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

Erena A. Edae,* Patrick F. Byrne, Harish Manmathan, Scott D. Haley, Marc Moragues, Marta S. Lopes, and Matthew P. Reynolds

Abstract

Functional markers are needed for key genes involved in drought tolerance to improve selection for crop yield under moisture stress conditions. The objectives of this study were to (i) characterize five drought tolerance candidate genes, namely dehydration responsive element binding 1A (*DREB1A*), enhanced response to abscisic acid (*ERA1-B* and *ERA1-D*), and fructan 1-exohydrolase (*1-FEH-A* and *1-FEH-B*), in wheat (*Triticum aestivum* L.) for nucleotide and haplotype diversity, Tajima's D value, and linkage disequilibrium (LD) and (ii) associate within-gene single nucleotide polymorphisms (SNPs) with phenotypic traits in a spring wheat association mapping panel (*n* = 126). Field trials were grown under contrasting moisture regimes in Greeley, CO, and Melkassa, Ethiopia, in 2010 and 2011. Genome-specific amplification and DNA sequence analysis of the genes identified SNPs and revealed differences in nucleotide and haplotype diversity, Tajima's D, and patterns of LD. *DREB1A* showed associations (false discovery rate adjusted probability value = 0.1) with normalized difference vegetation index, heading date, biomass, and spikelet number. Both *ERA1-A* and *ERA1-B* were associated with harvest index, flag leaf width, and leaf senescence. *1-FEH-A* was associated with grain yield, and *1-FEH-B* was associated with thousand kernel weight and test weight. If validated in relevant genetic backgrounds, the identified marker–trait associations may be applied to functional marker-assisted selection.

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 or 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 such as drought tolerance. Candidate gene association mapping is aimed at linking phenotypic variation with polymorphic sites in candidate genes to identify causative polymorphisms (Martinez-Gonzalez 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 phytohormone abscisic acid (ABA): ABA-independent and ABA-dependent. For example, ABA-independent dehydration responsive element binding (DREB) genes are

E.A. Edae, P.F. Byrne, H. Manmathan, S.D. Haley, and M. Moragues, Dep. of Soil and Crop Sciences, Colorado State Univ., Fort Collins, CO 80523; M.S. Lopes and M.P. Reynolds, CIMMYT, Int., Apdo. Postal 6-641, 06600 Mexico, DF, Mexico. Received 22 Apr. 2013. *Corresponding author (erenakae@yahoo.com).

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* L.), maize (*Zea mays* L.), *Arabidopsis thaliana* (L.) Heynh., and wheat (*Triticum aestivum* L.), 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 *A. thaliana* controlled by the stress-inducible *rd29a* promoter showed greater root branching, increased drought tolerance, and larger spike size than nontransgenic 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 hypogaea* L.) plants, *DREB1A* improved transpiration efficiency (Bhatnagar-Mathur 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* L.) 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, that is, 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 *A. thaliana* (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 *A. thaliana* 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. Compared to other marker types, single nucleotide polymorphisms (SNPs) are generally more abundant, stable, amenable to automation, efficient, and cost effective (Rafalski, 2002; Akhunov et al., 2010). Single nucleotide polymorphisms can be individually responsible for phenotypic variation of a trait or linked to causative SNPs (Langridge and Fleury, 2011). 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).

Different methods can be used to discover SNPs. However, the most straightforward approach is direct resequencing of amplicons of genes from different genotypes (Rafalski, 2002). Amplification of DNA segments with genome-specific primers for polyploids such as 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 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., semidwarfism 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). For successful functional marker development, prior information about the level of DNA polymorphism, extent of linkage disequilibrium (LD), 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 (i) characterize five drought tolerance candidate genes in wheat for nucleotide and haplotype diversity, Tajima's D value, and LD and (ii) determine the association between within-gene SNPs and phenotypic traits in a spring wheat association mapping panel.

Materials and Methods
Plant Materials
A total of 126 lines was selected from a set of 294 spring wheat lines of an association mapping panel, which was developed by the International Maize and Wheat Improvement Center (CIMMYT) from entries in the Elite Spring Wheat Yield Trial (26th, 27th, and 28th), Semiarid Wheat Yield Trial (first to 16th) and High

Temperature Wheat Yield Trial (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 subsampling. The lines in the subsample 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 (40.45° N, 104.64° W, and elevation 1427 m), in 2010 and 2011 and at Melkassa, Ethiopia (8.40° N, 39.33° E, and 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 to 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, unpublished bulletin, 2011).

On 5 Apr. 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 3 mo (January through July).

In both years each entry was replicated twice in a Latinized row–column design prepared with CycDesign 3.0 software (VSN International, 2005). 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 in wet soil and on 19 July 2011 in drier soil in an adjacent field. The experiment was laid out as an α 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 $(NH_4)2HPO_4$ (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 length (measured in centimeters from leaf collar to the tip) and maximum width (measured in centimeters 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 × flag leaf width × 0.75.

