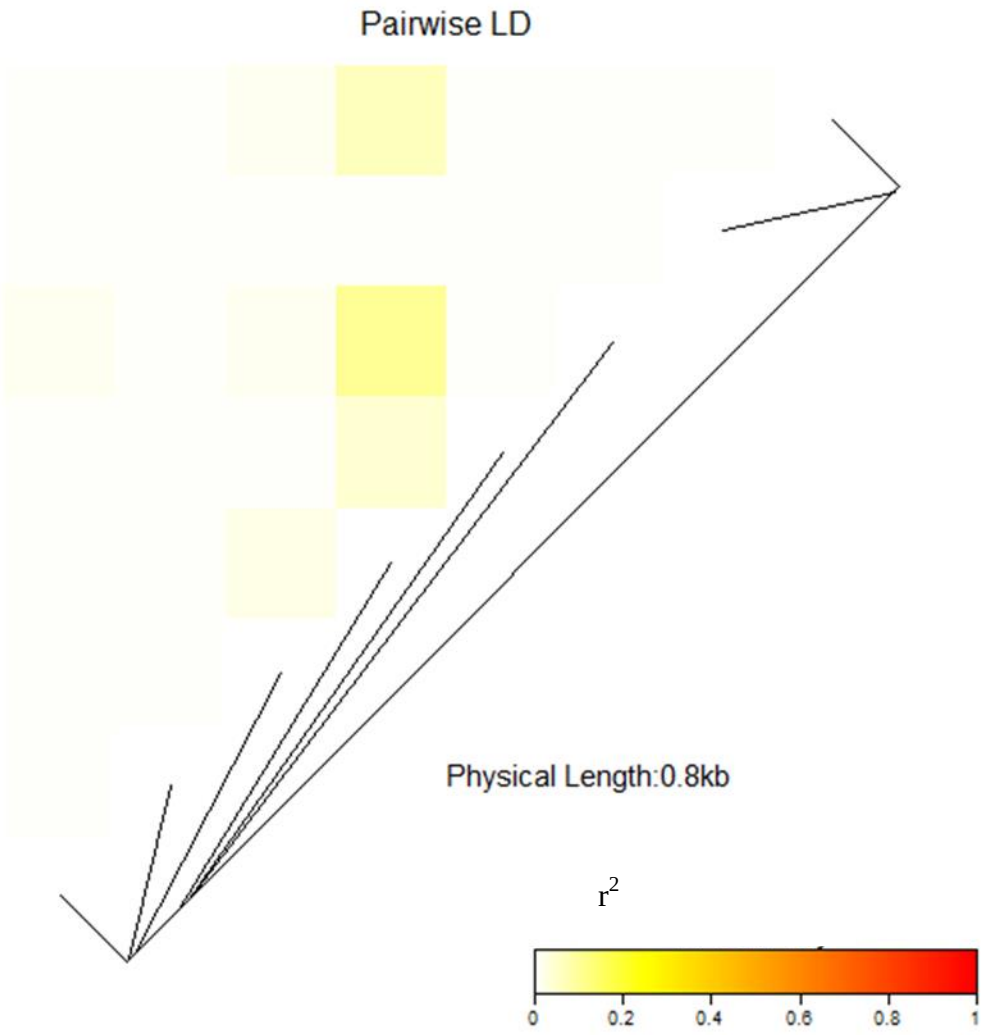

Figure 3.2. Graphical representation of linkage disequilibrium (LD) in the *ERA1-B* gene.

Figure 3.3. Graphical display of single nucleotide polymorphisms (SNPs) within the *ERA1-D* gene.

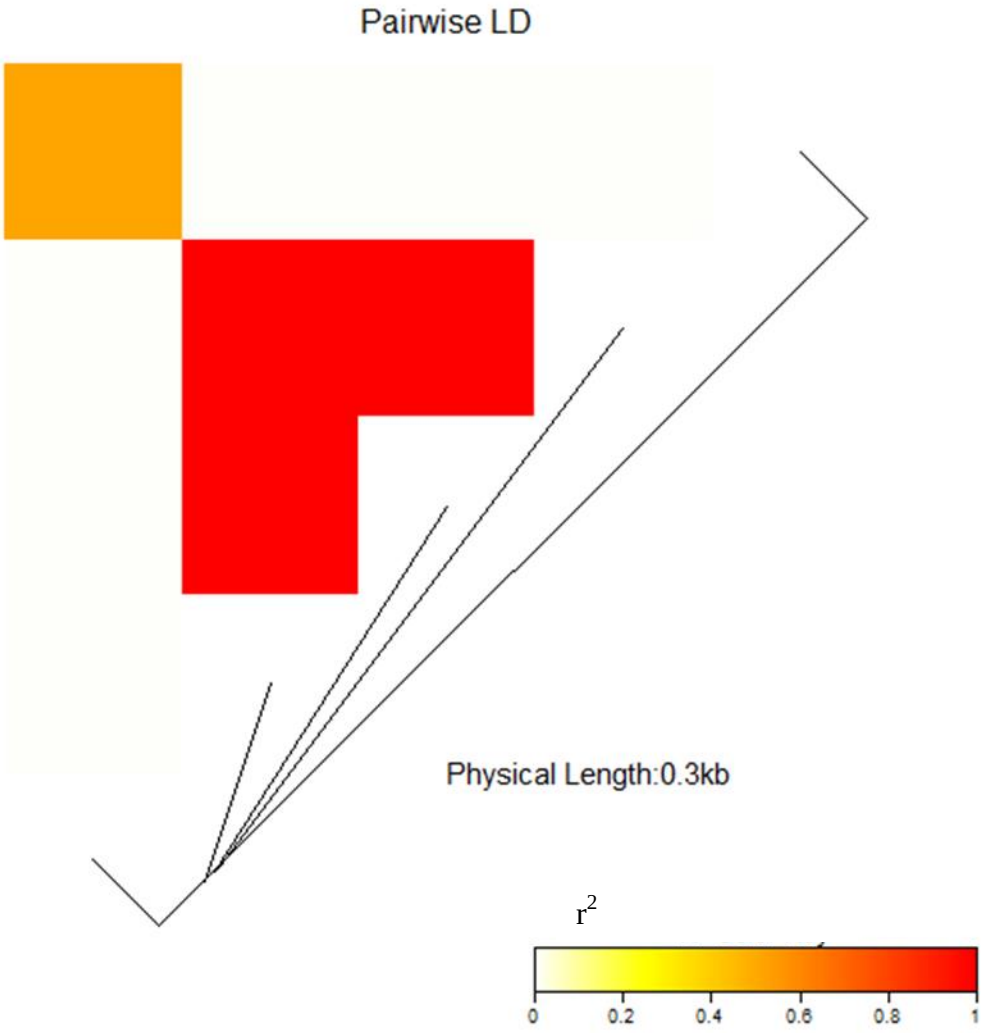

Figure 3.4. Graphical display of single nucleotide polymorphisms (SNPs) within the *1-FEH-A* gene.

The LD decay curves were fitted for chromosomes 3A and 6A using 37 and 53 DArT markers, respectively. The LD decayed below r^2=0.2 at ~3.69 cM for chromosome 3A, on which *DREB1A* is located, while LD decayed below r^2=0.2 at 2.27 cM for chromosome 6A, which harbors *1-FEH-A* (Figure 3.5 and Figure 3.6).

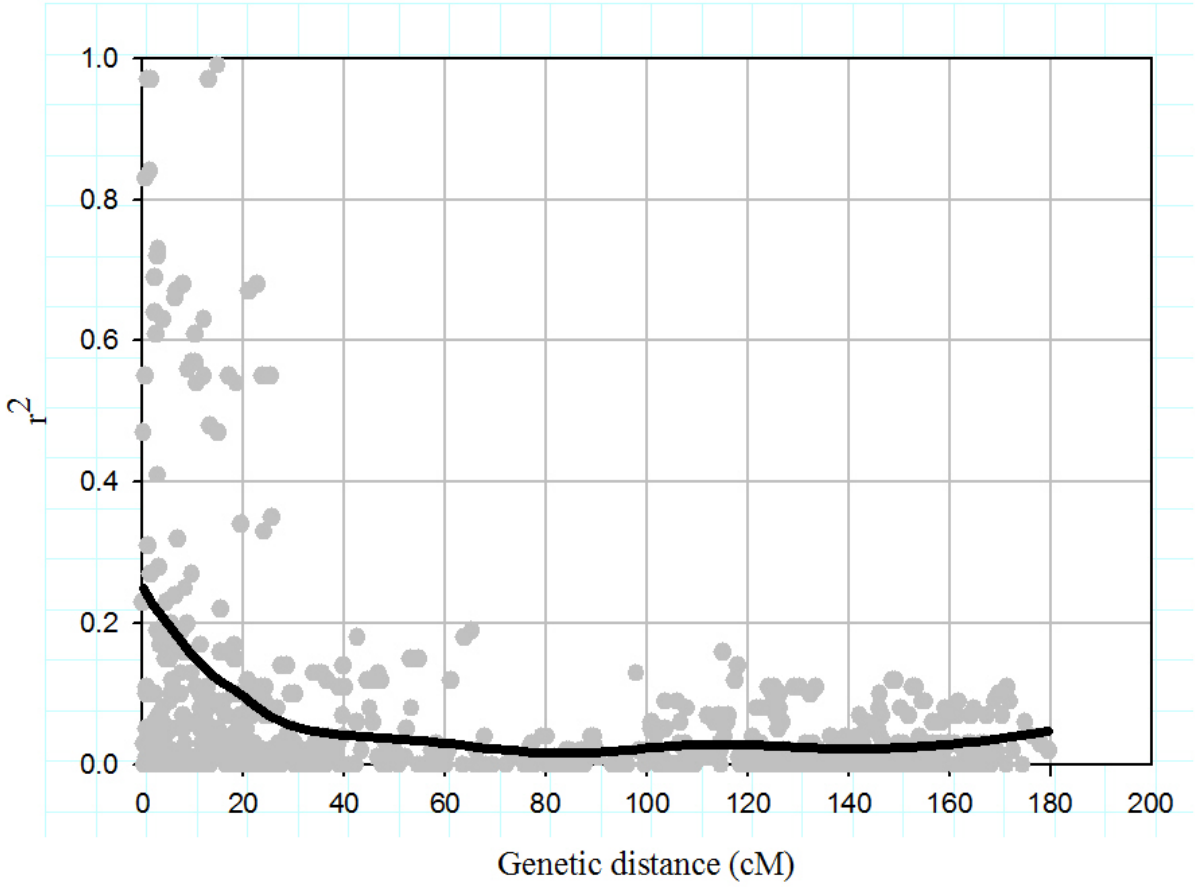

Figure 3.5. Linkage disequilibrium (LD) decay for chromosome 3A of hexaploid wheat.

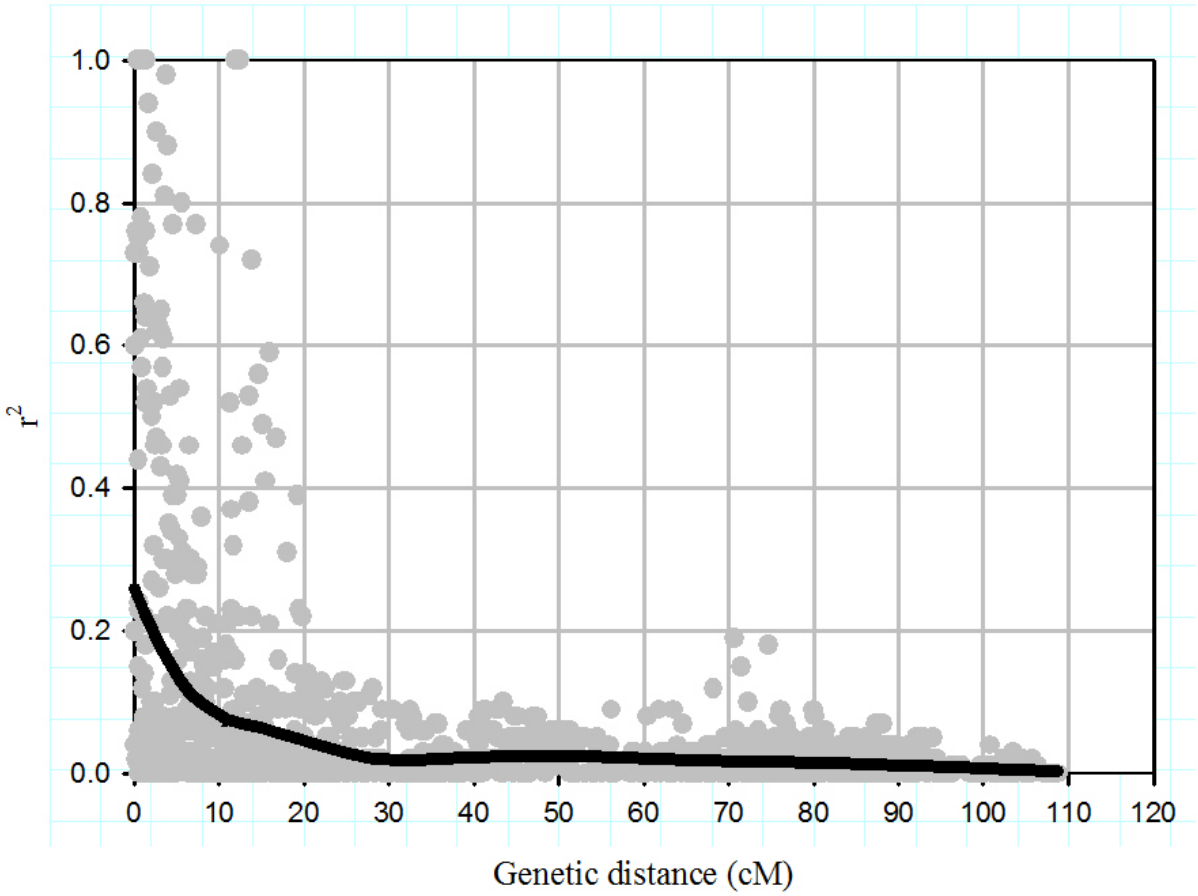

Figure 3.6. Linkage disequilibrium (LD) decay for Chromosome 6A of hexaploid wheat.

The use of functional markers in marker-assisted plant breeding depends on the degree to which economically important traits are affected by a gene. SNPs within *DREB1A* were associated with several traits, including final biomass, normalized vegetation index, days to heading, and spikelet number (Table 3.6). The percentage of phenotypic variation explained by those SNPs ranged from 6.4% for heading date to 9.7% for NDVI. Among the six SNPs detected in the coding region of *DREB1A*, only one SNP (*DREB1A_870*) showed association with a trait (NDVI at GRW10). The change of the nucleotide C to nucleotide T at this site was synonymous substitution, and NDVI mean for lines carried the C nucleotide (common allele) was higher than

that of lines carried the T nucleotide (rare allele) (Figure 3.7). Other SNPs that showed non-synonymous substitution in the coding region of *DREB1A* did not show associations with any phenotypic trait in this study. In association mapping, a QTL that explains about 10% of the phenotypic variation can be considered a major QTL according to definitions for QTL classes suggested by Flint-Garcia et al. (2005). The haplotypes based on combinations of the SNPs within *DREB1A* were also associated with yield and yield component traits such as kernel number m^{-2}, kernel weight, spikelet number, spike length and spike number m^{-2} (data not shown).

Table 3.6. Marker-trait associations for SNPs within five drought tolerance candidate genes and phenotypic traits in individual environments and combined across environments.

Gene	SNP name	Trait	Environments†	$_pFDR^{‡}$	R^2 (%)§
DREB1A	DREB1A_108	Spikelet number	GRW10	0.0518	7.4
	DREB1A_174	Days to heading	Combined data	0.054	7.5
	DREB1A_252	Days to heading	GRW10	0.085	6.9
	DREB1A_252	Days to heading	MLKW11	0.10	6.4
	DREB1A_870	Final biomass	GRD11	0.069	7.9
		NDVI	GRW10	0.014	9.7
ERA1-B	ERA1B_126	Plant height	GRW10	0.067	5.6
	ERA1B_AIN_172	Harvest index	GRW11	0.0378	9.3
	ERA1B_AIN_183	Flag leaf width	MLKW11	0.0046	10.2
	ERA1B_CIN_185	Harvest index	GRW10	0.0599	5.0
		Grain filling duration	GRW10	0.0059	9.7
		Grain filling duration	Combined data	0.044	7.14
		Leaf senescence	GRW10	0.029	6.6
		Spike length	GRW11	0.07024	5.2
	ERA1B_932	Spikes m^{-2}	Combined data	0.0618	6.1
		Spikes m^{-2}	GRW10	0.003	11.3
ERA1-D	ERA1D_235	Flag leaf width	GRD11	0.0331	8.6
	ERA1D_240	Kernel weight per spike	GRD11	0.0259	6.7
		Flag leaf width	GRD11	0.093	3.6
	ERA1D_241	Leaf senescence	GRD11	0.044	6.3
	ERA1D_1203	Kernel number per spike	GRW10	0.048	8.8
	ERA1D_1207	Flag leaf width	GRW10	0.0487	6.45
		Harvest index	GRD11	0.102	4.8
1-FEH-A	FEHA_127	Green leaf area	GRD11	0.064	4.0
		Flag leaf length	GRW10	0.0091	7.3
		Grain yield	GRW10	0.072	5.8
		Flag leaf length	Combined data	0.043	5.8
		Flag leaf area	GRW10	0.055	5.4

Table 3.6 Continued.

Gene	SNP name	Trait	Environments	PFDR‡	R^2 (%)§
	FEHA_145, FEHA_149 and FEHA_151	Spike length	GRW10	0.026	4.7
	FEHA_412	Green leaf area	GRD11	0.064	4.2
		NDVI	GRW10	0.0034	9.8
		Flag leaf length	GRW10	0.0091	7.0
		Final biomass	GRW10	0.0132	6.6
		Grain yield	GRW10	0.0513	4.3
		Kernel number per spike	Combined data	0.0546	4.5
		NDVI	Combined data	0.079	5.0
1-FEH-B	FEH-B-_561	Days to maturity	GRD11	0.0064	5.3
		Thousand Kernel Weight	GRW11	0.034	3.7
		Test weight	GRW11	0.048	3.3
		Days to heading	MLKW11	0.041	4.2

† GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011.
‡ False Discovery Rate adjusted *P*-value.
§ Percent phenotypic variance explained by the SNP.