Plant height was recorded as the average of three values measured in centimeters 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 difference 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.). A green leaf area index was obtained from a photo taken at heading at a height of approximately 0.50 m directly above each plot with a digital camera (Coolpix S8100; Nikon Corp.) and processed with Breedpix software (Casadesus et al., 2007). Leaf senescence was scored during the grain filling stage a week before physiological maturity on a scale from 0 to 10, in which 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 h in a 40°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 plus biomass grain weight) divided by the plot area and expressed as kilograms per hectare. 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 (Model 900-2; International Marketing and Design Corp.) and obtaining the weight of the seeds. Number of spikes per square meter was calculated by dividing the number of kernels per square meter by kernel number per spike. The number of kernels per square meter was obtained from the ratio of grain yield (expressed as g m^{-2}) to thousand kernel weight, multiplied by 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

Table 1. Forward and reverse primer sequences used to amplify five 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′

Instruments). 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 version 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 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 × 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. Therefore, 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 [SAS Institute, 2011]) as h^2 = genotypic variance/(genotypic variance + error variance/r), in which r = number of replications for a single environment. For combined data, heritability estimates were calculated as genotypic variance/[genotypic variance + (genotype × environment interaction variance/n) + (error variance/nr)], in which n = number of environments. Pearson phenotypic correlation coefficients among traits were obtained using the CORR procedure in SAS.

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, 1-FEH and DREB1A, were obtained from the National Center for Biotechnology Information GenBank database (http://www.ncbi.nlm.nih.gov/gene, accessed 26 Jan. 2011). 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 polymerase chain reaction (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. The primers used to amplify ERA1 and 1-FEH were designed for genome-specific amplification with primers designed from unique regions of the genes using primer3 software (Rozen and Skaletsky, 2000). Genome specificity of the primers was verified by PCR amplification of the corresponding gene in the hexaploid wheat progenitors *Triticum urartu* Tumanian ex Gandilyan (AA, $2n = 2x = 14$), *Aegilops speltoides* Tausch (BB, $2n = 2x = 14$), and *Aegilops tauschii* Coss. (DD, $2n = 2x = 14$). A complete list of the genome-specific primers is given in Table 1.

To extract DNA, leaf tissues were sampled from 2-wk-old seedlings of the 126 lines grown in a greenhouse. The leaf samples were immediately transferred to 2-mL tubes and stored at –80°C. The DNA was extracted following a standard cetyl trimethyl ammonium bromide 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 deoxyribonucleotide triphosphates (dNTPs), 1.5 mM $MgCl_2$, 0.5 U of VELOCITY DNA polymerase (Bioline USA Inc. [http://www.bioline.com/, accessed 17 Mar. 2011]), and 3% dimethyl sulfoxide was used. The PCR was performed on a MJ PTC-200 programmable thermal controller (MJ Research) as follows: initial denaturation at 98°C for 2 min; 30 cycles of 98°C for 1 min, an annealing step at 68°C for ERA1 and 64°C for 1-FEH for 1 min, and 72°C for 1.5 min; and final extension at 72°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). The PCR amplification was done on a MJ PTC-200 programmable thermal controller at an initial denaturation temperature of 94°C for 3 min followed by 34 amplification cycles at 94°C for 1 min, annealing temperature of 63°C for 1 min, and 72°C for 1.5 min, and final extension at 72°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 to 2% agarose gels, stained with ethidium bromide, and visualized under

ultraviolet light. The amplified PCR products were purified and sequenced on an Applied Biosystems Inc. sequencing instrument at Beckman Coulter Genomics (Beckman Coulter Genomics). 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 (DNASTAR, 2008). 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 LD using DnaSP version 5 software (Rozas et al., 1999). Tajima's D statistic was also calculated in DnaSP from the normalized differences between the number of segregating sites (S) and the average number of nucleotide diversity (π), and Tajima's D significance test was conducted with the assumption that D follows β distribution (Tajima, 1989). Genome-specific sequences of the *ERA1* gene were mapped to the long arms of chromosomes 3A, 3B, and 3D through a basic local alignment search tool (BLAST) search against the survey sequences of all individual chromosomes of bread wheat in the International Wheat Sequencing Consortium database (http://wheat-urgi.versailles.inra.fr/Seq-Repository, accessed 4 Feb. 2013).

Population Structure, Linkage Disequilibrium, and Marker–Trait Association Analysis

Diversity Array Technology (DArT) markers (Triticarte Pty. Ltd.; http://www.triticarte.com.au/, accessed 11 June 2013; 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 duration to test a number of populations (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) in which the change of k (Δk) was maximized.

Single nucleotide polymorphisms within each gene were used to determine pairwise LD with GGT2 computer software (Berloo, 2008). 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. The 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.