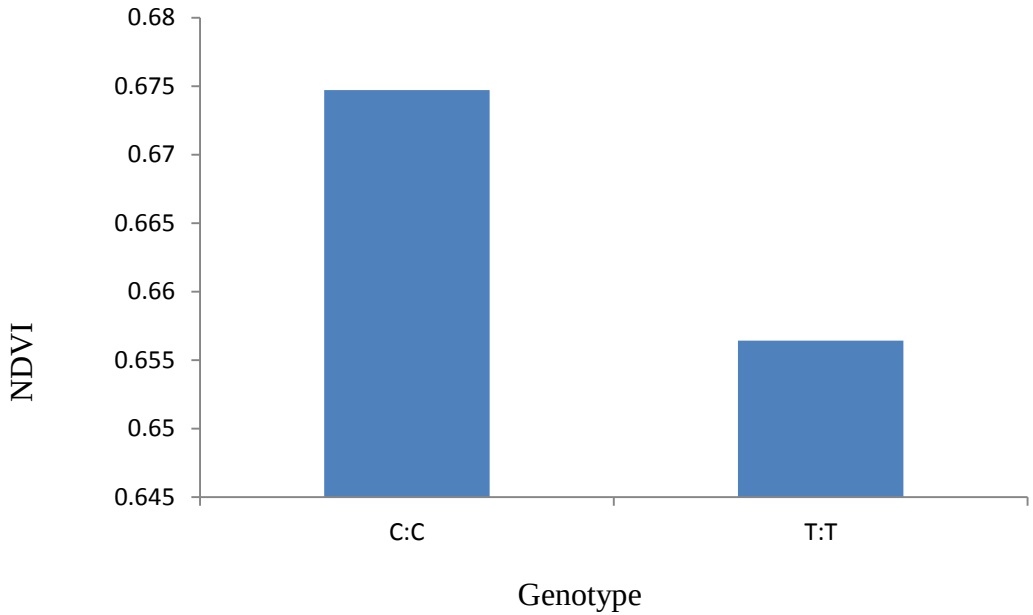

Figure 3.7. Mean of NDVI for two genotypic classes based on SNP (*DREB1A_870*) of *DREB1A* that associated with NDVI evaluated at Greeley under irrigated conditions in 2010

The SNPs within *ERA1-B* were associated with harvest index, spikes m^{-2}, and grain filling duration either in two environments or in one environment plus combined data across environments. SNPs in that gene were associated in a single environment with plant height, leaf senescence, spike length, and leaf width. These SNPs explained the largest phenotypic variation in spike m^{-2} (11.3%) followed by flag leaf width (10.2%), grain filling duration (9.7%), and harvest index (9.3%). Interestingly, the mean of spike m^{-2} of lines carried rare allele was higher than mean of spike m^{-2} for common allele (Figure 3.8). The haplotypes within *ERA1-B* also showed associations with harvest index, grain filling period, kernel weight $spike^{-1}$, and kernel number m^{-2}. *ERA1-D* was also associated with the yield component traits kernel weight $spike^{-1}$, kernel number $spike^{-1}$, harvest index, flag leaf width, and leaf senescence. The majority of the SNP-trait associations for *ERA1-D* were obtained under rainfed conditions. The largest phenotypic variation was explained by the SNP *ERA1D_1203* in *ERA1-D* for kernel number $spike^{-1}$, but the mean of lines carried rare allele was lower than the mean of kernel number spike-1 of common allele at this site (Figure 3.9). The haplotypes within *ERA1-D* were associated with yield, spike length, kernel number m^{-2} and flag leaf width (data not shown). Both *ERA1-B* and *ERA1-D* were associated with leaf senescence, harvest index, and flag leaf width (Table 3.6), suggesting the importance of *ERA1* for drought tolerance in wheat, as some of these traits (e.g., delayed leaf senescence) are related to productivity under dry conditions. Manmathan et al. (2013) recently reported reduced stomatal conductance, increased water use efficiency, and better relative water content in wheat plants silenced for *ERA1* via virus induced gene silencing compared to the control.

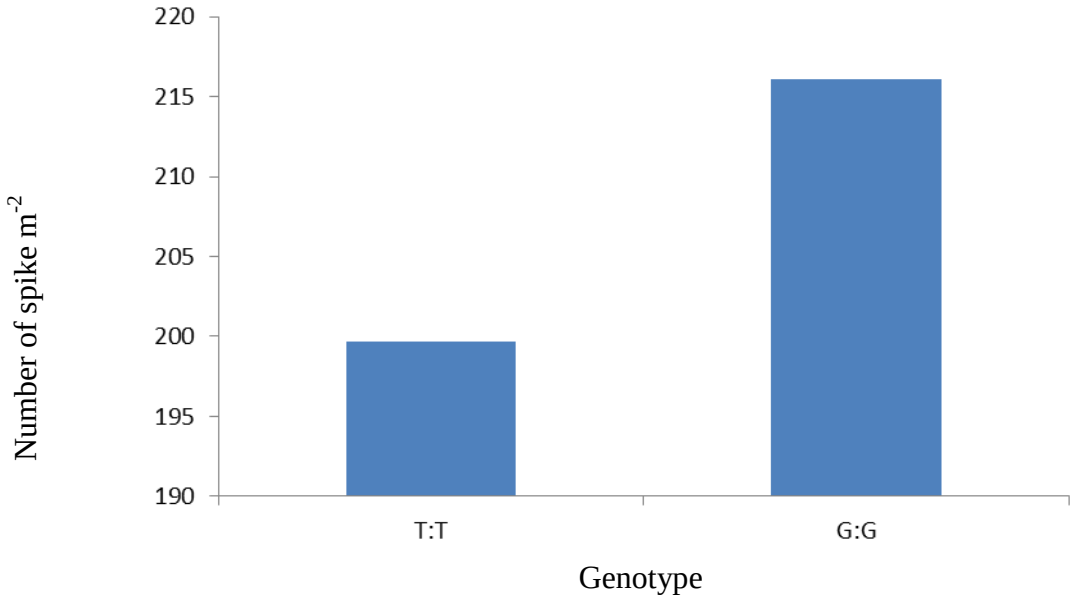

Figure 3.8. Mean of number of spikes m^{-2} for two genotypic classes based on SNP (*ERA1B_932*) of *ERA1-B* that associated with number of spikes m^{-2} evaluated at Greeley under irrigated conditions in 2010.

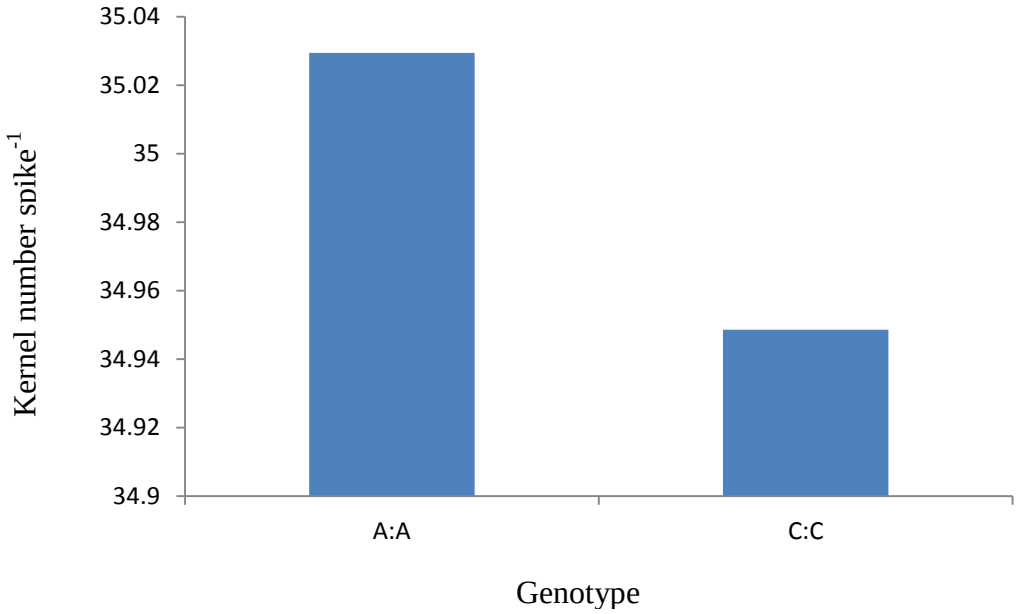

Figure 3.9. Mean of number of kernel number spike^{-1} for two genotypic classes based on SNP (*ERA1D_1203*) of *ERA1-D* that associated with kernel number of spike^{-1} evaluated at Greeley under irrigated conditions in 2010.

The *1-FEH-A* gene was associated with yield, kernel number spike^{-1}, spike length, NDVI, biomass, flag leaf length and area, and green leaf area index (Table 3.6). The SNPs in this gene explained the highest phenotypic variation for NDVI (9.8%) followed by flag leaf length (7.0%) and biomass (6.6%). The phenotypic mean of NDVI for the lines carried rare allele was lower than that of common allele (Figure 3.10). The *1-FEH-A* haplotype was also associated with yield, biomass, NDVI, plant height, days to heading, leaf length, spike length, and green leaf area index (data not shown). Only a single SNP was detected for *1-FEH-B* and this SNP was associated with days to maturity, kernel weight, test weight and days to heading (Table 3.6). The phenotypic mean of lines carried rare allele was lower that of common allele for thousand kernel weight (Figure 3.11). In genome-wide association analysis with DArT markers for the complete panel described here (n=294), we detected QTL on chromosome 6AS, where *1 FEH-A* resides, for several traits, including thousand kernel weight, plant height, flag leaf area and width. A previous bi-parental QTL mapping study detected QTL for stem water soluble carbohydrate, thousand kernel weight, and grain filling efficiency on chromosome 6AS (Yang et al., 2007). Thus, our results, supported by the previous study, suggest that *1-FEH* genes are associated with yield-related traits that are important in both irrigated and rainfed conditions.

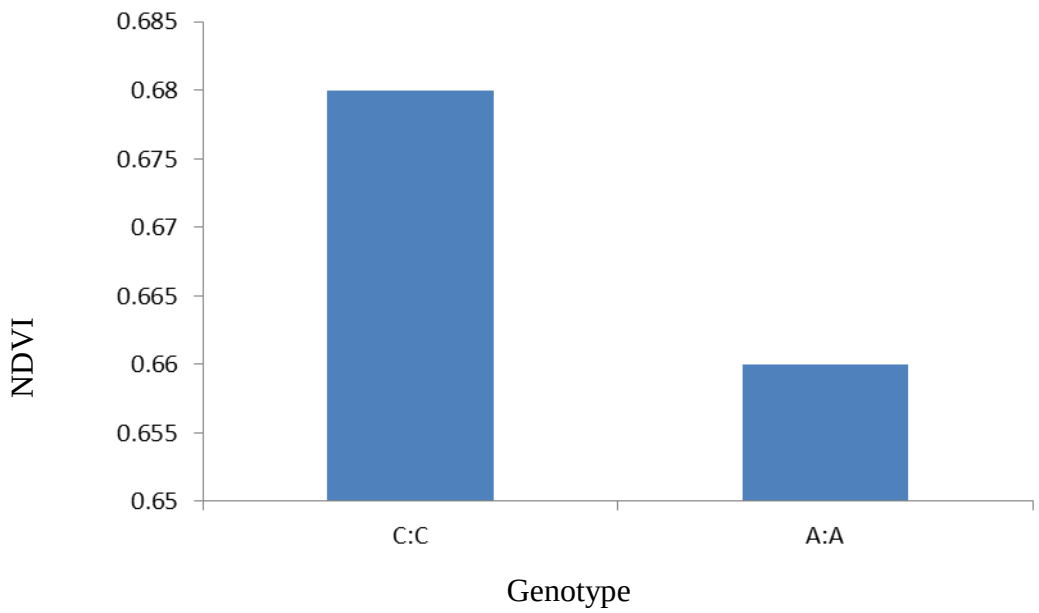

Figure 3.10. Mean of NDVI for two genotypic classes based on SNP (*1-FEHA_412*) of *1-FEH-A* that associated with NDVI data obtained from Greeley under irrigated conditions in 2010.

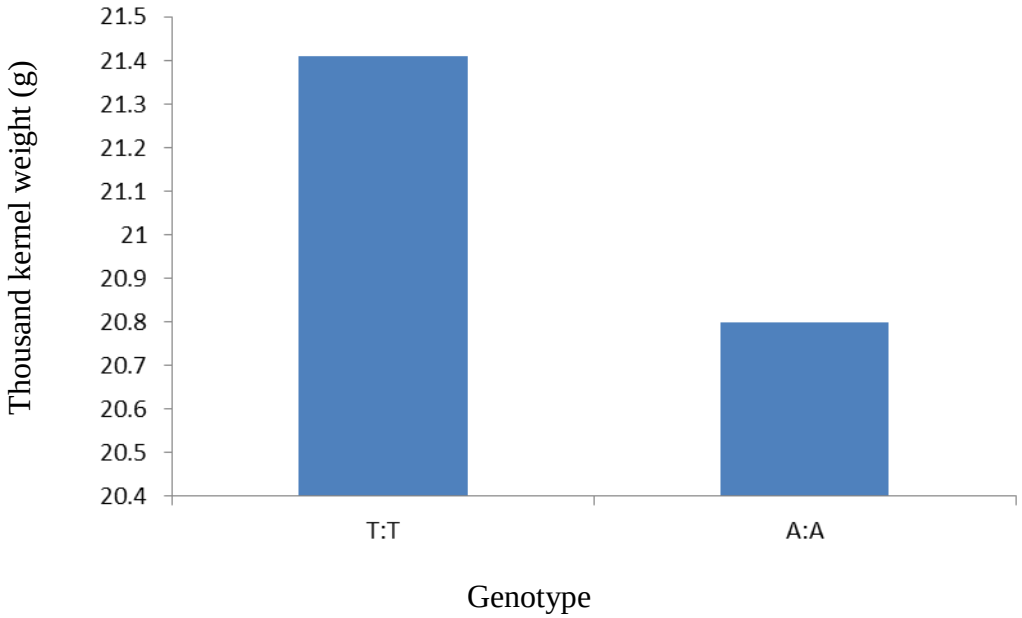

Figure 3.11. Mean of number of thousand kernel weight for two genotypic classes based on SNP (*1-FEH-B_561*) of *1-FEH-B* that associated with thousand kernel weight evaluated at Greeley under irrigated conditions in 2011.

Very few SNP-trait associations were detected for the Melkassa environments, and we are uncertain of the reason for this. One explanation may be the wide range of heading dates at that location (15 to 18 days), which may have confounded the effects of the candidate genes on the yield-related traits. Another factor is that the number of traits evaluated in Melkassa was less than the number evaluated in Greeley, so there were fewer opportunities to detect significant associations.

3.3 DISCUSSION

Information on nucleotide diversity and SNP density is very rare for hexaploid wheat. To our knowledge, this study is the first report on nucleotide diversity for drought tolerance genes of hexaploid wheat.

Orthologous genes on homoeologous chromosomes showed differences in almost all diversity parameters considered here, including nucleotide diversity, haplotype diversity, Tajima's D values, Fu and Li's test, and LD patterns. *ERA1-B* and *ERA1-D* were amplified from homoeologous chromosomes on the B and D genomes of hexaploid wheat, respectively. However, *ERA1-B* is less diverse than *ERA1-D* based on these diversity parameters. This is an unexpected result because both the A and B genomes of wheat are more diverse than the D genome (Chao et al., 2010) based on differences in LD decay rate among the genomes. In the current study also, the percentage of SNP pairs in LD for *ERA1-D* is higher than that of SNP pairs in LD for *ERA1-B* (Table 3.5). Similarly, *1-FEH-A* is more diverse than *1-FEH-B* and selection impact is high on the latter. Selection resulted in accumulation of rare SNPs with frequency <5% for *1-FEH-B*, while the number of rare alleles is in agreement with neutral expectation for *1-FEH-A*. Tajima's test may not provide complete information about the action

of selection as demographic processes such as population bottleneck, recombination, population structure, and sample size can bias the results (Figueiredo et al., 2010). We are unable to compare the extent of LD between these genes as only one SNP showed minor allele frequency >5% for *1-FEH-B*. However, *1-FEH-A* had a higher number of SNP pairs with significant LD of all the candidate genes in this study.

The average number of SNPs within a gene varies depending on the species, region of a chromosome, and selection pressure. Although the SNP frequency is greatly gene dependent, the average of one SNP per 136 bp obtained in this study is higher than SNP density reported for wheat by Ravel et al. (2007). However, it is far less than SNP density reported for other crops such as maize (1 SNP/104 bp), sorghum (*Sorghum bicolor*) (1 SNP/123 bp), sunflower (*Helianthus annuus*) (1 SNP/69 bp), and rice (1 SNP/113 bp and 1 SNP/100 bp) (Fusari et al., 2008). The presence of low genetic diversity in hexaploid wheat is partly explained by low effective recombination, as wheat is highly self-pollinated. In addition to this, both domestication and modern breeding for high yield and disease resistance have reduced genetic diversity in wheat (Reif et al., 2005; Akhunov et al., 2010).

In the context of our current study, LD is a non-random association of polymorphic sites (SNPs) within a gene. Graphical displays of LD (Figure 3.1- 3.4) in terms of r^2 showed the patterns of association among polymorphic sites within all tested genes. The orthologous genes amplified from different genomes of hexaploid wheat showed different LD patterns. This information is useful in deciding how many functional markers need to be developed per gene, as the degree of associations of SNPs within a gene is different for different SNPs. The chromosome-wide LD analysis with DArT markers also confirmed differences in the extent of LD among chromosomes harboring the candidate genes as expected. On average, LD decays

faster for chromosome 6A than chromosome 3A, implying a better chance of tagging *DREB1A* with linked genome-wide markers than the *1-FEH-A* gene provided that there is no change in relationship of average LD decay rates around the two genes.

In this study we found that SNPs that reside within a few base pairs were associated with different traits. A potential weakness of genome-wide QTL scanning is the possibility of over-looking SNPs at a locus that may be associated with a trait of interest, because QTL regions may not be represented with enough markers (Haseneyer et al., 2010).

Although previous reports indicated that the five drought tolerance candidate genes are stress-induced and confer drought tolerance under stress conditions, SNP-trait associations were detected both under dry and irrigated conditions for all genes in this study. It is possible that even the trials grown under wetter conditions experienced some degree of moisture stress, thereby inducing expression of the evaluated genes. Most of the detected associations were significant only in a single environment, which is consistent with the high level of genotype by environment interaction that occurred in this study. Therefore, the advantage of these genes for yield or drought tolerance will depend on variable environmental conditions, as the genes may show different expression patterns in different environments (Wei et al., 2009; Mochida et al., 2003).

All drought tolerance candidate genes showed associations with yield and yield components, morphological and phenological traits both at individual SNP and haplotype levels. The genes explained substantial amounts of phenotypic variation for yield component traits (e.g., spikes m^{-2}), morphological traits (e.g., flag leaf width), and drought tolerance-related indices (e.g., NDVI). However, before the SNPs identified in this study are converted into functional markers for use in breeding, confirmation of their benefits is needed.