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 for yield, variation due to genotype × environment interaction was about 5% of the total variation while about 3% 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 nonstressed conditions in 2011 (Table 2). The mean yield performance of genotypes in the nonstressed treatment at Melkassa was about threefold higher than that of the irrigated treatment at Greeley in the same year (Table 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 4 to 7 d in the Greeley environments for the subsampled population, the range at Melkassa was 15 to 18 d (Table 2).

The estimated heritability values for agronomic traits measured in each environment and combined data are presented in Table 2. The variance components from which heritability estimates were calculated are given in Supplemental Table S1. High heritability estimates were obtained for single kernel weight, test weight, and single kernel diameter. Grain yield showed higher and more consistent heritability estimates than yield component traits such as kernel number per square meter, spike number per square meter, kernel number per spikelet, kernel weight per spike, and biomass. Generally, 17 out of 25 traits (68%) had heritability estimates ≥50% in half or more test environments implying the reliability of the phenotypic measurements.

Table 2. Mean values and heritability estimates for 22 traits in a spring wheat association mapping panel (*n* = 126) evaluated in five environments.

| Trait[†] | Environments[‡] | | | | | | | | | | | |
| | GRW10 | | GRW11 | | GRD11 | | MLKW11 | | MLKD11 | | Combined | |
	Mean	Heritability	Mean	Heritability	Mean	Heritability	Mean	Heritability	Mean	Heritability	Mean	Heritability
YLD	2,179	0.60	1,528	0.43	1,301	0.40	4,420	0.68	3,904	0.61	2,690	0.61
TKW	35.45	0.74	21.37	0.31	25.11	0.51	25.53	0.65	22.6	0.65	25.97	0.81
TW	77.84	0.77	65.46	0.65	69.07	0.84					70.83	
DH	67.24	0.83	69.96	0.72	68.29	0.69	54.92	0.89	55.68	0.93	63.23	0.75
DM	103.22	0.82	104.04	0.66	99.58	0.56	92.78	0.83	95.16	0.67	98.96	0.69
GFD	35.98	0.67	34.00	0.37	30.59	0.38	37.86	0.58	39.47	0.63	35.58	0.55
KN	6,368	0.55	7,319	0.31	5,304	0.32	17,610	0.68	17,296	0.48	10,795	0.45
HI	0.25	0.38	0.29	0.58	0.36	0.16	0.28	0.68	0.23	0.57	0.28	0.45
PHT	62.79	0.83	63.38	0.78	49.53	0.76	81.91	0.43			68.12	0.83
NDVI	0.67	0.28	0.4	0.49	0.27	0.32					0.45	0.62
BM	7,798	0.32	4,315	0.074	3,863	0.23	16,246	0.40	17,237	0.32	9,903	0.48
SPN	16.4	0.45	16	0.44	15.8	0.54	16.58	0.73			16.31	0.57
LL	15.84	0.66	15.41	0.45	12.17	0.61	20.179	0.28	22.22	0.46	17.12	0.72
LW	1.51	0.75	1.31	0.42	1.23	0.38	1.29	0.27	1.34	0.23	1.34	0.62
LS	5.43	0.76			7.75	0.44					6.59	0.73
KNS	34.96	0.26	38.69	0.11	38.71	0.38						
KWS	1.24	0.28	0.84	0.11	0.97	0.19						
SN	200.45	0.15	198.94	0.005	140.62	0.28						
SL	9.33		8.89	0.83	8.73	0.60						
KNL	2.12	0.26	2.41	0.14	2.38	0.28						
SKD	2.87	0.12	2.53	0.79	2.56	0.78						
SKW	35.46	0.75	25.71	0.77	27.08	0.75						

[†]YLD, grain yield; TKW, thousand kernel weight; TW, test weight; DH, days to heading; DM, days to maturity; GFD, grain filling duration; KN, kernel number; HI, harvest index; PHT, plant height; NDVI, normalized difference vegetation index; BM, biomass; SPN, spikelet number; LL, flag leaf length; LW, flag leaf width; LA, flag leaf area; LS, leaf senescence; KNS, kernel number per spike; KWS, kernel weight per spike; SN, spikes number per square meter; SL, spike length; KNL, kernel number per spikelet; SKD, single kernel diameter; SKW, single kernel weight.

[‡]GRW10, Greeley wet 2010; GRW11, Greeley wet 2011; GRD11, Greeley dry 2011; MLKW11, Melkassa wet 2011; MLKD11, Melkassa dry 11.

The phenotypic correlation coefficients among major agronomic traits measured in Greeley under irrigated and rainfed conditions are presented in Table 3. Of these phenotypic correlation coefficients, about 45 and 47% were significant under irrigated and rainfed conditions, respectively. However, only 8 and 4% of the phenotypic correlation values were ≥0.5 under irrigated and rainfed conditions, respectively. This shows weak interdependency among many traits measured in this study. Grain yield showed consistently high and positive phenotypic correlation with kernel number per square meter, spike number per square meter, and test weight regardless of the moisture level. Thousand kernel weight had significant and negative correlation coefficients with spike number per square meter and days to heading both under irrigated and rainfed conditions.

Sequence Diversity and Single Nucleotide Polymorphism–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 π, that is, the average number of pairwise nucleotide differences per site (Nei, 1987), per site estimates of diversity parameter (θ), and haplotype diversity for each gene were determined (Table 4). A total of 37 SNPs with minor allele frequency (MAF) 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 4). 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).

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*. This indicates a departure from the assumptions of a neutral model and that those genes are either under selection or undergoing population expansion after a severe bottleneck (Table 4). The negative sign of Tajima's D for

Table 3. Phenotypic correlation coefficients among 17 traits in an association mapping panel (*n* = 126) evaluated at Greeley in 2011 under full irrigation (below diagonal) and moisture stress (above diagonal).

	YLD†	TKW	DH	LW	LL	DM	HI	PHT	SPN	BM	TW	NDVI	GA	KN	SL	KNS	SN
YLD		0.22 *	−0.16 ns‡	−0.02 ns	0.04 ns	−0.04 ns	0.25 **	0.38 **	−0.16 ns	0.20 *	0.38 **	0.27 **	0.24 **	0.71 **	0.19 *	−0.07 ns	0.66 **
TKW	0.26 **		−0.31 **	0.20 *	0.22 *	0.03 ns	−0.01 ns	0.40 **	−0.17 ns	0.18 *	0.32 **	0.18 *	−0.02 ns	0.48 **	0.43 **	−0.24 **	−0.26 **
DH	−0.27 **	−0.33 **		0.19 *	0.19 *	0.49 **	−0.10 ns	0.03 ns	0.07 ns	−0.07 ns	−0.03 ns	0.35 **	0.43 **	0.07 ns	−0.13 ns	0.17 ns	−0.01 ns
LW	−0.06 ns	0.27 **	0.06 ns		0.66 **	0.24 **	−0.06 ns	0.27 **	0.06 ns	0.08 ns	0.06 ns	0.24 **	0.09 ns	−0.13 ns	0.25 **	0.06 ns	−0.13 ns
LL	−0.02 ns	0.05 ns	−0.04 ns	0.65 **		0.23 **	−0.04 ns	0.32 **	−0.02 ns	0.05 ns	0.16 ns	0.31 **	0.20 *	−0.07 ns	0.25 **	0.08 ns	−0.09 ns
DM	−0.06 ns	−0.10 ns	0.45 **	0.16 ns	−0.03 ns		−0.17 ns	0.22 *	−0.01 ns	0.03 ns	0.24 **	0.46 **	0.40 **	−0.09 ns	0.20 *	−0.04 ns	−0.05 ns
HI	0.48 **	0.20 *	−0.35 **	−0.11 ns	−0.10 ns	−0.17 ns		−0.03 ns	0.11 ns	−0.52 **	0.21 *	−0.03 ns	0.02 ns	0.23 *	−0.11 ns	0.05 ns	0.14 ns
PHT	0.28 **	0.06 ns	0.16 ns	0.25 **	0.36 **	0.08 ns	−0.15 ns		−0.19 *	0.25 **	0.29 **	0.41 **	0.23 **	0.07 ns	0.53 **	−0.12 ns	0.13 ns
SPN	−0.10 ns	0.00 ns	0.06 ns	0.08 ns	−0.05 ns	0.01 ns	−0.16 ns	−0.12 ns		−0.07 ns	−0.13 ns	0.01 ns	0.03 ns	−0.04 ns	0.09 ns	0.58 **	−0.38 **
BM	0.28 **	0.07 ns	0.02 ns	0.02 ns	0.08 ns	0.12 ns	0.09 ns	0.23 **	0.08 ns		0.07 ns	0.09 ns	0.04 ns	0.04 ns	0.21 *	−0.01 ns	0.03 ns
TW	0.47 **	0.21 *	−0.21 *	0.04 ns	0.07 ns	−0.03 ns	0.4 **	0.15 ns	−0.11 ns	0.12 ns		0.33 **	0.24 **	0.11 ns	0.20 *	−0.12 ns	0.16 ns
NDVI	0.46 **	0.04 ns	0.22 *	0.05 ns	0.11 ns	0.28 **	0.01 ns	0.68 **	−0.15 ns	0.26 **	0.23 *		0.69 **	0.09 ns	0.27 **	0.02 ns	0.08 ns
GA	0.32 **	−0.12 ns	0.23 *	−0.09 ns	−0.07 ns	0.19 *	0.12 ns	0.29 **	−0.06 ns	0.26 **	0.17 ns	0.53 **		0.22 *	0.04 ns	0.04 ns	0.17 ns
KN	0.69 **	−0.46 **	−0.05 ns	−0.26 **	−0.05 ns	−0.02 ns	0.29 **	0.20 *	−0.08 ns	0.19 *	0.29 **	0.35 **	0.30 **		−0.13 ns	0.11 ns	0.81 **
SL	−0.04 ns	0.21 *	0.06 ns	0.42 **	0.39 **	0.08 ns	−0.29 **	0.52 **	0.22 *	0.11 ns	−0.06 ns	0.24 **	−0.01 ns	−0.18 *		0.05 ns	−0.15 ns
KNS	0.10 ns	0.02 ns	−0.13 ns	−0.00 ns	−0.03 ns	−0.05 ns	0.10 ns	−0.26 **	0.41 **	−0.03 ns	0.09 ns	−0.14 ns	−0.07 ns	0.09 ns	−0.09 ns		0.05 ns
SN	0.39 **	−0.38 **	0.05 ns	−0.18 *	−0.03 ns	−0.04 ns	0.13 ns	0.28 **	−0.22 *	0.17 ns	0.15 ns	0.31 **	0.25 **	0.66 **	−0.02 ns	−0.58 **	