In conclusion, gene sequence variability analysis of hexaploid wheat indicated the presence of sufficient polymorphic sites in the evaluated genes for development of functional markers. The homoeologous genes on different wheat genomes showed clear differences in nucleotide diversity, LD patterns, and SNP-trait associations. Since gene copies on different homoeologous chromosomes showed different SNP-trait associations, the development of functional markers requires consideration of the economic importance of a trait and the amount of phenotypic variation explained by each gene copy. Future research on *DREB1A*, *ERA1*, and *1-FEH* should validate the relative importance of the orthologous genes in different genetic backgrounds across a range of moisture conditions.

REFERENCES

Abdurakhmonov I., Abdukarimov A. (2008). Application of association mapping to understandingthe genetic diversity of plant germplasm resources. International Journal of Plant Genomics. doi:10.1155/2008/574927.

Adhikari T.B., Gurung S., Hansen J.M., Jackson E.W., Bonman J.M. (2012) Association mapping of quantitative trait loci in spring wheat landraces conferring resistance to bacterial leaf streak and spot blotch. Plant Genome 5:1-16. DOI: DOI 10.3835/plantgenome2011.12.0032.

Agrama H.A., Eizenga G.C., Yan W. (2007) Association mapping of yield and its components in rice cultivars. Molecular Breeding 19:341-356. DOI: DOI 10.1007/s11032-006-9066-6.

Ahuja I., de Vos R.C.H., Bones A.M., Hall R.D. (2010) Plant molecular stress responses face climate change. Trends in Plant Science 15:664-674. DOI: DOI 10.1016/j.tplants.2010.08.002.

Akbari M., Wenzl P., Caig V., Carling J., Xia L., Yang S.Y., Uszynski G., Mohler V., Lehmensiek A., Kuchel H., Hayden M.J., Howes N., Sharp P., Vaughan P., Rathmell B., Huttner E., Kilian A. (2006) Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. Theoretical and Applied Genetics 113:1409-1420. DOI: DOI 10.1007/s00122-006-0365-4.

Akhunov E., Nicolet C., Dvorak J. (2009) Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. Theoretical and Applied Genetics 119:507-517. DOI: DOI 10.1007/s00122-009-1059-5.

Akhunov E.D., Akhunova A.R., Anderson O.D., Anderson J.A., Blake N., Clegg M.T., Coleman-Derr D., Conley E.J., Crossman C.C., Deal K.R., Dubcovsky J., Gill B.S., Gu Y.Q., Hadam J., Heo H., Huo N.X., Lazo G.R., Luo M.C., Ma Y.Q., Matthews D.E., McGuire P.E., Morrell P.L., Qualset C.O., Renfro J., Tabanao D., Talbert L.E., Tian C., Toleno D.M., Warburton M.L., You F.M., Zhang W.J., Dvorak J. (2010) Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. Bmc Genomics 11. DOI: Artn 702 Doi 10.1186/1471-2164-11-702.

Andersen J., Lubberstedt T. (2003) Functional markers in plants. TRENDS in Plant Science 18:554-560.

Aranzana M.J., Kim S., Zhao K.Y., Bakker E., Horton M., Jakob K., Lister C., Molitor J., Shindo C., Tang C.L., Toomajian C., Traw B., Zheng H.G., Bergelson J., Dean C., Marjoram P., Nordborg M. (2005) Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. PLOS Genetics 1:531-539. DOI: ARTN e60 DOI 10.1371/journal.pgen.0010060.

Araus J.L., Slafer G.A., Royo C., Serret M.D. (2008) Breeding for Yield Potential and Stress Adaptation in Cereals. Critical Reviews in Plant Sciences 27:377-412.

Babar M.A., Reynolds M.P., Van Ginkel M., Klatt A.R., Raun W.R., Stone M.L. (2006) Spectral reflectance to estimate genetic variation for in-season biomass, leaf chlorophyll, and canopy temperature in wheat. Crop Science 46:1046-1057. DOI: DOI 10.2135/cropsci2005.0211.

Bagge M., Lubberstedt T. (2008) Functional markers in wheat: technical and economic aspects. Molecular Breeding 22:319-328. DOI: DOI 10.1007/s11032-008-9190-6.

Bagge M., Xia X.C., Lubberstedt T. (2007) Functional markers in wheat - Commentary. Current Opinion in Plant Biology 10:211-216. DOI: DOI 10.1016/j.pbi.2007.01.009.

Bao L., Wei L., Peirce J.L., Homayouni R., Li H., Zhou M., Chen H., Lu L., Williams R.W., Pfeffer L.M., Goldowitz D., Cui Y. (2006) Combining gene expression QTL mapping and phenotypic spectrum analysis to uncover gene regulatory relationships. Mammalian Genome 17:575-83. DOI: 10.1007/s00335-005-0172-2.

Barnabas B., Jager K., Feher A. (2008) The effect of drought and heat stress on reproductive processes in cereals. Plant Cell and Environment 31:11-38. DOI: DOI 10.1111/j.1365-3040.2007.01727.x.

Bayeh B. (2010) Assessment of Bread Wheat Production, Marketing and Selection of N-Efficient Bread Wheat (Tritium aestivum L.) Varieties for Higher Grain Yield and Quality in North Western Ethiopia. M.S. Thesis, Bahir Dar University, Ethiopia.

Beales J., Turner A., GriYths S., Snape J.W., Laurie D.A. (2007) A Pseudo-response regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 115:721-733. DOI: DOI 10.1007/s00122-007-0603-4.

Benjamini, Y., Hochberg Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 57: 289–300.

Benson J., Brown-Guedira G., Murphy J.P., Sneller C. (2012) Population Structure, Linkage Disequilibrium, and Genetic Diversity in Soft Winter Wheat Enriched for Fusarium Head Blight Resistance. Plant Genome 5:71-80. DOI: DOI 10.3835/plantgenome2011.11.0027.

Bernardo R. (1993) Estimation of coefficient of co-ancestry using molecular markers in maize. Theoretical and Applied Genetics 85: 1055–1062.

Berkman P.J., Visendi P., Lee H. C., Stiller J., Manoli S., Lorenc M. T., Lai K., Batley J., Fleury D., Simkov H., Kubalakov M., Weining S., Dolezel J., Edwards D. (2013) Dispersion and domestication shaped the genome of bread wheat. Plant Biotechnology Journal 11: 564–571.

Bertin P., Gallais A. (2001) Genetic variation of nitrogen use efficiency in a set of recombinant inbred lines. II. QTL detection and coincidences. Maydica 46:53–68.

Bhatnagar-Mathur P., Devi J.M., Lavanya M., Reddy D.S., Vadez V., Serraj R., Yamaguchi-Shinozaki K., Sharma K.K. (2007) Stress-inducible expression of at DREB1A in transgenic peanut (*Arachis hypogaea* L.) increases transpiration efficiency under water limiting conditions. Plant Cell Reports 26:2071–2082.

Blake N.K., Lanning S.P., Martin J.M., Sherman J.D., Talbert L.E. (2007) Relationship of flag leaf characteristics to economically important traits in two spring wheat crosses. Crop Science 47:491-496. DOI: DOI 10.2135/cropsci2006.05.0286.

Blum A. (2005) Drought resistance, water-use efficiency, and yield potential - are they compatible, dissonant, or mutually exclusive? Australian Journal of Agricultural Research 56:1159-1168. DOI: Doi 10.1071/Ar05069.

Blum A. (2011) Drought resistance - is it really a complex trait? Functional Plant Biology 38:753-757. DOI: Doi 10.1071/Fp11101.

Blum A., Shpiler L., Golan G., Mayer J. (1989) Yield Stability and Canopy Temperature of Wheat Genotypes under Drought-Stress. Field Crops Research 22:289-296. DOI: Doi 10.1016/0378-4290(89)90028-2.

Bonnet D., Smith J.B. (1976) Nuclear DNA amounts in angiosperms. Phil. Trans. Roy. Soc. Lond. B. 274:227-274.

Borner A., Schumann E., Furste A., Coster H., Leithold B., Roder M.S., Weber W.E. (2002) Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 105:921-936. DOI: DOI 10.1007/s00122-002-0994-1.

Borner A., Worland A.J. (2002) Does the Chinese dwarf wheat variety 'XN0004' carry Rht21? Cereal Research Communications 30:25-29.

Bradbury P.J., Zhang Z., Kroon D.E., Casstevens T.M., Ramdoss Y., Buckler E.S. (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633-2635. DOI: DOI 10.1093/bioinformatics/btm308.

Breseghello F., Sorrells M.E. (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172:1165-1177. DOI: DOI 10.1534/genetics.105.044586.

Bryan G.J., Collins A.J., Stephenson P., Orry A., Smith J.B., Gale M.D. (1997) Isolation and characterisation of microsatellites from hexaploid bread wheat. Theoretical and Applied Genetics 94:557-563. DOI: DOI 10.1007/s001220050451.

Cadalen T., Boeuf C., Bernard S., Bernard M. (1997) An inter varietal molecular marker map in *Triticum aestivum* L and comparison with a map from a wide cross. Theoretical and Applied Genetics 94:367-377. DOI: DOI 10.1007/s001220050425.

Campbell K.G., Bergman C.J., Gualberto D.G., Anderson J.A., Giroux M.J., Hareland G., Fulcher R.G., Sorrells M.E., Finney P.L. (1999) Quantitative trait loci associated with kernel traits in a soft x hard wheat cross. Crop Science 39:1184-1195.

Casadesus J., Kaya Y., Bort J., Nachit M.M., Araus J.L., Amor S., Ferrazzano G., Maalouf F., Maccaferri M., Martos V., Ouabbou H., Villegas D. (2007) Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in water-limited environments. Annals of Applied Biology 150:227-236. DOI: DOI 10.1111/j.1744-7348.2007.00116.x.

Cattivelli L., Rizza F., Badeck F.W., Mazzucotelli E., Mastrangelo A.M., Francia E., Mare C., Tondelli A., Stanca A.M. (2008) Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. Field Crops Research 105:1-14. DOI: DOI 10.1016/j.fcr.2007.07.004.

Chao S.M., Dubcovsky J., Dvorak J., Luo M.C., Baenziger S.P., Matnyazov R., Clark D.R., Talbert L.E., Anderson J.A., Dreisigacker S., Glover K., Chen J.L., Campbell K., Bruckner P.L., Rudd J.C., Haley S., Carver B.F., Perry S., Sorrells M.E., Akhunov E.D. (2010) Population- and genome-specific patterns of linkage disequilibrium and SNP

variation in spring and winter wheat (*Triticum aestivum* L.). Bmc Genomics 11. DOI: Artn 727 Doi 10.1186/1471-2164-11-727.

Chao S.M., Zhang W.J., Akhunov E., Sherman J., Ma Y.Q., Luo M.C., Dubcovsky J. (2009) Analysis of gene-derived SNP marker polymorphism in US wheat (*Triticum aestivum* L.) cultivars. Molecular Breeding 23:23-33. DOI: DOI 10.1007/s11032-008-9210-6.

Chao S.M., Zhang W.J., Dubcovsky J., Sorrells M. (2007) Evaluation of genetic diversity and genome-wide linkage disequilibrium among US wheat (*Triticum aestivum* L.) germplasm representing different market classes. Crop Science 47:1018-1030. DOI: DOI 10.2135/cropsci2006.06.0434.

Chen F., Liu S., Zhao F., Xu C., Xia G. (2010) Molecular characterization of the low-molecular weight glutenin subunit genes of tall wheatgrass and functional properties of one clone Ee34. Amino Acids 38:991-9. DOI: 10.1007/s00726-009-0307-y.

Chen Y.H., Carver B.F., Wang S.W., Cao S.H., Yan L.L. (2010) Genetic regulation of developmental phases in winter wheat. Molecular Breeding 26:573-582. DOI: DOI 10.1007/s11032-010-9392-6.

Ching A., Caldwell K.S., Jung M., Dolan M., Smith O.S., Tingey S., Morgante M., Rafalski A.J. (2002) SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. Bmc Genetics 3. DOI: Artn 19 Doi 10.1186/1471-2156-3-19.

Chu C.G., Xu S.S., Friesen T.L., Faris J.D. (2008) Whole genome mapping in a wheat doubled haploid population using SSRs and TRAPs and the identification of QTL for agronomic traits. Molecular Breeding 22:251-266. DOI: DOI 10.1007/s11032-008-9171-9.

Cleveland W. S. (1979) “Robust locally-weighted regression and smoothing scatterplots”. Journal of the American Statistical Association, 74, 829-836.

Collins N.C., Shirley N.J., Saeed M., Pallotta M., Gustafson J.P. (2008) An ALMT1 gene cluster controlling aluminum tolerance at the Alt4 locus of rye (Secale cereale L.). Genetics 179:669-682. DOI: DOI 10.1534/genetics.107.083451.

Comadran J., Thomas W.T.B., van Eeuwijk F.A., Ceccarelli S., Grando S., Stanca A.M., Pecchioni N., Akar T., Al-Yassin A., Benbelkacem A., Ouabbou H., Bort J., Romagosa I., Hackett C.A., Russell J.R. (2009) Patterns of genetic diversity and linkage disequilibrium in a highly structured *Hordeum vulgare* association-mapping population for the Mediterranean basin. Theoretical and Applied Genetics 119:175-187. DOI: DOI 10.1007/s00122-009-1027-0.

Cook J.P., McMullen M.D., Holland J.B., Tian F., Bradbury P., Ross-Ibarra J., Buckler E.S., Flint-Garcia S.A. (2012) Genetic Architecture of Maize Kernel Composition in the Nested Association Mapping and Inbred Association Panels. Plant Physiology 158:824-834. DOI: DOI 10.1104/pp.111.185033.

Cooper M., van Eeuwijk F.A., Hammer G.L., Podlich D.W., Messina C. (2009) Modeling QTL for complex traits: detection and context for plant breeding. Current Opinion in Plant Biology 12:231-240. DOI: DOI 10.1016/j.pbi.2009.01.006.

Crossa J., Burgueno J., Dreisigacker S., Vargas M., Herrera-Foessel S.A., Lillemo M., Singh R.P., Trethowan R., Warburton M., Franco J., Reynolds M., Crouch J.H., Ortiz R. (2007) Association analysis of historical bread wheat germplasm using additive genetic

covariance of relatives and population structure. Genetics 177:1889-1913. DOI: DOI 10.1534/genetics.107.078659.

Cuthbert J.L., Somers D.J., Brule-Babel A.L., Brown P.D., Crow G.H. (2008) Molecular mapping of quantitative trait loci for yield and yield components in spring wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 117:595-608. DOI: DOI 10.1007/s00122-008-0804-5.

Cutler S., Ghassemian M., Bonetta D., Cooney S., McCourt P. (1996) A protein farnesyl transferase involved in abscisic acid signal transduction in Arabidopsis. Science 273:1239–1241.

Daud H.M. and Gustafson J.P. (1996) Molecular evidence for Triticum speltoides as a B-genome progenitor of wheat (Triticum aestivum). Genome 39:543-548. DOI: Doi 10.1139/G96-069.

Devos K.M. and Gale M.D. (1997) Comparative genetics in the grasses. Plant Molecular Biology 35: 3-15.

Dhuyvetter K.C., Thompson C.R., Norwood C.A., Halvorson A.D. (1996) Economics of dryland cropping systems in the great plains: A review. Journal of Production Agriculture 9:216-222.

Dodig D., Zoric M., Kobiljski B., Savic J., Kandic V., Quarrie S., Barnes J. (2012) Genetic and Association Mapping Study of Wheat Agronomic Traits Under Contrasting Water Regimes. International Journal of Molecular Sciences 13:6167-6188. DOI: Doi 10.3390/Ijms13056167.

Dolferus R., Ji X.M., Richards R.A. (2011) Abiotic stress and control of grain number in cereals. Plant Science 181:331-341. DOI: DOI 10.1016/j.plantsci.2011.05.015.

Dubcovsky J., Loukoianov A., Fu D.L., Valarik M., Sanchez A., Yan L.L. (2006) Effect of photoperiod on the regulation of wheat vernalization genes VRN1 and VRN2. Plant Molecular Biology 60:469-480. DOI: DOI 10.1007/s11103-005-4814-2.

Duggan B.L., Richards R.A., van Herwaarden A.F. (2005) Agronomic evaluation of a tiller inhibition gene (tin) in wheat. II. Growth and partitioning of assimilate. Australian Journal of Agricultural Research 56:179-186. DOI: Doi 10.1071/Ar04153.

El-Feki W. (2010) Mapping quantitative trait loci for bread making quality and agronomic traits in winter wheat under different soil moisture levels. Ph.D. dissertation, Colorado State University, U.S.A.

Ellis M.H., Bonnett D.G., Rebetzke G.J. (2007) A 192bp allele at the Xgwm261 locus is not always associated with the Rht8 dwarfing gene in wheat (*Triticum aestivum* L.). Euphytica 157:209-214. DOI: DOI 10.1007/s10681-007-9413-7.

Ersoz E., Yu J., Buckler E. (2009) Applications of Linkage Disequilibrium and Association Mapping in Maize. A.L. Kriz, B.A. Larkins (eds.), Molecular Genetic Approaches to Maize Improvement 173. Biotechnology in Agriculture and Forestry, Vol. 63, Pp 173-195.

Evanno G., Regnaut S., Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620. DOI: DOI 10.1111/j.1365-294X.2005.02553.x.

Farooq M., Bramley H., Palta J., Siddique K. (2011) Heat Stress in Wheat during Reproductive and Grain-filling Phases. Critical Reviews in Plant Sciences, 30:491–507.

Fereres E., Gonzalez-Dugo V. (2009) Improving productivity to face water scarcity in irrigated agriculture. *Sadras V. and Calderini D. (eds*.). Crop Physiology, Applications for Crop Improvement and Agronomy. Pp 123-143.

Fereres E., Soriano M.A. (2007) Deficit irrigation for reducing agricultural water use. Journal of Experimental Botany 58:147-159. DOI: Doi 10.1093/Jxb/Erl165.

Figueiredo M.L., Wentworth K.M., Sandgren E.P. (2010) Quantifying growth and ransformation frequency of oncogene-expressing mouse hepatocytes in vivo. Hepatology 52:634-43. DOI: 10.1002/hep.23682.

Figueiredo A. F., Sine B., Chantereau J., Mestres C., Fliedel G., Rami J. F., Glaszmann J. C., Deu M., Courtois B. (2010) Variability of grain quality in sorghum: association with polymorphism in Sh2, Bt2, SssI, Ae1, Wx and O2. Theoretical Applied Genetics 121: 1171-1185.

Fisher R.A., Maurer R. (1978) Drought resistance in spring wheat cultivars: I. Grain yield responses. Australian Journal of Agricultural Research 29:897-912.

Fleury D., Jefferies S., Kuchel H., Langridge P. (2010) Genetic and genomic tools to improve drought tolerance in wheat. Journal of Experimental Botany 61:3211-3222. DOI: Doi 10.1093/Jxb/Erq152.

Flint-Garcia S.A., Thornsberry J.M., Buckler E.S. (2003) Structure of linkage disequilibrium in plants. Annual Review of Plant Biology 54:357-374. DOI: DOI 10.1146/annurev.arplant.54.031902.134907.

Flint-Garcia S.A, Thuillet A.C., Yu J., Pressoir G., Romero S.M., Mitchell S.E., Doebley J., Kresovich S., Goodman M.M, Buckler S. (2005) Maize association population: a high-resolution platform for quantitative trait locus dissection. The Plant Journal 44, 1054–1064.

Flood R.G., Halloran G.M. (1983) Developmental Responses in Tetraploid Wheat (*Triticum-Turgidum*). Canadian Journal of Botany-Revue Canadienne De Botanique 61:2539-2545.

Freeman K.W., Raun W.R., Johnson G.V., Mullen R.W., Stone M.L., Solie J.B. (2003) Late-season prediction of wheat grain yield and grain protein. Communications in Soil Science and Plant Analysis 34:1837-1852. DOI: Doi 10.1081/Css-120023219.

Fu D.L., Szucs P., Yan L.L., Helguera M., Skinner J.S., von Zitzewitz J., Hayes P.M., Dubcovsky J. (2005) Large deletions within the first intron in VRN-1 are associated with spring growth habit in barley and wheat. Molecular Genetics and Genomics 273:54-65. DOI: DOI 10.1007/s00438-004-1095-4.

Fusari, C., Lia V., Hopp E., Heinz R., Paniego N. (2008) Identification of single nucleotide polymorphisms and analysis of linkage disequilibrium in sunflower elite inbred lines using the candidate gene approach. BMC Plant Biology 8. DOI: Artn 7 Doi 0.1186/1471-2229-8-7.

Gaju O., Reynolds M.P., Sparkes D.L., Foulkes M.J. (2009) Relationships between large-spike phenotype, grain number, and yield potential in spring wheat. Crop Science 49:961-973. DOI: DOI 10.2135/cropsci2008.05.0285.

Giordani T., Buti M., Natali L., Pugliesi C., Cattonaro F., Morgante M., Cavallini A. (2011) An analysis of sequence variability in eight genes putatively involved in drought response in sunflower (*Helianthus annuus* L.). Theoretical Applied Genetics 122:1039-1049.

Glover K. D., Rudd J.C., Devkota R.N., Hall R.G., Jin Y., Rickertsen J.R. (2006) Registration of 'Granger' wheat. Crop Science 46:1390-1391.

Gonzalez-Martinez S.C., Ersoz E., Brown G.R., Wheeler N.C., Neale D.B. (2006) DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L. Genetics 172:1915-1926. DOI: DOI 10.1534/genetics.105.047126.

Gonzalez-Martinez S.C., Huber D., Ersoz E., Davis J.M., Neale D.B. (2008) Association genetics in *Pinus taeda* L. II. Carbon isotope discrimination. Heredity 101:19-26. DOI: Doi 10.1038/Hdy.2008.21.

Graybosch R.A., Peterson C.J. (2010) Genetic Improvement in Winter Wheat Yields in the Great Plains of North America, 1959-2008. Crop Science 50:1882-1890. DOI: DOI 10.2135/cropsci2009.11.0685.

Green A.J., Berger G., Griffey C.A., Pitman R., Thomason W., Balota M., Ahmed A. (2012) Genetic yield improvement in soft red winter wheat in the Eastern United States from 1919 to 2009. Crop Science 52:2097-2108. DOI: DOI 10.2135/cropsci2012.01.0026.

Griffiths S., Simmonds J., Leverington M., Wang Y., Fish L., Sayers L., Alibert L., Orford S., Wingen L., Herry L., Faure S., Laurie D., Bilham L., Snape J. (2009) Meta-QTL analysis of the genetic control of ear emergence in elite European winter wheat germplasm. Theoretical and Applied Genetics 119:383-395.

Griffiths S., Simmonds J., Leverington M., Wang Y.K., Fish L., Sayers L., Alibert L., Orford S., Wingen L., Snape J. (2012) Meta-QTL analysis of the genetic control of crop height in elite European winter wheat germplasm. Molecular Breeding 29:159-171. DOI: DOI 10.1007/s11032-010-9534-x.

Gupta P.K., Langridge P., Mir R.R. (2010) Marker-assisted wheat breeding: present status and future possibilities. Molecular Breeding 26:145-161. DOI: DOI 10.1007/s11032-009-9359-7.

Gupta P.K., Rustgi S., Kulwal P.L. (2005) Linkage disequilibrium and association studies in higher plants: Present status and future prospects. Plant Molecular Biology 57:461-485. DOI: DOI 10.1007/s11103-005-0257-z.

Gupta P.K., Varshney R.K., Sharma P.C., Ramesh B. (1999) Molecular markers and their applications in wheat breeding. Plant Breeding 118:369-390. DOI: DOI 10.1046/j.1439-0523.1999.00401.x.

Gutierrez, M., M.P. Reynolds, Raun W.R., Stone M.L., Klatt A.R. (2010). Spectral water indices for assessing yield in elite bread wheat genotypes grown under well irrigated, water defi cit stress, and high temperature conditions. Crop Sci. 50:197–214.

Habash D.Z., Kehel Z., Nachit M. (2009) Genomic approaches for designing durum wheat ready for climate change with a focus on drought. Journal of Experimental Botany 60:2805-2815. DOI: Doi 10.1093/Jxb/Erp211.

Hamblin M.T., Close T.J., Bhat P.R., Chao S.M., Kling J.G., Abraham K.J., Blake T., Brooks W.S., Cooper B., Griffey C.A., Hayes P.M., Hole D.J., Horsley R.D., Obert D.E., Smith K.P., Ullrich S.E., Muehlbauer G.J., Jannink J.L. (2010) Population structure and linkage disequilibrium in US barley germplasm: implications for association mapping. Crop Science 50:556-566. DOI: DOI 10.2135/cropsci2009.04.0198.

Hanocq E., Laperche A., Jaminon O., Laine A.L., Le Gouis J. (2007) Most significant genome regions involved in the control of earliness traits in bread wheat, as revealed by QTL meta-analysis. Theoretical and Applied Genetics 114:569-584. DOI: DOI 10.1007/s00122-006-0459-z.

Hansen K.A., Martin J.M., Lanning S.P., Talbert L.E. (2005) Correlation of genotype performance for agronomic and physiological traits in space-planted versus densely seeded conditions. Crop Science 45:1023-1028. DOI: DOI 10.2135/cropsci2004.0194.

Hao C.Y., Perretant M.R., Choulet F., Wang L.F., Paux E., Sourdille P., Zhang X.Y., Feuillet C., Balfourier F. (2010) Genetic diversity and linkage disequilibrium studies on a 3.1-Mb genomic region of chromosome 3B in European and Asian bread wheat (*Triticum aestivum* L.) populations. Theoretical and Applied Genetics 121:1209-1225. DOI: DOI 10.1007/s00122-010-1382-x.

Haudry A., Cenci A., Ravel C., Bataillon T., Brunel D., Poncet C., Hochu I., Poirier S., Santoni S., Gle´mink S., David J. (2007) Grinding up Wheat: A Massive Loss of Nucleotide Diversity since Domestication. Mol. Biol. Evol. 24: 1506-1517.

Haseneyer G., Stracke S., Piepho H.P., Sauer S., Geiger H.H., Graner A. (2010) DNA polymorphisms and haplotype patterns of transcription factors involved in barley endosperm development are associated with key agronomic traits. BMC Plant Biology 10. DOI: Artn 5 Doi 10.1186/1471-2229-10-5.

Hay R.K.M. (1995) Harvest Index - a Review of Its Use in Plant-Breeding and Crop Physiology. Annals of Applied Biology 126:197-216. DOI: DOI 10.1111/j.1744-7348.1995.tb05015.x.

Hazratkulova S., Sharma R.C., Alikulov S., Islomov S., Yuldashev T., Ziyaev Z., Khalikulov Z., Ziyadullaev Z., Turok J. (2012) Analysis of genotypic variation for normalized difference vegetation index and its relationship with grain yield in winter wheat under terminal heat stress. Plant Breeding 131:716-721. DOI: Doi 10.1111/Pbr.12003.

Hill W.G., Robertson A. (1968) Linkage disequilibrium in finite populations. Theor Appl Genet 38:226–231

Hincha D.K., Zuther E., Heyer A.G. (2003) The preservation of liposomes by raffinose family oligosaccharides during drying is mediated by effects on fusion and lipid phase transitions. Biochimica Et Biophysica Acta-Biomembranes 1612:172-177. DOI: Doi 10.1016/S0005-2736(03)00116-0.

Hirschorn J.N, Daly M.J. (2005) Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6:95–108.

Holland J.B. (2006) Estimating genotypic correlations and their standard errors using multivariate restricted maximum likelihood estimation with SAS Proc MIXED. Crop Science 46:642-654. DOI: DOI 10.2135/cropsci2005.0191.

Holland J.B. (2007) Genetic architecture of complex traits in plants. Current Opinion in Plant Biology 10:156-161. DOI: DOI 10.1016/j.pbi.2007.01.003.

Huang X.H., Wei X.H., Sang T., Zhao Q.A., Feng Q., Zhao Y., Li C.Y., Zhu C.R., Lu T.T., Zhang Z.W., Li M., Fan D.L., Guo Y.L., Wang A., Wang L., Deng L.W., Li W.J., Lu Y.Q., Weng Q.J., Liu K.Y., Huang T., Zhou T.Y., Jing Y.F., Li W., Lin Z., Buckler E.S., Qian Q.A., Zhang Q.F., Li J.Y., Han B. (2010) Genome-wide association studies of 14 agronomic traits in rice landraces. Nature Genetics 42:961-U76. DOI: Doi 10.1038/Ng.695.

Huang X.Q., Brule-Babel A. (2012) Sequence diversity, haplotype analysis, association mapping and functional marker development in the waxy and starch synthase IIa genes for grain-yield-related traits in hexaploid wheat (*Triticum aestivum* L.). Molecular Breeding 30:627-645. DOI: DOI 10.1007/s11032-011-9649-8.

Huang X.Q., Cloutier S., Lycar L., Radovanovic N., Humphreys D.G., Noll J.S., Somers D.J., Brown P.D. (2006) Molecular detection of QTL for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (*Triticum aestivum* L.). Theoretical and Applied Genetics 113:753-766. DOI: DOI 10.1007/s00122-006-0346-7.

Huang X.Q., Coster H., Ganal M.W., Roder M.S. (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 106:1379-1389. DOI: DOI 10.1007/s00122-002-1179-7.

Hussain S.S., Rivandi A. (2007) Molecular breeding for drought tolerance in plants: wheat prospective. Proc. Pakistan Acad. Sci. 44(1): 35-62.

Ingvarsson P.K., Street N.R. (2011) Association genetics of complex traits in plants. New Phytologist 189:909-22. DOI: 10.1111/j.1469-8137.2010.03593.x.

Inoue T., Inanaga S., Sugimoto Y., An P., Eneji A.E. (2004) Effect of drought on ear and flag leaf photosynthesis of two wheat cultivars differing in drought resistance. Photosynthetica 42:559-565. DOI: Doi 10.1007/S11099-005-0013-2.

Jaccoud D., Peng K., Feinstein D., Kilian A. (2001) Diversity arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Research 29:e25.

Ji X.M., Shiran B., Wan J.L., Lewis D.C., Jenkins C.L.D., Condon A.G., Richards R.A., Dolferus R. (2010) Importance of pre-anthesis anther sink strength for maintenance of grain number during reproductive stage water stress in wheat. Plant Cell and Environment 33:926-942. DOI: DOI 10.1111/j.1365-3040.2010.02130.x.

Joshi A.K., Kumari M., Singh V.P., Reddy C.M., Kumar S., Rane J., Chand R. (2007) Stay green trait: variation, inheritance and its association with spot blotch resistance in spring wheat (*Triticum aestivum* L.). Euphytica 153:59-71. DOI: DOI 10.1007/s10681-006-9235-z.

Jung M., Ching A., Bhattramakki D., Dolan M., Tingey S., Morgante M., Rafalski A. (2004) Linkage disequilibrium and sequence diversity in a 500-kbp region around the adh1 locus in elite maize germplasm. Theoretical and Applied Genetics 109:681-689. DOI: DOI 10.1007/s00122-004-1695-8.

Kadam S., Singh K., Shukla S., Goel S., Vikram P., Pawar V., Gaikwad K., Khanna-Chopra R., Singh N. (2012) Genomic associations for drought tolerance on the short arm of wheat chromosome 4B. Funct Integr Genomics 12:447-64. DOI: 10.1007/s10142-012-0276-1.

Kirigwi F.M., Van Ginkel M., Brown-Guedira G., Gill B.S., Paulsen G.M., Fritz A.K. (2007) Markers associated with a QTL for grain yield in wheat under drought. Molecular Breeding 20:401-413. DOI: DOI 10.1007/s11032-007-9100-3.

Kobayashi F., Ishibashi, M., Takumi, S. (2008) Transcriptional activation of Cor/Lea genes and increase in abiotic stress tolerance through expression of a wheat DREB2 homolog in transgenic tobacco. Transgenic Res. 17:755–767 doi:10.1007/s11248-007-9158-z. PMID:18034365.

Korzun V., Roder M., Ganal M., Worland A., Law C. (1998) Genetic analysis of the dwarfing gene (Rht8) in wheat.Part I. Molecular mapping of Rht8 on the short arm of chromosome 2D of bread wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 96:1104-1109.

Kuchel H., Hollamby G., Langridge P., Williams K., Jefferies S.P. (2006) Identification of genetic loci associated with ear-emergence in bread wheat. Theoretical and Applied Genetics 113:1103-1112. DOI: DOI 10.1007/s00122-006-0370-7.

Kuchel H., Williams K.J., Langridge P., Eagles H.A., Jefferies S.P. (2007) Genetic dissection of grain yield in bread wheat. I. QTL analysis. Theoretical and Applied Genetics 115:1029-1041. DOI: DOI 10.1007/s00122-007-0629-7.

Kulwal P., Ishikawa G., Benscher D., Feng Z.Y., Yu L.X., Jadhav A., Mehetre S., Sorrells M.E. (2012) Association mapping for pre-harvest sprouting resistance in white winter wheat. Theoretical and Applied Genetics 125:793-805. DOI: DOI 10.1007/s00122-012-1872-0.

Kumar N., Kulwal P.L., Balyan H.S., Gupta P.K. (2007) QTL mapping for yield and yield contributing traits in two mapping populations of bread wheat. Molecular Breeding 19:163-177. DOI: DOI 10.1007/s11032-006-9056-8.

Kumar U., Joshi A.K., Kumari M., Paliwal R., Kumar S., Roder M.S. (2010) Identification of QTL for stay green trait in wheat (*Triticum aestivum* L.) in the 'Chirya 3' x 'Sonalika' population. Euphytica 174:437-445. DOI: DOI 10.1007/s10681-010-0155-6.

Kumari M., Pudake R.N., Singh V.P., Joshi A. K. (2012) Association of staygreen trait with canopy temperature depression and yield traits under terminal heat stress in wheat (*Triticum aestivum L.*) Euphytica DOI:10.1007/s10681-012-0780-3.

Kumari M., Singh V.P., Tripathi R., Joshi A.K. (2007) Variation for stay green trait and its association with canopy temperature depression and yield traits under terminal heat stress in wheat. Wheat Production in Stressed Environments 12:357-363.

Langridge P., Fleury D. (2011) Making the most of 'omics' for crop breeding. Trends in Biotechnology 29:33-40. DOI: DOI 10.1016/j.tibtech.2010.09.006.

Latini A., Rasi C., Sperandei M., Cantale C., Lannetta M., Dettori M., Ammar K., Galeffi P. (2007) Identification of a DREB-related gene in *Triticum durum* and its expression under stress conditions. Ann Appl Biol 150: 187-195.

Levitt J. (1972) Responses of plants to environmental stresses. New York: Academic Press.

Lewontin RC (1988) On measures of gametic disequilibrium. Genetics 120:849–852.

Landjeva S., Korzun V., Borner A. (2007) Molecular markers: actual and potential contributions to wheat genome characterization and breeding. Euphytica 156:271-296. DOI: DOI 10.1007/s10681-007-9371-0.

Li S.S., Jia J.Z., Wei X.Y., Zhang X.C., Li L.Z., Chen H.M., Fan Y.D., Sun H.Y., Zhao X.H., Lei T.D., Xu Y.F., Jiang F.S., Wang H.G., Li L.H. (2007) A intervarietal genetic map and QTL analysis for yield traits in wheat. Molecular Breeding 20:167-178. DOI: DOI 10.1007/s11032-007-9080-3.

Lin F., Xue S.L., Tian D.G., Li C.J., Cao Y., Zhang Z.Z., Zhang C.Q., Ma Z.Q. (2008) Mapping chromosomal regions affecting flowering time in a spring wheat RIL population. Euphytica 164:769-777. DOI: DOI 10.1007/s10681-008-9724-3.

Lipka A.E., Tian F., Wang Q.S., Peiffer J., Li M., Bradbury P.J., Gore M.A., Buckler E.S., Zhang Z.W. (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics 28:2397-2399. DOI: DOI 10.1093/bioinformatics/bts444.

Liu L., Wang L., Yao J. , Zheng Y., Zhao C. (2010) Association mapping of six agronomic traits on chromosome 4A of wheat (*Triticum aestivum* L.). Molecular Plant Breeding. DOI: 10.5376/mpb.2010.01.0005.

Lopes M.S., Reynolds M.P. (2012) Stay-green in spring wheat can be determined by spectral reflectance measurements (normalized difference vegetation index) independently from phenology. Journal of Experimental Botany 63:3789-3798. DOI: Doi 10.1093/Jxb/Ers071.

Lopes M.S., Reynolds M.P., Jalal-Kamali M.R., Moussa M., Feltaous Y., Tahir I.S.A., Barma N., Vargas M., Mannes Y., Baum M. (2012) The yield correlations of selectable physiological traits in a population of advanced spring wheat lines grown in warm and drought environments. Field Crops Research 128:129-136. DOI: DOI 10.1016/j.fcr.2011.12.017.

Lothier J., Lasseur B., Le Roy K., Van Laere A., Prud'homme M.P., Barre P., Van den Ende W., Morvan-Bertrand A. (2007) Cloning, gene mapping, and functional analysis of a fructan 1-exohydrolase (*1-FEH*) from *Lolium perenne* implicated in fructan synthesis rather than in fructan mobilization. Journal of Experimental Botany 58:1969-1983. DOI: Doi 10.1093/Jxb/Erm053.

Lu Y.L., Xu J., Yuan Z.M., Hao Z.F., Xie C.X., Li X.H., Shah T., Lan H., Zhang S.H., Rong T.Z., Xu Y.B. (2012) Comparative LD mapping using single SNPs and haplotypes identifies QTL for plant height and biomass as secondary traits of drought tolerance in maize. Molecular Breeding 30:407-418. DOI: DOI 10.1007/s11032-011-9631-5.