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

†YLD, grain yield; TKW, thousand kernel weight; DH, days to heading; LW, leaf width; LL, leaf length; DM, days to maturity; HI, harvest index; PHT, plant height; SPN, spikelet number; BM, biomass; TW, test weight; NDVI, normalized difference vegetation index; GA, green leaf area; KN, kernel number; SL, spike length; KNS, kernel number per spike; SN, spike number.

‡ns, nonsignificant.

all candidate genes shows the accumulation of too many low frequency SNPs with respect to predictions of the neutral theory (Fusari et al., 2008; Giordani et al., 2011). However, the estimate of Tajima's D was nonsignificant for the *1-FEH-A* gene, indicating the absence of a selection footprint for this gene.

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 5. 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 (Fig. 1, 2, 3, and 4) indicating the inconsistency of LD

within a gene region. 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 approximately 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* (Fig. 5 and 6).

The use of functional markers in marker-assisted plant breeding depends on the degree to which economically important traits are affected by a gene. The SNPs within *DREB1A* were associated with several traits, including final biomass, normalized vegetation index, days to heading, and spikelet number (Table 6). The percentage of phenotypic variation explained by those SNPs ranged from 6.4% for heading date to 9.7% for NDVI. In association mapping, a QTL that explains about 10% of the phenotypic variation can be considered

Table 4. Summary of measures of nucleotide variability in sequences of five drought tolerance candidate gene.

Gene	Sample number	Length (bp)	SNPs (MAF > 0.05)†	Nucleotide diversity	Theta per site	Haplotype diversity	Tajima's D
DREB1A	126	971	16	0.180	0.392	0.948	−1.809*
ERA1-B	122	1410	8 (5 indels)	0.00094	0.0065	0.508	−2.649***
ERA1-D	121	1388	7	0.0023	0.011	0.826	−2.457***
1-FEH-A	126	601	5	0.00224	0.0035	0.45	−0.896 NS‡
1-FEH-B	124	668	1	0.00078	0.0049	0.153	−2.307***

*Significant at the 0.05 probability level.

***Significant at the 0.001 probability level.

†SNP, single nucleotide polymorphism; MAF, minor allele frequency (>0.05).

‡NS, nonsignificant.

Table 5. Linkage disequilibrium (LD) analysis of five drought tolerance candidate genes.

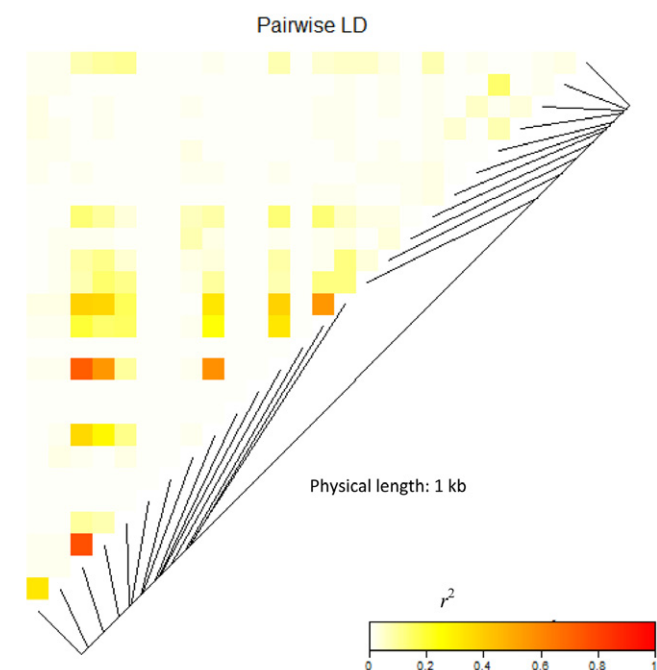| Gene	Number of pairwise comparisons	Number of significant pairwise comparisons (Fisher exact test, P < 0.01)	Percent of significant pairwise comparisons	SNP pairs in complete LD†
DREB1A	121	17	14	None
ERA1-B	28	2	7	None
ERA1-D	21	5	24	None
1-FEH-A	10	4	40	3 (30%)
1-FEH-B	–	–	–	–

†SNP, single nucleotide polymorphism.