Lupton F.G.H. (1966) Translocation of Photosynthetic Assimilates in Wheat. Annals of Applied Biology 57:355-&. DOI: DOI 10.1111/j.1744-7348.1966.tb03829.x.

Ma Z.Q., Zhao D.M., Zhang C.Q., Zhang Z.Z., Xue S.L., Lin F., Kong Z.X., Tian D.G., Luo Q.Y. (2007) Molecular genetic analysis of five spike-related traits in wheat using RIL

and immortalized F-2 populations. Molecular Genetics and Genomics 277:31-42. DOI: DOI 10.1007/s00438-006-0166-0.

Maccaferri M., Sanguineti M.C., Corneti S., Ortega J.L.A., Ben Salem M., Bort J., DeAmbrogio E., del Moral L.F.G., Demontis A., El-Ahmed A., Maalouf F., Machlab H., Martos V., Moragues M., Motawaj J., Nachit M., Nserallah N., Ouabbou H., Royo C., Slama A., Tuberosa R. (2008) Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. Genetics 178:489-511. DOI: DOI 10.1534/genetics.107.077297.

Maccaferri M., Sanguineti M.C., Demontis A., El-Ahmed A., del Moral L.G., Maalouf F., Nachit M., Nserallah N., Ouabbou H., Rhouma S., Royo C., Villegas D., Tuberosa R. (2011) Association mapping in durum wheat grown across a broad range of water regimes. Journal of Experimental Botany 62:409-438. DOI: Doi 10.1093/Jxb/Erq287.

Maccaferri M., Sanguineti M.C., Mantovani P., Demontis A., Massi A., Ammar K., Kolmer J.A., Czembor J.H., Ezrati S., Tuberosa R. (2010) Association mapping of leaf rust response in durum wheat. Molecular Breeding 26:189-228. DOI: DOI 10.1007/s11032-009-9353-0.

Maccaferri M., Sanguineti M.C., Noli E., Tuberosa R. (2005) Population structure and long-range linkage disequilibrium in a durum wheat elite collection. Molecular Breeding 15:271-289. DOI: DOI 10.1007/s11032-004-7012-z.

Manmathan H., Shaner D., Snelling J., Tisserat N., Lapitan N. (2013) Virus-induced gene silencing of *Arabidopsis thaliana* gene homologues in wheat identifies genes conferring improved drought tolerance. Journal of Experimental Botany. doi:10.1093/jxb/ert003.

Marone D., Panio G., Ficco D.B.M., Russo M.A., De Vita P., Papa R., Rubiales D., Cattivelli L., Mastrangelo A.M. (2012) Characterization of wheat DArT markers: genetic and functional features. Molecular Genetics and Genomics 287:741-753. DOI: DOI 10.1007/s00438-012-0714-8.

Marti J., Bort J., Slafer G.A., Araus J.L. (2007) Can wheat yield be assessed by early measurements of Normalized Difference Vegetation Index? Annals of Applied Biology 150:253-257. DOI: DOI 10.1111/j.1744-7348.2007.00126.x.

Martinez-Gonzalez S.C., Huber D., Ersoz E., Davis J. M., Neale D. B. (2008) Association genetics in *Pinus taeda* L.II. Carbon isotope discrimination. Heredity 101:19-26.

Marza F., Bai G.H., Carver B.F., Zhou W.C. (2006) Quantitative trait loci for yield and related traits in the wheat population Ning7840 x Clark. Theoretical and Applied Genetics 112:688-698. DOI: DOI 10.1007/s00122-005-0172-3.

Mather K.A., Caicedo A.L., Polato N.R., Olsen K.M., McCouch S., Purugganan M.D. (2007) The extent of linkage disequilibrium in rice (*Oryza sativa* L.). Genetics 177:2223-2232. DOI: DOI 10.1534/genetics.107.079616.

McCartney C.A., Somers D.J., Humphreys D.G., Lukow O., Ames N., Noll J., Cloutier S., McCallum B.D. (2005) Mapping quantitative trait loci controlling agronomic traits in the spring wheat cross RL4452 x 'AC Domain'. Genome 48:870-883.

McCouch S.R., Zhao K.Y., Wright M., Tung C.W., Ebana K., Thomson M., Reynolds A., Wang D., DeClerck G., Ali M.L., McClung A., Eizenga G., Bustamante C. (2010) Development

of genome-wide SNP assays for rice. Breeding Science 60:524-535. DOI: Doi 10.1270/Jsbbs.60.524.

McIntyre C.L., Mathews K.L., Rattey A., Chapman S.C., Drenth J., Ghaderi M., Reynolds M., Shorter R. (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. Theoretical and Applied Genetics 120:527-541. DOI: DOI 10.1007/s00122-009-1173-4.

Mergoum M., Frohberg R.C., Ali S., Singh P.K., Rasmussen J.B., Miller J.D. (2006) Registration of Spring Wheat Germplasm ND 735 Combining Tan Spot, Leaf, and Stem Rusts. Crop Science 46:1003-1004.

Mir R., Zaman-Allah M., Sreenivasulu N., Trethowan R.,Varshney R. (2012) Integrated genomics, physiology and breeding approaches for improving drought tolerance in crops. Theoretical and Applied Genetics. DOI 10.1007/s00122-012-1904-9.

Mochida K., Yamazaki Y., Ogihara Y. (2003) Discrimination of homoeologous gene expression in hexaploid wheat by SNP analysis of contigs grouped from a large number of expressed sequence tags. Mol Genet Genomics. 270:371-377.

Mulat T. (2004) Chemical desiccation and tolerance and nonstructural carbohydrate dynamics in winter wheat. Ph.D. dissertation, Colorado State University, U.S.A.

Munns R., Richards R.A. (2010) Recent advances in breeding wheat for drought and salt tolerance. *M. A. Jenks et al. (eds.),* Advances in Molecular Breeding Toward Drought and Salt Tolerance in Crops, 565-585.

Naruoka Y., Talbert L.E., Lanning S.P., Blake N.K., Martin J.M., Sherman J.D. (2011) Identification of quantitative trait loci for productive tiller number and its relationship to agronomic traits in spring wheat. Theoretical and Applied Genetics 123:1043-1053. DOI: DOI 10.1007/s00122-011-1646-0.

Nei M. (1987) Molecular evolutionary genetics. Columbia University Press, New York.

Neumann K., Kobiljski B., Dencic S., Varshney R.K., Borner A. (2011) Genome-wide association mapping: a case study in bread wheat (Triticum aestivum L.). Molecular Breeding 27:37-58. DOI: DOI 10.1007/s11032-010-9411-7.

Nevo E., Chen G.X. (2010) Drought and salt tolerances in wild relatives for wheat and barley improvement. Plant Cell and Environment 33:670-685. DOI: DOI 10.1111/j.1365-3040.2009.02107.x.

Nezhad K.Z., Weber W.E., Roder M.S., Sharma S., Lohwasser U., Meyer R.C., Saal B., Borner A. (2012) QTL analysis for thousand-grain weight under terminal drought stress in bread wheat (*Triticum aestivum* L.). Euphytica 186:127-138. DOI: DOI 10.1007/s10681-011-0559-y.

Osborne B.G., Anderssen R.S. (2003) Single-kernel characterization principles and applications. Cereal Chemistry 80:613-622. DOI: Doi 10.1094/Cchem.2003.80.5.613.

Passioura J. (2007) The drought environment: physical, biological and agricultural perspectives. Journal of Experimental Botany 58:113-117. DOI: Doi 10.1093/Jxb/Erl212.

Passioura J.B. (1977) Grain-Yield, Harvest Index, and Water-Use of Wheat. Journal of the Australian Institute of Agricultural Science 43:117-120.

Passioura J.B. (2012) Phenotyping for drought tolerance in grain crops: when is it useful to breeders? Functional Plant Biology 39:851-859. DOI: Doi 10.1071/Fp12079.

Pei Z.M., Ghassemian M., Kwak C. M., McCourt P., Schroeder J. I. (1998) Role of famesyîtransferase in ABA regulation of guard cell anion channels and plant water loss. Science 282:287-290.

Peleg Z., Fahima T., Krugman T., Abbo S., Yakir D., Korol A.B., Saranga Y. (2009) Genomic dissection of drought resistance in durum wheat x wild emmer wheat recombinant inbreed line population. Plant Cell and Environment 32:758-779. DOI: DOI 10.1111/j.1365-3040.2009.01956.x.

Pellegrineschi A., Reynolds M., Pacheco M., Brito R.M., Almeraya R., Yamaguchi-Shinozaki K., Hoisington D. (2004) Stress-induced expression in wheat of the Arabidopsis thaliana DREB1A gene delays water stress symptoms under greenhouse conditions. Genome 47:493-500. DOI: Doi 10.1139/G03-140.

Peng J.H., Bai Y., Haley S.D., Lapitan N.L.V. (2009) Microsatellite-based molecular diversity of bread wheat germplasm and association mapping of wheat resistance to the Russian wheat aphid. Genetica 135:95-122. DOI: DOI 10.1007/s10709-008-9262-x.

Pfeiffer W., Trethowan R., Van G., Ortiz M., Rajaram S. (2005) Breeding for abotic stress tolerance in wheat. In: Ashraf M., Harris P. (eds) Abiotic Stress: Plant Resistance Through Breeding and molecular Approaches. The Haworth Press, Inc. NY Pp 401-489.

Pinto R.S., Reynolds M.P., Mathews K.L., McIntyre C.L., Olivares-Villegas J.J., Chapman S.C. (2010) Heat and drought adaptive QTL in a wheat population designed to minimize

confounding agronomic effects. Theoretical Applied Genetics 121:1001-21. DOI: 10.1007/s00122-010-1351-4.

Poland J.A., Rife T.W. (2012) Genotyping-by-Sequencing for Plant Breeding and Genetics. Plant Genome 5:92-102. DOI: DOI 10.3835/plantgenome2012.05.0005.

Prasad A.K., Singh R.P., Tare V., Kafatos M. (2007) Use of vegetation index and meteorological parameters for the prediction of crop yield in India. International Journal of Remote Sensing 28:5207-5235.

Prasad B., Carver B.F., Stone M.L., Babar M.A., Raun W.R., Klatt A.R. (2007a) Genetic analysis of indirect selection for winter wheat grain yield using spectral reflectance indices. Crop Science 47:1416-1425. DOI: DOI 10.2135/cropsci2006.08.0546.

Prasad B., Carver B.F., Stone M.L., Babar M.A., Raun W.R., Klatt A.R. (2007b) Potential use of spectral reflectance indices as a selection tool for grain yield in winter wheat under Great Plains conditions. Crop Science 47:1426-1440. DOI: DOI 10.2135/cropsci2006.07.0492.

Pritchard J.K., Stephens M., Donnelly P. (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Quarrie S.A., Quarrie S.P., Radosevic R., Rancic D., Kaminska A., Barnes J.D., Leverington M., Ceoloni C., Dodig D. (2006) Dissecting a wheat QTL for yield present in a range of environments: from the QTL to candidate genes. Journal of Experimental Botany 57:2627-2637. DOI: Doi 10.1093/Jxb/Erl026.

Quarrie S.A., Steed A., Calestani C., Semikhodskii A., Lebreton C., Chinoy C., Steele N., Pljevljakusic D., Waterman E., Weyen J., Schondelmaier J., Habash D.Z., Farmer P.,

Saker L., Clarkson D.T., Abugalieva A., Yessimbekova M., Turuspekov Y., Abugalieva S., Tuberosa R., Sanguineti M.C., Hollington P.A., Aragues R., Royo A., Dodig D. (2005) A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring X SQ1 and its use to compare QTL for grain yield across a range of environments. Theoretical and Applied Genetics 110:865-880. DOI: DOI 10.1007/s00122-004-1902-7.

Rafalski A. (2002) Applications of single nucleotide polymorphisms in crop genetics. Current Opinion in Plant Biology 5:94-100. DOI: Doi 10.1016/S1369-5266(02)00240-6.

Rampino P., Pataleo S., Gerardi C., Mita G., Perrotta C. (2006) Drought stress response in wheat: physiological and molecular analysis of resistant and sensitive genotypes. Plant Cell and Environment 29:2143-2152. DOI: DOI 10.1111/j.1365-3040.2006.01588.x.

Ravel C., Praud S., Canaguier A., Dufour P., Giancola S., Balfourier F., Chalhoub B., Brunel D., Linossier L., Dardevet M., Beckert M., Rousset M., Murigneux A., Charmet G. (2007) DNA sequence polymorphisms and their application to bread wheat quality. Euphytica 158:331-336.

Ravi K., Vadez V., Isobe S., Mir R.R, Guo Y., Nigam S.N, Gowda M.V.C., Radhakrishnan T., Bertioli D.J., Knapp S.J., Varshney R.K. (2011) Identification of several small main-effect QTL and a largenumber of epistatic QTL for drought tolerance in groundnut (*Arachis hypogaea* L.). Theor Appl Genet 122:1119–1132.

Rebetzke G.J., Ellis M.H., Bonnett D.G., Mickelson B., Condon A.G., Richards R.A. (2012) Height reduction and agronomic performance for selected gibberellin-responsive

dwarfing genes in bread wheat (*Triticum aestivum* L.). Field Crops Research 126:87-96. DOI: DOI 10.1016/j.fcr.2011.09.022.

Reif J.C., Zhang P., Dreisigacker S., Warburton M.L., van Ginkel M., Hoisington D., Bohn M., Melchinger A.E. (2005) Wheat genetic diversity trends during domestication and breeding. Theoretical and Applied Genetics 110:859-864. DOI: DOI 10.1007/s00122-004-1881-8.

Remington D.L., Thornsberry J.M., Matsuoka Y., Wilson L.M., Whitt S.R., Doeblay J., Kresovich S., Goodman M.M., Buckler E.S. (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proceedings of the National Academy of Sciences of the United States of America 98:11479-11484. DOI: DOI 10.1073/pnas.201394398.

Reynolds M., Bonnett D., Chapman S.C., Furbank R.T., Manes Y., Mather D.E., Parry M.A.J. (2011) Raising yield potential of wheat. I. Overview of a consortium approach and breeding strategies. Journal of Experimental Botany 62:439-452. DOI: Doi 10.1093/Jxb/Erq311.

Reynolds A., Lopez-Tejeira F., Cassagne D., Garcia-Vidal F.J., Jouanin C., Sanchez-Dehesa J. (1999) Spectral properties of opal-based photonic crystals having a SiO2 matrix. Physical Review B 60:11422-11426. DOI: DOI 10.1103/PhysRevB.60.11422.

Reynolds, M.P., Manes, Y., Izanloo, A. and Langridge, P. (2009) Phenotyping for physiological breeding and gene discovery in wheat. Annals of Applied Biology: 155: 309–320.

Reynolds M.P., Mujeeb-Kazi A., Sawkins M. (2005) Prospects for utilising plant-adaptive mechanisms to improve wheat and other crops in drought- and salinity-prone environments. Annals of Applied Biology 146:239-259. DOI: DOI 10.1111/j.1744-7348.2005.040058.x.

Roussel V., Leisova L., Exbrayat F., Stehno Z., Balfourier F. (2005) SSR allelic diversity changes in 480 European bread wheat varieties released from 1840 to 2000. Theoretical and Applied Genetics 111:162-170. DOI: DOI 10.1007/s00122-005-2014-8.

Rousset M., Bonnin I., Remoue C., Falque M., Rhone B., Veyrieras J.B., Madur D., Murigneux A., Balfourier F., Le Gouis J., Santoni S., Goldringer I. (2011) Deciphering the genetics of flowering time by an association study on candidate genes in bread wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 123:907-926. DOI: DOI 10.1007/s00122-011-1636-2.

Rozas J, Librado P., Sánchez-DelBarrio J. C., Messeguer X., Rozas R. (1999) DnaSP version 5: an integrated program for population genetics and molecular evolution analysis. Bioinformatics 15:174-175.

Saini H.S., Aspinall, D. (1982) Abnormal sporogenesis in wheat (*Triticum aestivum* L.) induced by short periods of high temperature. Ann. Bot. 49:835–846.

Saint Pierre, C., Crossa J., Bonnet D., Yamaguchi-Shinozaki K., Reynolds M. (2012) Phenotyping transgenic wheat for drought resistance. Journal of Experimental Botany 63 (5):1799-1808.

Salekdeh G.H., Reynolds M., Bennett J., Boyer J. (2009) Conceptual framework for drought phenotyping during molecular breeding. Trends in Plant Science 14:488-496. DOI: DOI 10.1016/j.tplants.2009.07.007.

SAS Institute (2006) SAS for Mixed Models. SAS Inst., Cary, NC.

SAS Institute (2011) The SAS system for Windows. Version 9.3. SAS Inst., Cary, NC.

Scarth R., Law C.N. (1983) The Location of the Photoperiod Gene, Ppd2 and an Additional Genetic-Factor for Ear-Emergence Time on Chromosome-2b of Wheat. Heredity 51:607-619. DOI: Doi 10.1038/Hdy.1983.73.

Schneider K., Anderson L. (2010) Yield Gap and Productivity Potential in Ethiopian Agriculture: Staple Grains & Pulses. University of Washington, EPAR Brief No.98.

Setter T.L., Yan J.B., Warburton M., Ribaut J.M., Xu Y.B., Sawkins M., Buckler E.S., Zhang Z.W., Gore M.A. (2011) Genetic association mapping identifies single nucleotide polymorphisms in genes that affect abscisic acid levels in maize floral tissues during drought. Journal of Experimental Botany 62:701-716. DOI: Doi 10.1093/Jxb/Erq308.

Shah M.M., Gill K.S., Baenziger P.S., Yen Y., Kaeppler S.M., Ariyarathne H.M. (1999) Molecular mapping of loci for agronomic traits on chromosome 3A of bread wheat. Crop Science 39:1728-1732.

Shearman V.J., Sylvester-Bradley R., Scott R.K., Foulkes M.J. (2005) Physiological processes associated with wheat yield progress in the UK. Crop Science 45:175-185.

Shinozaki K., Yamaguchi-Shinozaki K. (2007) Gene networks involved in drought stress response and tolerance. Journal of Experimental Botany 58:221-227. DOI: Doi 10.1093/Jxb/Erl164.

Shinozaki K., Yamaguchi-Shinozaki Y. (2007) Gene networks involved in drought stress response and tolerance. Journal of Experimental Botany 58(2):221-227.

Silva S.A., Carvallo F.I., Oliveira A.C., Coimbra J.L., Vasconsellos N.J., Lorencetti C. (2000) Genetic basis of stay green traits. Journal of New seeds 2:55-68.

Sinclair T.R. (2012) Is transpiration efficiency a viable plant trait in breeding for crop improvement? Functional Plant Biology 39:359-365. DOI: Doi 10.1071/Fp11198.

Sinclair T.R., Jamieson P.D. (2006) Grain number, wheat yield, and bottling beer: An analysis. Field Crops Research 98:60-67. DOI: DOI 10.1016/j.fcr.2005.12.006.

Sneller C.H., Mather D.E., Crepieux S. (2009) Analytical Approaches and Population Types for Finding and Utilizing QTL in Complex Plant Populations. Crop Science 49:363-380. DOI: DOI 10.2135/cropsci2008.07.0420.

Sorrells M., Yu J. (2009) Linkage Disequilibrium and Association Mapping in the Triticeae. C. Feuillet, G. J. Muehlbauer (eds.), Genetics and Genomics of the Triticeae, Plant Genetics and Genomics: Crops and Models 7. DOI 10. 1007/978-387-77489-3_22.

Spano G., Di Fonzo N., Perrotta C., Platani C., Ronga G., Lawlor D.W., Napier J.A., Shewry P.R. (2003) Physiological characterization of 'stay green' mutants in durum wheat. Journal of Experimental Botany 54:1415-1420. DOI: Doi 10.1093/Jxb/Erg150.

Spielmeyer W., McIntosh R.A., Kolmer J., Lagudah E.S. (2005) Powdery mildew resistance and Lr34/Yr18 genes for durable resistance to leaf and stripe rust co-segregate at a locus on the short arm of chromosome 7D of wheat. Theoretical and Applied Genetics 111:731-735. DOI: DOI 10.1007/s00122-005-2058-9.

Srivalli S., Khanna-Chopra R. (2009) Delayed wheat flag leaf senescence due to removal of spikelets is associated with increased activities of leaf antioxidant enzymes, reduced glutathione/oxidized glutathione ratio and oxidative damage to mitochondrial proteins. Plant Physiology and Biochemistry 47:663-670. DOI: DOI 10.1016/j.plaphy.2009.03.015.

Su Z.Q., Hao C.Y., Wang L.F., Dong Y.C., Zhang X.Y. (2011) Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (*Triticum aestivum* L.). Theoretical and Applied Genetics 122:211-223. DOI: DOI 10.1007/s00122-010-1437-z.

Sun X.Y., Wu K., Zhao Y., Kong F.M., Han G.Z., Jiang H.M., Huang X.J., Li R.J., Wang H.G., Li S.S. (2009) QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. Euphytica 165:615-624. DOI: DOI 10.1007/s10681-008-9794-2.

Tabor H.K., Risch N.J., Myers R.M. (2002) Candidate-gene approaches for studying complex genetic traits: practical considerations. Nature Reviews Genetics 3:391-A396. DOI: Doi 10.1038/Nrg796.

Tajima F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.

Thomas H., Howarth C.J. (2000) Five ways to stay green. Journal of Experimental Botany 51:329-337. DOI: DOI 10.1093/jexbot/51.suppl_1.329.

Tian D.C., Araki H., Stahl E., Bergelson J., Kreitman M. (2002) Signature of balancing selection in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 99:11525-11530. DOI: DOI 10.1073/pnas.172203599.

Trethowan R.M., Reynolds M. (2007) Drought resistance: Genetic approaches for improving productivity under stress. Wheat Production in Stressed Environments 12:289-299.

Ugarte C., Calderini D.F., Slafer G.A. (2007) Grain weight and grain number responsiveness to pre-anthesis temperature in wheat, barley and triticale. Field Crops Research 100:240-248. DOI: DOI 10.1016/j.fcr.2006.07.010.

Vadez V., Rao S., Sharma K.K., Bhatnagar-Mathur P., Devi M. J. (2007) *DREB1A* allows for more water uptake in groundnut by a large modification in the root/shoot ratio under water deficit. International Arachis Newsletter 27:27-31.

Vadez V., Rao J. S., Bhatnagar-Mathur P., Sharma K. K. (2013) DREB1A promotes root development in deep soil layers and increases water extraction under water stress in groundnut. Plant Biology 15:45-52.

Verma V., Foulkes M.J., Worland A.J., Sylvester-Bradley R., Caligari P.D.S., Snape J.W. (2004) Mapping quantitative trait loci for flag leaf senescence as a yield determinant in winter wheat under optimal and drought-stressed environments. Euphytica 135:255-263. DOI: Doi 10.1023/B:Euph.0000013255.31618.14.

Vijayalakshmi K., Fritz A.K., Paulsen G.M., Bai G.H., Pandravada S., Gill B.S. (2010) Modeling and mapping QTL for senescence-related traits in winter wheat under high temperature. Molecular Breeding 26:163-175. DOI: DOI 10.1007/s11032-009-9366-8.

Wang L.F., Ge H.M., Hao C.Y., Dong Y.S., Zhang X.Y. (2012) Identifying Loci Influencing 1,000-Kernel Weight in Wheat by Microsatellite Screening for Evidence of Selection during Breeding. Plos One 7. DOI: ARTN e29432 DOI 10.1371/journal.pone.0029432.

Wang R.X., Hai L., Zhang X.Y., You G.X., Yan C.S., Xiao S.H. (2009) QTL mapping for grain filling rate and yield-related traits in RILs of the Chinese winter wheat population Heshangmai X Yu8679. Theoretical and Applied Genetics 118:313-325. DOI: DOI 10.1007/s00122-008-0901-5.

Wanjugi, H., Coleman-Derr D., Huo N., Anderson O., Gu Y.Q. (2009) Rapid development of PCR-based genome-specific repetitive DNA junction markers in wheat. Genome 52: 576–587.

Warburton M.L., Crossa J., Franco J., Kazi M., Trethowan R., Rajaram S., Pfeiffer W., Zhang P., Dreisigacker S., van Ginkel M. (2006) Bringing wild relatives back into the family: recovering genetic diversity in CIMMYT improved wheat germplasm. Euphytica 149:289-301. DOI: DOI 10.1007/s10681-005-9077-0.

Waugh R., Jannink J.L., Muehlbauer G.J., Ramsay L. (2009) The emergence of whole genome association scans in barley. Current Opinion in Plant Biology 12:218-222. DOI: DOI 10.1016/j.pbi.2008.12.007.

Wei B., Jing R., Wang C., Chen J., Mao X., Chang X., Jia.J. (2009) Dreb1 genes in wheat (*Triticum aestivum* L.): development of functional markers and gene mapping based on SNPs. Mol Breeding 23:13-22.

Wilson L.M, Whitt S.R, Ibanez A.M, Rocheford T.R, Goodman M.M, Buckler E.S. (2004) Dissection of maize kernel composition and starch production by candidate gene association. Plant Cell 16:2719–2733.

Witcombe J.R., Hollington P.A., Howarth C.J., Reader S., Steele K.A. (2008) Breeding for abiotic stresses for sustainable agriculture. Philosophical Transactions of the Royal Society B-Biological Sciences 363:703-716. DOI: DOI 10.1098/rstb.2007.2179.

Worland A.J., Appendino M.L., Sayers E.J. (1994) The Distribution, in European Winter Wheats, of Genes That Influence Ecoclimatic Adaptability While Determining Photoperiodic Insensitivity and Plant Height. Euphytica 80:219-228. DOI: Doi 10.1007/Bf00039653.

Worland A.J., Borner A., Korzun V., Li W.M., Petrovic S., Sayers E.J. (1998) The influence of photoperiod genes on the adaptability of European winter wheats (Reprinted from Wheat: Prospects for global improvement, 1998). Euphytica 100:385-394. DOI: Doi 10.1023/A:1018327700985.

Worland, A.J. (1996) The influence of flowering time genes on environmental adaptability in European wheats. Euphytica 89:49-57.

Xu P., Wu X., Wang B., Luo J., Liu Y., Ehlers J.D., Close T.J., Roberts P.A., Lu Z., Wang S., Li G. (2012) Genome wide linkage disequilibrium in Chinese asparagus bean (*Vigna.*

unguiculata ssp. sesquipedialis) germplasm: implications for domestication history and genome wide association studies. Heredity 109:34-40. DOI: Doi 10.1038/Hdy.2012.8.

Xu W.W., Subudhi P.K., Crasta O.R., Rosenow D.T., Mullet J.E., Nguyen H.T. (2000) Molecular mapping of QTL conferring stay-green in grain sorghum (*Sorghum bicolor* L. Moench). Genome 43:461-469. DOI: Doi 10.1139/Gen-43-3-461.

Yan L., Fu D., Li C., Blechl A., Tranquilli G., Bonafede M., Sanchez A., Valarik M., Yasuda S., Dubcovsky J. (2006) The wheat and barley vernalization gene VRN3 is an orthologue of FT. Proceedings of the National Academy of Sciences of the United States of America 103:19581-19586. DOI: DOI 10.1073/pnas.0607142103.

Yan L., Helguera M., Kato K., Fukuyama S., Sherman J., Dubcovsky J. (2004) Allelic variation at the *VRN-1* promoter region in polyploid wheat. Theoretical and Applied Genetics 109:1677-1686. DOI: DOI 10.1007/s00122-004-1796-4.

Yan L., Loukoianov A., Tranquilli G., Helguera M., Fahima T., Dubcovsky J. (2003) Positional cloning of the wheat vernalization gene VRN1. Proceedings of the National Academy of Sciences of the United States of America 100:6263-6268. DOI: DOI 10.1073/pnas.0937399100.

Yan L.L., Loukoianov A., Blechl A., Tranquilli G., Ramakrishna W., SanMiguel P., Bennetzen J.L., Echenique V., Dubcovsky J. (2004) The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. Science 303:1640-1644. DOI: DOI 10.1126/science.1094305.

Yang, D., Jing R., Chang X., Li W. (2007) Identification of quantitative trait loci and environmental interactions for accumulation and remobilization of water-soluble carbohydrates in wheat (*Triticum aestivum* L.) stems. Genetics 176:571-584.

Yao J., Wang L.X., Liu L.H., Zhao C.P., Zheng Y.L. (2009) Association mapping of agronomic traits on chromosome 2A of wheat. Genetica 137:67-75. DOI: DOI 10.1007/s10709-009-9351-5.

Yoshida T., Nishida H., Zhu J., Nitcher R., Distelfeld A., Akashi Y., Kato K., Dubcovsky J. (2010) *Vrn-D4* is a vernalization gene located on the centromeric region of chromosome 5D in hexaploid wheat. Theoretical and Applied Genetics 120:543-552. DOI: DOI 10.1007/s00122-009-1174-3.

Yu J.M., Pressoir G., Briggs W.H., Bi I.V., Yamasaki M., Doebley J.F., McMullen M.D., Gaut B.S., Nielsen D.M., Holland J.B., Kresovich S., Buckler E.S. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nature Genetics 38:203-208. DOI: Doi 10.1038/Ng1702.

Yu L.X., Liu S.X., Anderson J.A., Singh R.P., Jin Y., Dubcovsky J., Brown-Guidera G., Bhavani S., Morgounov A., He Z.H., Huerta-Espino J., Sorrells M.E. (2012) Haplotype diversity of stem rust resistance loci in uncharacterized wheat lines (vol 26, pg 667, 2010). Molecular Breeding 30:613-614. DOI: DOI 10.1007/s11032-011-9680-9.

Yu L.X., Lorenz A., Rutkoski J., Singh R.P., Bhavani S., Huerta-Espino J., Sorrells M.E. (2011) Association mapping and gene-gene interaction for stem rust resistance in CIMMYT spring wheat germplasm. Theoretical and Applied Genetics 123:1257-1268. DOI: DOI 10.1007/s00122-011-1664-y.

Zhang H.P., Turner N.C., Poole M.L. (2012) Increasing the harvest index of wheat in the high rainfall zones of southern Australia. Field Crops Research 129:111-123. DOI: DOI 10.1016/j.fcr.2012.02.002.

Zhang J., Huang S., Fosu-Nyarko J., Dell B., McNeil M., Waters I., Moolhuijzen P., Conocono E., Appels R. (2008) The genome structure of the *1-FEH* genes in wheat (*Triticum aestivum* L.): new markers to track stem carbohydrates and grain filling QTL in breeding. Molecular Breeding 22:339-351. DOI: DOI 10.1007/s11032-008-9179-1.

Zhang L.Y., Marchand S., Tinker N.A., Belzile F. (2009) Population structure and linkage disequilibrium in barley assessed by DArT markers. Theoretical and Applied Genetics 119:43-52. DOI: DOI 10.1007/s00122-009-1015-4.

Zhang, J., Huang S., Fosu-Nyarko J., Dell B., McNeil M., Waters I., Moolhuijzen P., Conocono E., Appels R. (2008) The genome structure of the *1-FEH* genes in wheat (*Triticum aestivum* L.): New markers to track stem carbohydrates and grain filling QTL in breeding. Mol Breeding 22:339-351.

Zhao H.H., Fernando R.L., Dekkers J.C.M. (2007) Effects of population structure on power and precision of regression-based linkage disequilibrium mapping of QTL. Journal of Animal Science 85:63-63.

Zheng S.S., Byrne P.F., Bai G.H., Shan X.Y., Reid S.D., Haley S.D., Seabourn B.W. (2009) Association analysis reveals effects of wheat glutenin alleles and rye translocations on dough-mixing properties. Journal of Cereal Science 50:283-290. DOI: DOI 10.1016/j.jcs.2009.06.008.

Ziegelhoffer, E., Leonard J., Elliot M. (2000) Cloning of the Arabidopsis WIGGUM gene identifies a role for farnesylation in meristem development. PNAS 97:7633-7638.

APPENDIX

Table A.1. Lists of lines in the spring wheat association mapping (WAMII) evaluated in five environments.

GID	Entry	Cross Name
450975	9001	COOK/VEE//DOVE/SERI/3/BJY/COC
610288	9002	JUP/ZP//COC/3/PVN/4/GEN
41868	9003	PFAU/VEE#5
3895	9004	SERI M 82
42893	9005	VORONA/GEN
601549	9006	KAUZ/GEN
80836	9007	KEA/TOW//LIRA
613415	9008	LIRA/URES//MILO 9G19-2-26
552808	9009	MYNA/VUL//PRL
3828077	9010	JUPARE C 2001
68315	9011	TIA.3
85599	9012	CHOIX M 95
222973	9013	KAUZ*2/FN//KAUZ
295261	9014	KAUZ*2/MNV//KAUZ
80662	9015	PAT10/ALD//PAT72300/3/PVN/4/URES/5/PFAU
130819	9016	PRINIA
222939	9017	CAR422/ANA//URES

Table A.1. Continued

GID	Entry	Cross Name
4248	9018	CIANO T 79
222912	9019	HD2206/HORK//BUC/BUL
144712	9020	PAT10/ALD//PAT72300/3/PVN/4/BOW
42423	9021	VORONA/CNO79
595693	9022	AZ//KAL/BB/3/PGO
16122	9023	BACANORA T 88
547311	9024	FCT/3/GOV/AZ//MUS/4/DOVE/BUC
294897	9025	INIFAP M 97
295213	9026	KAUZ*2//DOVE/BUC/3/KAUZ
1065811	9027	KAUZ*2/TRAP//KAUZ
294568	9028	LIRA/BUC
294800	9029	PARA2//JUP/BJY/3/VEE/JUN/4/2*KAUZ
434375	9030	RHEA
294705	9031	TOBARITO M 97
458377	9032	TRAP#1/BOW
294548	9033	TURACO/CHIL
1066760	9034	KAUZ*3//TC*6/RL5406(RL6043)
1066924	9035	KAUZ*2//TC*6/RL6081/3/KAUZ

Table A.1. Continued.

GID	Entry	Cross Name
1066245	9036	KAUZ*2/YACO//KAUZ
421503	9037	PRL/VEE#6
217743	9038	UP 2338
377174	9039	ATTILA//ALTAR 84/AOS/3/ATTILA
1339454	9040	FANG60/7C
358957	9041	HP 1761
30709	9042	HYBRID DELHI 2172
369673	9043	KAUZ//ALTAR 84/AOS
1101307	9044	MNCH/3*BCN
16004	9045	RAYON F 89
1339633	9046	SERI/7C
370194	9047	SERI/NKT//2*KAUZ
342263	9048	STAR//KAUZ/STAR
342253	9049	TILHI
3828077	9050	JUPARE C 2001
342152	9051	URES/RAYON
1301292	9052	COMARA/TEG//WEAVER/3/LAJ3302
1314999	9053	PICUS/4/CS(5A)/5RL-1//BUC/BJY/3/ALD/PVN/5/LAJ3302

Table A.1. Continued.

GID	Entry	Cross Name
523085	9054	TIA.2/KAUZ
358954	9055	HP 1731
1212664	9056	IAS62/ALDAN//2*SKAUZ
1811686	9057	KAUZ/RAYON
2460025	9058	KEA/TAN/4/TSH/3/KAL/BB//TQFN/5/PAVON/6/SW89.3064
2460334	9059	PASTOR/2*SITTA
391994	9060	BHRIKUTI
2668073	9061	BL 1724
3628874	9062	BOW/PRL*3/6/WRM/4/FN/3*TH//K58/2*N/3/AUS-6869/5/PELOTAS-ARTHUR/7/HE1/3*CNO79//2*SERI
3599378	9063	BUC/PRL//WEAVER
3582667	9064	CHUM18/5*BCN
2460291	9065	LAJ3302/2*MO88
3613474	9066	MILAN/3/JUP/BJY//URES
2668072	9067	NL 750
376804	9068	PUNJAB 96
1302305	9069	RABE/2*MO88
3591880	9070	SW89-5124*2/FASAN
3600263	9071	TIA.4/WL6572//RL6043/3*GEN/3/LUAN

Table A.1. Continued.

GID	Entry	Cross Name
1082613	9072	CAZO/KAUZ//KAUZ
4085042	9073	CHIR1//SHA5/WEAVER
3616330	9074	HUITES/4/CS/TH.SC//3*PVN/3/MIRLO/BUC
3597332	9075	KAUZ//BOW/NKT
2463808	9076	KAUZ/WEAVER
3586080	9077	MINO
1563731	9078	OASIS/SKAUZ//4*BCN
4097301	9079	SHA3/SERI//SHA4/LIRA/3/CHIR1/4/SHA7//PRL/VEE#6/3/FASAN
3578100	9080	SW89.5181/KAUZ
3592850	9081	W462//VEE/KOEL/3/PEG//MRL/BUC
3619633	9082	GUAM92/KAUZ
4320047	9083	CHEN/AEGILOPS SQUARROSA (TAUS)//BCN/3/2*KAUZ
3605299	9084	CHEN/AEGILOPS SQUARROSA (TAUS)//BCN/3/KAUZ
4319277	9085	CMH84.3379/CMH78.578//MILAN
4048654	9086	HUAYTU CIAT
4318107	9087	OTUS
3617481	9088	TARACHI F 2000
3612293	9089	SITE/MO/4/NAC/TH.AC//3*PVN/3/MIRLO/BUC

Table A.1. Continued.