Figure 1. Patterns of linkage disequilibrium (LD) in pairwise combinations of single nucleotide polymorphisms in the *DREB1A* gene. The beginning of the sequence is at the lower left corner. Red color indicates a greater degree of LD.

a major QTL according to definitions for QTL classes suggested by Flint-Garcia et al. (2005).

The SNPs within *ERA1-B* were associated with harvest index, spikes per square meter, and grain filling duration either in two environments or in one environment plus combined data across environments. In that gene, SNPs 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 per square meter (11.3%) followed by flag leaf width (10.2%), grain filling duration (9.7%), and harvest index (9.3%). *ERA1-D* was also associated with the yield component traits kernel weight per spike, kernel number per spike, harvest index, flag leaf width, and leaf senescence. The majority of the SNP–trait associations for *ERA1-D* were obtained under rainfed conditions. Both *ERA1-B* and *ERA1-D* were associated with leaf senescence, harvest index, and flag leaf width (Table 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.

The *1-FEH-A* gene was associated with yield, kernel number per spike, spike length, NDVI, biomass, flag leaf length and area, and green leaf area index (Table 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%). 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 6). In genomewide 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, and flag leaf area and width. A previous biparental QTL mapping study detected QTL for stem water soluble carbohydrate, thousand kernel weight, and grain filling efficiency on
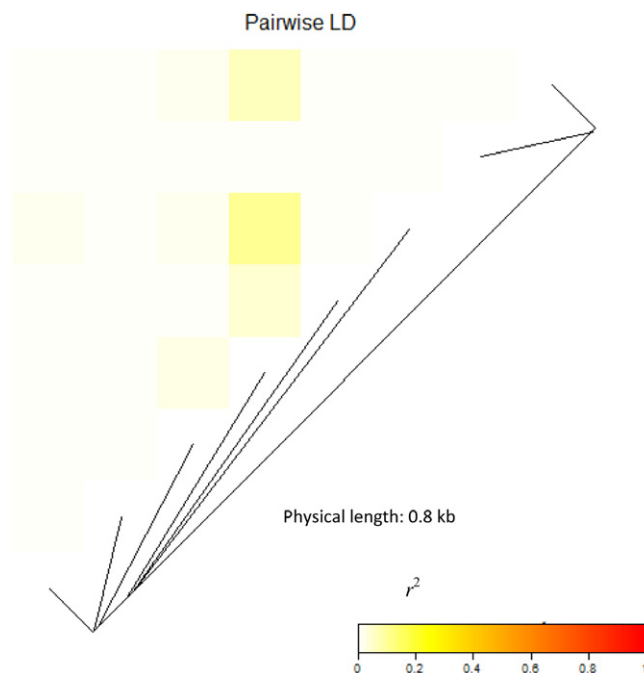
Figure 2. Patterns of linkage disequilibrium (LD) in the *ERA1-B* gene. The beginning of the sequence is at the lower left corner. Red color indicates a greater degree of LD.
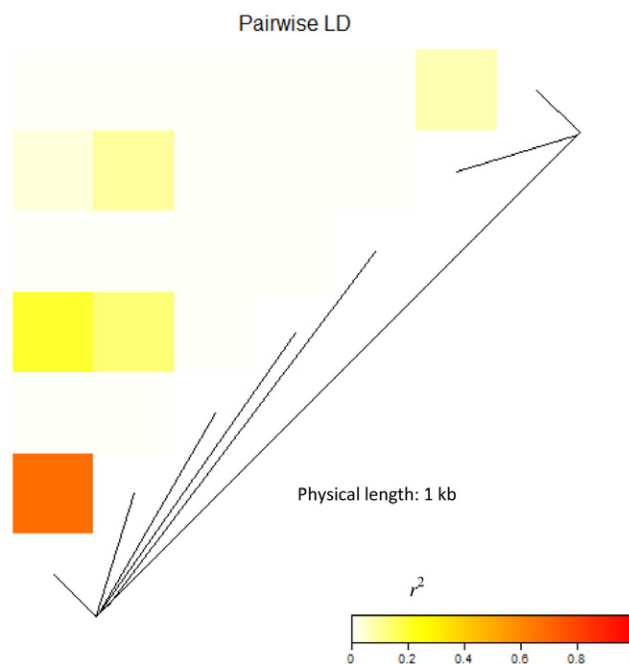