GID	Entry	Cross Name
3592403	9090	SKAUZ*2/FCT
2406044	9091	SURUTU-CIAT
3587319	9092	TAURUM
1491661	9093	INQALAB 91
3827755	9094	SKAUZ*2/FCT
3585839	9095	CNDO/R143//ENTE/MEXI_2/3/AEGILOPS SQUARROSA (TAUS)/4/WEAVER/5/2*KAUZ
4316539	9096	KETUPA*2/PASTOR
4755104	9097	WEAVER/3/SAPI/TEAL//HUI/4/CROC_1/AE.SQUARROSA (213)//PGO/5/SKAUZ*2/SRMA
4755489	9098	KAUZ*2/TRAP//KAUZ/3/PASTOR/4/SKAUZ*2/SRMA
4755706	9099	REH/HARE//2*BCN/3/CROC_1/AE.SQUARROSA (213)//PGO/4/HUITES
3828077	9100	JUPARE C 2001
4755979	9101	ATTILA/3*BCN//BAV92/3/TILHI
4757265	9102	CROC_1/AE.SQUARROSA (205)//BORL95/3/PASTOR
4760307	9103	CHEN/AE.SQ//WEAVER/3/SSERI1
4763836	9104	BAV92/3/OASIS/SKAUZ//4*BCN/4/PASTOR
4881156	9105	CROC_1/AE.SQUARROSA (205)//KAUZ/3/2*KAUZ*2/YACO//KAUZ
4905340	9106	WBLL1*2/KUKUNA
4905617	9107	ROELFS F2007

Table A.1. Continued

GID	Entry	Cross Name
3820458	9108	HD2687
3822974	9109	PBW450
4754390	9110	MILAN/S87230//BAV92
4756035	9111	ATTILA/3*BCN*2//BAV92
4757869	9112	TOBA97/PASTOR
4905071	9113	FRET2*2/4/SNI/TRAP#1/3/KAUZ*2/TRAP//KAUZ
4940559	9114	ALTAR 84/AE.SQUARROSA (221)//3*BORL95/3/URES/JUN//KAUZ/4/WBLL1
4835368	9115	TUKURU//BAV92/RAYON
4835640	9116	SUNSU/CHIBIA
5106646	9117	WBLL1*2/4/YACO/PBW65/3/KAUZ*2/TRAP//KAUZ
5106632	9118	WBLL1*2/VIVITSI
449286	9119	F6.74/BUN//SIS/3/YR/PAM
41372	9120	PASTOR
14103	9121	GALVEZ S 87
2457	9122	PAVON
88208	9123	VEE#8/5/VEE/4/KLTO//S12/J9281.67/3/MO/JUP
88522	9124	IRENA
41830	9125	BB//TOB/CNO67/3/HUAC/4/TI-R/3/BB/PL//SX

Table A.1. Continued.

GID	Entry	Cross Name
42174	9126	CULIACAN T 89
88509	9127	KITE/PGO
88442	9128	PSN/BOW//SERI
88701	9129	GIM/LIRA
88710	9130	URES/JUN//KAUZ
72533	9131	PROINTA FEDERAL
126306	9132	HI.1077
42274	9133	SITTA
640876	9134	URES//BUC/FLK/3/KAUZ
86005	9135	ARIVECHI M 92
82710	9136	TIA.1
85861	9137	RL6043/4*NAC
80512	9138	F60314.76/MRL//CNO79
268922	9139	ATTILA
222478	9140	PFAU/VEE#9
170046	9141	BAU/OPATA
85587	9142	PRINIA
66363	9143	GRANERO INTA

Table A.1.Continued.

GID	Entry	Cross Name
431445	9144	HUITES F 95
222492	9145	KEA/BUC//FCT
4970584	9146	KITE/GLEN
270393	9147	CHIL/BUC
324452	9148	FILIN
515272	9149	PRL/SARA//TSI/VEE#5
3828077	9150	JUPARE C 2001
270402	9151	PJN/BOW//OPATA
270428	9152	VEE/PJN//TUI
324715	9153	PARA2//JUP/BJY/3/VEERY#5.4/JUN/4/TUI
741166	9154	VEE/PJN//2*TUI
270453	9155	URES/BBL//KAUZ/3/KAUZ
346047	9156	ALTAR 84/AEGILOPS SQUARROSA (TAUS)//OPATA
851963	9157	DUCULA//HUI/TUB/3/CAZO
902339	9158	SHUHA
1995922	9159	PROINTA GRANAR
1370653	9160	TZPP/SERI//BUC
1493157	9161	PIFED/DERN

Table A.1. Continued.

GID	Entry	Cross Name
1987914	9162	HIDHAB
35054	9163	W3918A/JUP
346095	9164	JUN/GEN
346200	9165	K134(60)/VEE//BOW/PVN
346303	9166	ESDA/4/BD120/3/GTA/MXP//RUFF/FGO
346403	9167	VEE#8//JUP/BJY/3/F3.71/TRM/4/BCN/5/KAUZ
346459	9168	CNDO/R143//ENTE/MEXI_2/3/AEGILOPS SQUARROSA (TAUS)/4/WEAVER
346479	9169	HXL-F86/2*BAU
1706327	9170	VI/PIFED//VEE#8
393392	9171	GOV/AZ//MUS/3/SARA
358192	9172	MON/IMU//ALD/PVN
1093624	9173	TUI*2/MILAN
217385	9174	KAUZ*2/BOW//KAUZ
1812971	9175	SIMORGH
766786	9176	CAR853/COC//VEE/3/BOW/4/TUI/5/TUI
781213	9177	NANJING 8646/KAUZ//BCN
778966	9178	TUI/3/TMP64/TWN//SDY/4/RAYON
763509	9179	VEE#5/SARA//DUCULA

Table A.1. Continued.

GID	Entry	Cross Name
1328764	9180	CLC89//ESDA/KAUZ/3/BJY/COC//PRL/BOW
2244167	9181	NS-732/HER//KAUZ
1403850	9182	3VASKAR/G303.1M.1.3.2.2.2//KAUZ/3/SKAUZ/4/KAUZ
1405071	9183	TODY/3/JUP/BJY//SARA/4/TRAP#1/BOW/5/NL456/VEE#5
1403557	9184	CROC_1/AE.SQUARROSA (205)//JUP/BJY/3/SKAUZ/4/KAUZ
1498555	9185	PASTOR/3/VEE#5//DOVE/BUC
1558746	9186	VEE#5//PF70354/MUS/3/PIFED/4/OR791432/VEE#3.2
1658710	9187	KAUZ/5/PAT10/ALD//PAT72300/3/PVN/4/BOW
1661139	9188	MRL/BUC//LIRA/5/BB//TOB/CNO67/3/HUAC/4/TI-R/3/BB/PL//SX
1812527	9189	SAAR
2454848	9190	MNCH/3*BCN
1389162	9191	JUP/BJY//URES/3/HD2206/HORK//BUC/BUL
1395073	9192	SITE/PIOS
2478018	9193	PASTOR//SITE/MO/3/CHEN/AEGILOPS SQUARROSA (TAUS)//BCN
3616959	9194	FILIN/IRENA/5/CNDO/R143//ENTE/MEXI_2/3/AEGILOPS SQUARROSA (TAUS)/4/WEAVER
3567684	9195	CROC_1/AE.SQUARROSA (205)//KAUZ/3/ENEIDA
3630926	9196	F60314.76/MRL//CNO79/3/CHIL/PRL
3669874	9197	PASTOR/BAV92

Table A.1. Continued.

GID	Entry	Cross Name
2671579	9198	BARBET1
2671697	9199	MILVUS2
3828077	9200	JUPARE C 2001
2601142	9201	WEEBILL1
2601477	9202	KAMBARA1
2672707	9203	BABAX.1B.1B*3/PRL
2672710	9204	PEWIT1
2448313	9205	FRET2
2673150	9206	WEEBILL4
3686320	9207	ATTILA*2/9/KT/BAGE//FN/U/3/BZA/4/TRM/5/ALDAN/6/SERI/7/VEE#10/8/OPATA
3686333	9208	ATTILA*2/PBW65
3686491	9209	BABAX/KS93U76//BABAX
3607146	9210	SUJATA/SERI
2478027	9211	PASTOR/3/MUNIA//CHEN/ALTAR 84/5/CNDO/R143//ENTE/MEXI_2/3/AEGILOPS SQUARROSA (TAUS)/4/WEAVER
3630912	9212	URES/PRL//BAV92
3855011	9213	VOROBEY
3855085	9214	SOROCA
3855762	9215	FILIN/2*PASTOR

Table A.1. Continued.

GID	Entry	Cross Name
3827936	9216	URES/JUN//KAUZ/3/BAV92
3853128	9217	NAI60/HN7//BUC/3/PSN/BOW//TUI
3686338	9218	ATTILA*2/4/CAR//KAL/BB/3/NAC
2673154	9219	KAMBARA2
3822784	9220	PRL/2*PASTOR
3827649	9221	PBW65/2*PASTOR
3872312	9222	ATTILA*2/PASTOR
3868699	9223	SERI*3//RL6010/4*YR/3/PASTOR/4/BAV92
3888096	9224	PASTOR//HXL7573/2*BAU
3827938	9225	SOKOLL
4314513	9226	CROC_1/AE.SQUARROSA (213)//PGO/3/BAV92
4556647	9227	MILAN/KAUZ//PRINIA/3/BAV92
4315350	9228	ALTAR 84/AE.SQUARROSA (221)//PASTOR/3/PASTOR
4563437	9229	MILAN/KAUZ//PASTOR
4563443	9230	FLORKWA-1/DHARWAR DRY
4563455	9231	CNDO/R143//ENTE/MEXI_2/3/AEGILOPS SQUARROSA (TAUS)/4/WEAVER/5/PASTOR
4563461	9232	VEBOW/IRENA
4563470	9233	PASTOR/DHARWAR DRY

Table A.1. Continued.

GID	Entry	Cross Name
4563487	9234	BJY/COC//PRL/BOW/3/FRTL
4342318	9235	PASTOR//HXL7573/2*BAU
4569255	9236	SRMA/TUI//PASTOR
4577785	9237	SKAUZ/PASTOR/3/CROC_1/AE.SQUARROSA (224)//OPATA
4577847	9238	CNO79//PF70354/MUS/3/PASTOR/4/BAV92
4577963	9239	MILAN/KAUZ/3/URES/JUN//KAUZ/4/CROC_1/AE.SQUARROSA (224)//OPATA
4753157	9240	KABY/BAV92/3/CROC_1/AE.SQUARROSA (224)//OPATA
4753188	9241	BOW//BUC/BUL/3/KAUZ/4/BAV92/5/MILAN/KAUZ
4578411	9242	PASTOR//MILAN/KAUZ/3/VEE/PJN//2*TUI
4578503	9243	BJY/COC//PRL/BOW/3/MILAN/KAUZ/4/BAV92
4578860	9244	KAUZ/BAV92/3/BJY/COC//PRL/BOW
4564440	9245	FRAME/BUCHIN
4799210	9246	TEMPORALERA M 87*2/KONK
4883041	9247	FRAME*2/3/URES/JUN//KAUZ
4882998	9248	CROC_1/AE.SQUARROSA (224)//OPATA/3/PASTOR/4/PASTOR*2/OPATA
4885594	9249	RL6043/4*NAC//2*PASTOR
3828077	9250	JUPARE C 2001
4885599	9251	CROC_1/AE.SQUARROSA (224)//OPATA/3/PASTOR/4/JARU

Table A.1. Continued.

GID	Entry	Cross Name
4879809	9252	ALTAR 84/AE.SQ//2*OPATA/3/PIFED
4963944	9253	KRICHAUFF/2*PASTOR
4878600	9254	KABY//2*ALUBUC/BAYA
4878569	9255	CNO79//PF70354/MUS/3/PASTOR/4/CROC_1/AE.SQUARROSA (224)//OPATA
4878677	9256	BUC/MN72253//PASTOR/3/BAV92
4961148	9257	SCA/AE.SQUARROSA (409)//PASTOR/3/PASTOR
4961206	9258	CROC_1/AE.SQUARROSA (224)//OPATA/3/BJY/COC//PRL/BOW/4/BJY/COC//PRL/BOW
4961235	9259	CHEN/AE.SQ//2*OPATA/3/BAV92/4/JARU
4961444	9260	TIE CHUAN 1*2/3/HE1/3*CNO79//2*SERI
4934637	9261	ALTAR 84/AEGILOPS SQUARROSA (TAUS)//OPATA/3/ATTILA
4774392	9262	OASIS/5*BORL95/5/CNDO/R143//ENTE/MEXI75/3/AE.SQ/4/2*OCI
3844835	9263	PASTOR//TRAP#1/BOW/3/CHEN/AEGILOPS SQUARROSA (TAUS)//BCN
3822784	9264	PRL/2*PASTOR
5535278	9265	ND643/2*WAXWING
5535298	9266	ND643//2*PRL/2*PASTOR
5535434	9267	KIRITATI//2*PRL/2*PASTOR
5535482	9268	KIRITATI//2*ATTILA*2/PASTOR
5552132	9269	SAAR/WBLL1

Table A.1. Continued.

GID	Entry	Cross Name
5552181	9270	CHONTE
5551747	9271	CHEWINK
5551765	9272	WHEAR/KIRITATI/3/C80.1/3*BATAVIA//2*WBLL1
5551787	9273	WHEAR/4/SNI/TRAP#1/3/KAUZ*2/TRAP//KAUZ/5/C80.1/3*BATAVIA//2*WBLL1
5551820	9274	WHEAR/KUKUNA/3/C80.1/3*BATAVIA//2*WBLL1
5551860	9275	WHEAR/JARU/3/C80.1/3*BATAVIA//2*WBLL1
5551870	9276	WHEAR/TUKURU/3/C80.1/3*BATAVIA//2*WBLL1
5551892	9277	WHEAR/KURUKU/3/C80.1/3*BATAVIA//2*WBLL1
5551918	9278	WHEAR//2*PRL/2*PASTOR
5551926	9279	WHEAR//2*PRL/2*PASTOR
5534314	9280	CNDO/R143//ENTE/MEXI_2/3/AEGILOPS SQUARROSA (TAUS)/4/WEAVER/5/2*KAUZ/6/PRL/2*PASTOR
5534324	9281	PRL/2*PASTOR/4/CHOIX/STAR/3/HE1/3*CNO79//2*SERI
5534339	9282	PRL/2*PASTOR/4/CHOIX/STAR/3/HE1/3*CNO79//2*SERI
5534344	9283	PFAU/MILAN/5/CHEN/AEGILOPS SQUARROSA (TAUS)//BCN/3/VEE#7/BOW/4/PASTOR
5534451	9284	PRL/SARA//TSI/VEE#5/3/TILHI/4/ATTILA/2*PASTOR
5551628	9285	ELVIRA/5/CNDO/R143//ENTE/MEXI75/3/AE.SQ/4/2*OCI/6/VEE/PJN//KAUZ/3/PASTOR
5534403	9286	HEILO//MILAN/MUNIA
4755013	9287	KAUZ//ALTAR 84/AOS/3/MILAN/KAUZ/4/HUITES

Table A.1. Continued.

GID	Entry	Cross Name
5398757	9288	QUAIU #1
5398462	9289	PAURAQUE
5398279	9290	FRET2*2/BRAMBLING
5398125	9291	BECARD
5343251	9292	CROC_1/AE.SQUARROSA (205)//BORL95/3/PRL/SARA//TSI/VEE#5/4/FRET2
5343245	9293	CROC_1/AE.SQUARROSA (205)//BORL95/3/PRL/SARA//TSI/VEE#5/4/FRET2
5344026	9294	BETTY/3/CHEN/AE.SQ//2*OPATA

Table A.2. Meteorological data for Greeley in 2010.

Month	Tmax (^{O}C)	Tmin (^{O}C)	Precipitation (mm)	Wind speed (mph)	Soil Temp (^{O}C)	RH (%)†
10-Jan	4.73	-10.40	0.76	11.91	-2.56	45.25
10-Feb	4.66	-9.30	6.604	14.28	-1.48	38.56
10-Mar	12.65	-4.06	7.11	20.09	1.51	30.13
10-Apr	16.69	0.26	84.33	25.37	6.05	25.67
10-May	20.57	4.34	50.4	24.48	9.99	27.45
10-Jun	28.15	11.86	80.52	21.72	17.15	26.67
10-Jul	31.14	13.63	41.66	19.32	19.48	25.41
Total			271.02			

† RH (%), Relative humidity in percentage

Table A.3. Meteorological data of Greeley 2011.

Month	Tmax (^{O}C)	Tmin (^{O}C)	Prec (mm)	Wind speed (mph)	Soil T (^{O}C)	RH (%) †
11-Jun	4.55	-11.94	0.51	17.25	-1.05	42.58
11-Feb	6.61	-12.94	1.27	20.81	-1.61	30.79
11-Mar	14.81	-3.66	5.08	22.27	2.85	20.48
11-Apr	16.99	0.24	21.34	28.91	6.25	21.99
11-May	18.87	4.23	97.28	22.36	9.39	33.59
11-Jun	28.99	10.59	20.32	22.19	15.24	18.82
11-Jul	32.64	15.04	27.18	19.32	20.08	22.39

† RH (%), Relative humidity in percentage

Table A.4. Metrological data of the experimental year (January 2011-February 2012) at Melkassa, Ethiopia.

Month	T_{min} (^{O}C)	T_{max} (^{O}C)	Soil T (^{O}C)	Wind speed (m/sec)	RH (%) †	total Rainfall (mm)
January	-	-	25.4	2.14	48	0.0
February	-	-	27.0	2.33	38	1.5
March	9.3	28.3	27.3	2.64	36	37.9
April	10.7	32.8	28.8	2.41	38	45.6
May	11.9	31.6	28.2	2.14	47	38.2
June	11.5	31.0	28.2	2.27	52	102.0
July	10.3	27.6	25.2	2.31	64	126.4
August	10.4	26.2	24.8	1.93	71	208.8
September	8.8	26.7	23.5	1.25	71	197.5
October	4.5	28.6	25.5	1.79	41	0.0

† RH (%), Relative humidity in percentage

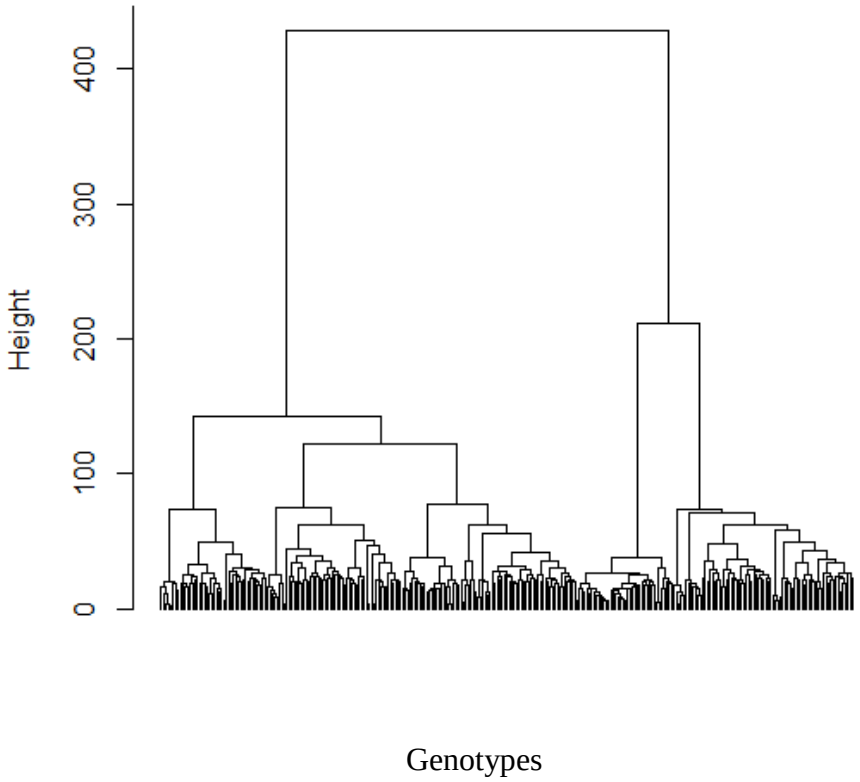

Figure A.1. Dendrogram of 287 spring wheat with 1864 DArT markers

Table A.5. Genotypic correlation among yield and yield component traits at Greeley 2010 under full irrigation.

Trait†	TKW	HI	YLD	KN	SL	SPN	BM	KNS	KWS	SN	KNL	TW	SKW	SKH
TKW	1													
HI	-0.06 ns	1												
YLD	0.22 **	0.754 **	1											
KN	-0.42 **	0.745 **	0.770 **	1										
SL	0.39 **	-0.11 ns	-0.11 ns	-0.36 **	1									
SPN	-0.41 **	0.09 ns	-0.15 *	0.12 *	0.03 ns	1								
BM	0.33 **	-0.17 **	0.47 **	0.18 **	0.33 **	-0.25 **	1							
KNS	-0.20 **	0.50 **	0.47 **	0.60 **	-0.21 **	0.41 **	-0.03 ns	1						
KWS	0.45 **	0.47 **	0.58 **	0.28 **	0.05 ns	0.12 *	0.20 **	0.77 **	1					
SN	-0.23 **	0.52 **	0.19 **	0.32 **	-0.02 ns	-0.48 **	0.36 **	-0.54 **	-0.66 **	1				
KNL	-0.05 ns	0.51 **	0.55 **	0.58 **	-0.24 **	0.01 ns	0.04 ns	0.91 **	0.80 **	-0.43 **	1			
TW	0.00 ns	0.10 ns	0.37 **	0.30 **	-0.03 ns	-0.03 ns	0.36 **	0.29 **	0.28 **	0.068 ns	0.31 **	1		
SKW	0.98 **	-0.07 ns	0.23 **	-0.43 **	0.42 **	-0.42 **	0.36 **	-0.23 **	0.43 **	-0.21 **	-0.08 ns	0.01 ns	1	
SKH	-0.32 **	0.10 ns	0.01 ns	0.25 **	-0.15 *	0.05 ns	0.01 ns	0.11 ns	-0.13 ns	0.29 **	0.07 ns	0.13 *	-0.33	1

†SKD, Single kernel diameter; SKW, Single kernel weight; KNL, Kernel number per spikelet; KWS; Kernel weight per spike; TW, Test weight; SPN, Spikelet number; SL, spike length; GN, Grain number; TKW, Thousand kernel weight; HI, Harvest index; SN, Spikes number m^{-2}; SKH, single kernel hardness; KNS=Kernel number $spike^{-1}$.

*=significant at 5%, **=significant at 1%, ns=non-significant.

Table A.6. Genotypic correlation among morphological, phenological and drought related traits at Greeley in 2010 under full irrigation.

Trait†	LL	LW	DH	DM	PHT	LS	NDVI	GA	LA	GFP	YLD	BM
LL	1											
LW	0.49**	1										
DH	0.04 ns	0.010 Ns	1									
DM	-0.02 ns	0.12	0.68 **	1								
PHT	0.37 **	0.17	0.15	0.19	1							
LS	-0.18	-0.17	-0.52 **	-0.78 **	-0.47 **	1						
NDVI	0.65 **	-0.09 Ns	0.27 **	0.37 **	0.70 **	-0.84 **	1					
GA	0.16	-0.14	0.18	0.42 **	0.13	-0.59 **	0.99 **	1				
LA	0.90 **	0.82 **	0.04 ns	0.05	0.31 **	-0.22 **	0.39 **	0.02	1			
GFD	-0.06	0.14	-0.39 **	0.41 **	0.06	-0.33 **	0.14	0.32 **	0.03	1		
YLD	-0.11	-0.18	-0.32 **	-0.21 **	-0.05	0.20 **	-0.13	-0.08	-0.20	0.12	1	
BM	0.26 **	0.01 Ns	-0.05	0.18	0.63 **	-0.53 **	0.36 **	0.42 **	0.16	0.27 **	0.47 **	1

†LL, Flag leaf length; LW, Flag leaf width; DH, Days to heading; DM, Days to maturity; PHT, Plant height; LS, Leaf senescence; NDVI, Normalized difference Vegetation index; GA, Green area; LA, Leaf area; GFD, Grain filling duration; YLD, Grain yield; BM, Biomass; *=significant at 5%, **=significant at 1%, ns=non-significant.

Table A.7. Genotypic correlation among yield and yield component traits at Greeley in 2011 under full irrigation condition (below diagonal) and moisture stress (above diagonal).

Trait†	TKW	HI	YLD	KN	SL	SPN	TW	SN	BM	KNS	KWS	KNL	SKW	SKD	SKH
TKW	1	0.02 ns	0.15 **	-0.57 **	0.39 **	-0.38 **	0.42 **	-0.17 **	0.21 **	0.63 **	0.50 **	-0.61 **	0.99	0.92 **	-0.29 **
HI	0.54 **	1	0.12 *	0.00	-0.13 *	0.38 **	0.27 **	-0.14 *	-0.51 **	-0.08 ns	-0.04 ns	-0.22 **	-0.30 **	-0.19 **	-0.13 *
YLD	0.11 ns	0.73 **	1	0.71 **	0.20 **	-0.12 *	0.38 **	0.75 **	0.63 **	0.10 ns	0.35 **	0.27 **	0.12 *	0.00	-0.13 *
KN	-0.09 ns	0.44 **	0.99 **	1	-0.12 ns	0.12 *	0.00ns	0.82 **	0.30 **	0.43 **	-0.13 *	0.57 **	-0.66 **	-0.65 **	0.14 *
SL	0.22**	-0.27 **	0.00	-0.10 ns	1	-0.14 *	0.23 **	0.11 Ns	0.29 **	-0.26 **	0.17 **	-0.30 **	0.52 **	0.44 **	-0.16 **
SPN	-0.22 **	-0.30 **	-0.45 **	-0.51 **	0.01 ns	1	-0.23 **	-0.38 **	-0.30 **	0.78 **	0.46 **	0.42 **	-0.41 **	-0.40	-0.03 ns
TW	0.36 **	0.51 **	0.48 **	0.51 **	-0.07 ns	-0.26	1	0.29 **	0.23 **	-0.37 **	0.07 ns	-0.37 **	0.42 **	0.40 **	0.06 ns
SN	-0.44 **	0.04 ns	0.99 **	0.99 **	0.36 **	-0.99 **	0.07 ns	1	0.33 **	-0.20 **	-0.36 **	0.03 ns	-0.30 **	-0.33 **	0.10 ns
BM	0.65 **	0.76 **	0.84**	0.14 **	0.30 **	-0.63 **	0.47 **	-0.01 Ns	1	0.16 **	0.46 **	0.49 **	0.15 **	-0.07 ns	-0.16 **
KNS	-0.75 **	0.45	0.23**	0.85 **	-0.61 **	0.41 **	0.20 **	0.66 **	-0.99 **	1	0.38 **	0.89 **	-0.58 **	-0.54 **	0.14 *

† TKW, Thousand kernel weight; HI, Harvest index; YLD, Grain yield; KN, Kernel number; SL, Spike length; SPN, Spikelet number; TW, Test weight; SN, Spike number; BM, Biomass; KNS, Kernel number per spike; KWS, Kernel weight per spike; KNL, Kernel number per spikelet; SKW, Single kernel weight; SKD, Single kernel diameter; SKH, Single kernel hardness.
*=significant at 5%, **=significant at 1%, ns=non-significant.

Table A.7. Continued.

Trait†	TKW	HI	YLD	KN	SL	SPN	TW	SN	BM	KNS	KWS	KNL	SKW	SKD	SKH
KWS	0.56 **	0.91 **	0.25**	0.56 **	-0.24 **	0.21 **	0.52 **	-0.46 **	-0.33 **	0.25 **	1	0.25 **	0.73 **	0.58 **	-0.22 **
KNL	-0.51 **	0.62 **	0.56**	0.99 **	-0.56 **	-0.28 **	0.39 **	0.99 **	-0.93 **	0.76 **	0.21 **	1	-0.52 **	-0.46 **	0.22 **
SKW	0.999 **	0.41 **	0.29**	-0.46 **	0.29 **	-0.40 **	0.36 **	-0.99 **	0.34 **	-0.19 **	0.99 **	0.11 ns	1	0.99 **	-0.23 **
SKD	0.99 **	0.23 **	0.18**	-0.52 **	0.21 **	-0.46 **	0.36 **	-0.99 **	-0.15 *	-0.17 **	0.99 **	0.16 **	0.84 **	1	0.09 ns
SKH	-0.23 **	-0.21 **	-0.21**	0.12 ns	-0.13 *	0.01 ns	0.19 *	0.02 Ns	-0.17 **	0.12 *	-0.14 *	0.10 ns	-0.30 **	0.09 ns	1

† TKW, Thousand kernel weight; HI, Harvest index; YLD, Grain yield; KN, Kernel number; SL, Spike length; SPN, Spikelet number; TW, Test weight; SN, Spike number; BM, Biomass; KNS, Kernel number per spike; KWS, Kernel weight per spike; KNL, Kernel number per spikelet; SKW, Single kernel weight; SKD, Single kernel diameter; SKH, Single kernel hardness.
*=significant at 5%, **=significant at 1%, ns=non-significant.

Table A.8. Genotypic correlation among phenological, morphological and drought related traits at Greeley in 2011 under full irrigation condition (below diagonal) and moisture stress (above diagonal**).**

Trait†	DH	DM	LL	LW	GA	GFP	LA	NDVI	YLD
DH	1	0.75 **	0.16 **	0.41 **	0.43 **	0.03 ns	0.25 **	0.49 **	-0.04 ns
DM	0.77 **	1	0.12 *	0.35 **	0.23 **	0.69 **	0.22 **	0.71 **	-0.06 ns
LL	0.10 ns	0.14 *	1	0.80 **	-0.03 ns	0.00	0.97 **	0.39 **	0.10 ns
LW	0.20 **	0.25 **	0.70 **	1	-0.34 **	0.10 ns	0.91 **	0.017 ns	-0.06 ns
GA	0.33 **	0.30 **	-0.13 *	-0.14 *	1	-0.16 **	-0.15 *	0.62 **	0.34 **
GFD	-0.51 **	0.12 *	0.01 Ns	-0.01 ns	-0.1 ns	1	0.06 ns	0.42 **	-0.03 ns
LA	0.14 *	0.18 **	0.95 **	0.86 **	-0.14 *	0.00	1	0.27 **	0.06 ns
NDVI	0.50 **	0.39 **	0.31 **	0.14 *	0.57 **	-0.28 **	0.27 **	1	0.42 **

† DH, Days to heading; DM, Days to maturity; LL, Leaf length; LW, Leaf width; GA, Green area; GFD, Grain filling duration; LA, Leaf area; NDVI, Normalized difference vegetation index; YLD, Grain yield.
*=significant at 5%, **=significant at 1%, ns=non-significant.

Table A.9. Genotypic correlation among agronomic traits at Melkassa under stressed (below diagonal) and non-stressed (above diagonal) in 2011.

Trait†	TKW	DH	DM	HI	Yld	KN	BM	LA	LL	LW	SL	GFP	GA	SPN
TKW	1	-0.55**	-0.47**	0.71**	0.77**	0.77**	0.27**	0.02ns	-0.17**	0.19**	0.19**	0.12*	0.14*	-0.40**
DH	-0.35**	1	0.82**	-0.62**	-0.54**	-0.54**	0.03ns	-0.02ns	0.001ns	-0.031ns	0.13*	-0.28**	0.54**	0.63**
DM	0.68**	0.88**	1	0.59**	-0.44**	-0.44**	0.15*	0.01ns	-0.03ns	0.06ns	0.24**	0.32**	0.54**	0.62**
HI	0.61**	-0.81**	-0.801**	1	0.91**	0.91**	0.081ns	-0.319**	-0.530**	-0.06ns	0.009ns	0.036ns	-0.103ns	-0.40**
YLD	-0.07ns	-0.68**	-0.671**	0.902**	1	0.99**	0.49**	-0.268**	-0.47**	0.016ns	0.12*	0.13*	0.24**	-0.24**
KN	-0.04ns	-0.40**	-0.56**	0.57**	0.74**	1	0.50**	-0.27**	-0.47**	0.016ns	0.12*	0.13*	0.24**	-0.24**
BM	0.34**	0.11ns	0.20**	-0.01**	0.39**	0.53**	1	0.012ns	-0.09ns	0.14*	0.27**	0.17**	0.81**	0.23**
LA	0.25**	0.02ns	0.25**	-0.12**	0.03ns	-0.17**	0.40**	1	0.91**	0.88**	0.29**	0.03ns	-0.18**	0.05ns
LL	0.31**	-0.13*	0.03ns	0.01Ns	0.06ns	-0.08ns	0.14**	0.90**	1	0.59**	0.10ns	-0.04ns	-0.47**	-0.02NS
LW	0.15**	0.39**	-0.25**	-0.03ns	-0.22**	0.53**	0.85**	0.52**	1	0.44**	0.12*	0.20**	0.10ns	
SL	-0.03ns	0.08ns	0.08**	0.20**	0.37**	0.35**	0.77**	0.43**	0.10ns	0.59**	1	0.20**	0.24**	0.66**
GFD	0.68**	-0.90**	-0.56**	0.63**	0.55**	0.16**	0.00	0.15*	0.19**	0.06ns	-0.13*	1	0.01ns	0.03ns
GA	-0.23**	0.66**	0.93**	-0.56**	-0.22**	-0.10Ns	0.68**	0.58**	0.29**	0.66**	0.99**	-0.30**	1	0.45**

†TKW, Thousand kernel weight; DH, Days to heading; DM, Days to maturity; HI, Harvest index; YLD, Grain yield; KN, Kernel number; BM, Biomass; LA, Leaf area; LL, Leaf length; LW, Leaf width; SL, Spike length; GFD; Grain filling duration; GA, Green area; SPN, Spikelet number.
*=significant at 5%, **=significant at 1%, ns=non-significant.

Table A.10. Summary of linkage disequilibrium greater than critical value (r^2>0.2641) for wheat chromosomes.

Chromosome	Total pairs	% r^2 at P<0.01†	% r^2 > 0.2‡	[a] % r^2>0.2641	Average LD at r^2>0.2641	Average LD for all pairs
1A	2016	31.30	7.09	6.15	0.5723	0.06
1B	1771	54.60	19.42	16.37	0.5984	0.14
1D	190	35.26	14.21	13.16	0.5289	0.09
2A	1035	26.67	7.05	6.38	0.6351	0.06
2B	2628	23.17	3.31	2.32	0.5297	0.03
2D	595	36.81	18.99	17.14	0.6935	0.14
3A	666	45.50	7.50	6.16	0.6206	0.08
3B	2556	23.12	3.36	2.74	0.5624	0.03
3D	276	42.39	25.72	24.28	0.5561	0.15
4A	903	62.35	11.74	8.08	0.5243	0.10
4B	253	33.20	4.74	4.35	0.6616	0.05
5A	210	20.00	5.24	4.29	0.8206	0.05
5B	1485	26.60	3.03	2.15	0.5423	0.03
6A	1378	35.78	7.47	6.09	0.5620	0.06

† Percent of r^2 at 0.2 is significant at P<0.001.
‡ Percent of r2 at 0.2641 is significant at P<0.001.

Table A.10. Continued.

Chromosome	Total pairs	% r^2 at P<0.01†	% r^2 > 0.2‡	[a] % r^2>0.2641	Average LD at r^2>0.2641	Average LD for all pairs
6B	2485	33.12	6.24	4.39	0.5236	0.05
6D	55	30.91	14.55	7.27	0.912	0.10
7A	1035	22.51	2.61	2.5	0.5301	0.03
7B	903	44.85	14.51	12.29	0.5807	0.10
7D	300	36.33	15.00	11.67	0.6222	0.10

† Percent of r^2 at 0.2 is significant at P<0.001.
‡ Percent of r2 at 0.2641 is significant at P<0.001.

LISTS OF ABBREVIATIONS

Abbreviation	Description
AFLP	Amplified fragment length polymorphism
BLUE	Best linear unbiased estimators
BM	Biomass
BLZ	Barley leucine zippers
BPBF	Barley prolamin box binding factor
CID	Carbon isotope discrimination
DH	Days to heading
DM	Days to maturity
DSI	Drought susceptibility index
GA	Green leaf area
GAMyB	Gibberellin-regulated Myb factor
GBSSI	Granule bound starch synthase
GXE	Genotype by Environment
GFD	Grain filling duration
GRD11	Greeley dry 2011
GRW10	Greeley wet 2010
GRW11	Greeley wet 2011
GYLD	Grain yield
HI	Harvest index
SKH	Kernel hardness
KN	Kernel number
KNL	Kernel number per spikelet
KWS	Kernel weight per spike
KNS	Kernel number per spike
LL	Flag leaf length
LA	Flag leaf area

Abbreviations	Description
LW	Flag leaf width
LS	Leaf senescence
MAS	Marker assisted selection
MLKW11	Melkassa wet 11
MELKD11	Melkassa dry 11
NDVI	Normalized vegetation index
PHT	Plant height
QTL	Quantitative Trait Locus
RAPD	Random amplified polymorphic DNA
RFLPs	Restriction fragment length Polymorphism
SL	Spike length
SN	Spike number per m^2
SKD	Single kernel diameter
SKW	Single kernel weight
SPN	Spikelet number per spike
SSIIa	Soluble starch synthase
SSR	Simple sequence repeat
VRN	Vernalization
TAGW2	*Triticum aestivum* grain weight gene
TKW	Thousand kernel weight
TW	Test weight