Figure 3. Patterns of linkage disequilibrium (LD) in the *ERA1-D* gene. The beginning of the sequence is at the lower left corner. Red color indicates a greater degree of LD.
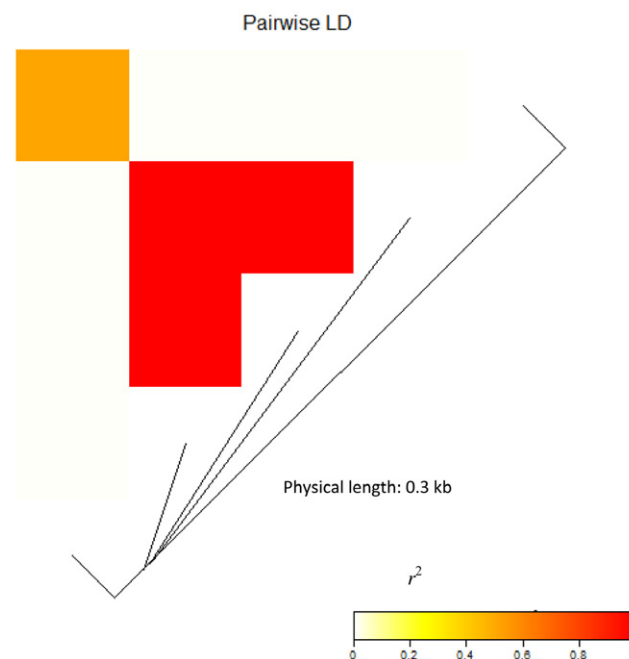

Figure 4. Patterns of linkage disequilibrium (LD) in the *1-FEH-A* gene. The beginning of the sequence is at the lower left corner. Red color indicates a greater degree of LD.

chromosome 6AS (Yang et al., 2007). Therefore, 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.

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 d), 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.

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, 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 5). Similarly, *1-FEH-A* is more diverse than *1-FEH-B*, and the polymorphisms in the former are also in agreement with neutral expectation. Although the significance of Tajima's D test may indicate the presence of selection footprints, this test may not provide complete information about the action of selection, as demographic processes such
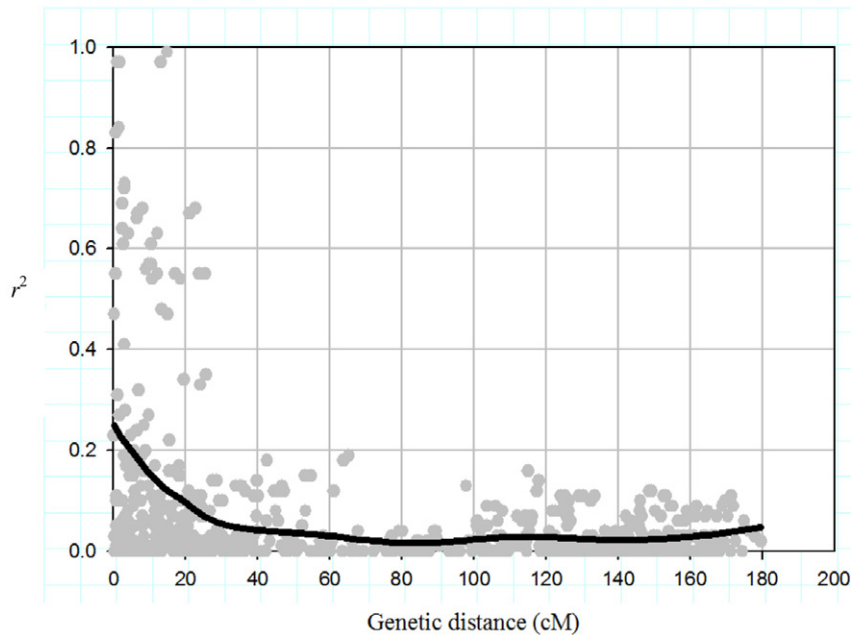
Figure 5. Linkage disequilibrium decay curve for chromosome 3A based on 37 Diversity Array Technology markers.
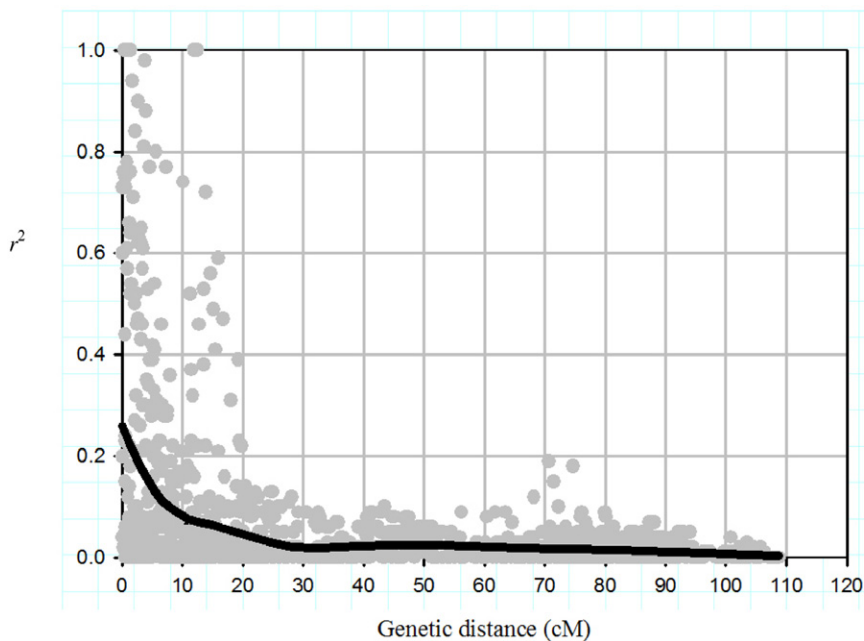

Figure 6. Linkage disequilibrium decay curve for chromosome 6A based on 53 Diversity Array Technology markers.

as population bottlenecks, 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 MAF >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 roughly in agreement with a previous report for wheat (Ravel et al., 2007). However, it is far less than SNP density reported for other crops such as maize (1 SNP per 104 bp), sorghum [*Sorghum bicolor* (L.) Moench] (1 SNP per 123 bp), sunflower (*Helianthus annuus* L.) (1 SNP per 69 bp), and rice (1 SNP per 113 bp and 1 SNP per 100 bp) (Fusari et al., 2008). Although the SNP density of *DREB1A* is high compared to the remaining drought tolerance candidate genes in this study, the value of one SNP per 61 bp is within the range of the SNP densities reported for four transcription factors involved in barley (*Hordeum vulgare* L.) endosperm development, which ranged from one SNP

per 31 bp to one SNP per 74 bp (Haseneyer et al., 2010). 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 nonrandom association of polymorphic sites (SNPs) within a gene. Graphical displays of LD (Fig. 1, 2, 3, and 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 chromosomewide 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 genomewide 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 genomewide QTL scanning is the possibility of overlooking 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 under both 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 × 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).

Several SNPs within drought tolerance candidate genes showed associations with yield and yield components and morphological and phenological traits. 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.

Table 6. Marker–trait associations for single nucleotide polymorphisms (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	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	GRW111	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	0.044	7.14
		Leaf senescence	GRW10	0.029	6.6
		Spike length	GRW11	0.07024	5.2
	ERA1B_932	Spikes per square meter	Combined	0.0618	6.1
		Spikes per square meter	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	0.043	5.8
		Flag leaf area	GRW10	0.055	5.4
	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	0.0546	4.5
		NDVI	Combined	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.

[‡]pFDR, false discovery rate adjusted probability value.

[§]Percent phenotypic variance explained by the SNP.

[¶]NDVI, normalized difference vegetation index.

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 associations of SNPs with phenotypic traits. 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.

Supplemental Information Available

Supplemental material is available at http://www.crops.org/publications/tpg.

References

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

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

Andersen, J., and T. Lubberstedt. 2003. Functional markers in plants. Trends Plant Sci. 18:554–560. doi:10.1016/j.tplants.2003.09.010

Bagge, M., and T. Lubberstedt. 2008. Functional markers in wheat: Technical and economic aspects. Mol. Breed. 22:319–328. doi:10.1007/s11032-008-9190-6

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R. Stat. Soc., B 57:289–300.

Berloo, V.R. 2008. GGT2.0: Versatile software for visualization and analysis of genetic data. J. Hered. 99(2):232–236.

Bernardo, R. 1993. Estimation of coefficient of co-ancestry using molecular markers in maize. Theor. Appl. Genet. 85:1055–1062. doi:10.1007/BF00215047

Bhatnagar-Mathur, P., J.M. Devi, M. Lavanya, D.S. Reddy, V. Vadez, R. Serraj, K. Yamaguchi-Shinozaki, and K.K. Sharma. 2007. Stress-inducible expression of At DREB1A in transgenic peanut (*Arachis hypogaea* L.) increases transpiration efficiency under water limiting conditions. Plant Cell Rep. 26:2071–2082. doi:10.1007/s00299-007-0406-8

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

Casadesus, J., Y. Kaya, J. Bort, M.M. Nachit, J.L. Araus, S. Amor, G. Ferrazzano, F. Maalouf, M. Maccaferri, V. Martos, H. Ouabbou,

and D. Villegas. 2007. Using vegetation indices derived from conventional digital cameras as selection criteria for wheat breeding in water-limited environments. Ann. Appl. Biol. 150:227–236. doi:10.1111/j.1744-7348.2007.00116.x

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

Chao, S.M., J. Dubcovsky, J. Dvorak, M.C. Luo, P.S. Baenziger, R. Matnyazov, D.R. Clark, L.E. Talbert, J.A. Anderson, S. Dreisigacker, K. Glover, J.L. Chen, K. Campbell, P.L. Bruckner, J.C. Rudd, S. Haley, B.F. Carver, S. Perry, M.E. Sorrells, and E.D. Akhunov. 2010. Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (*Triticum aestivum* L.). BMC Genomics 11:727. doi:10.1186/1471-2164-11-727

Cutler, S., M. Ghassemian, D. Bonetta, S. Cooney, and P. McCourt. 1996. A protein farnesyl transferase involved in abscisic acid signal transduction in Arabidopsis. Science 273:1239–1241. doi:10.1126/science.273.5279.1239

DNASTAR. 2008. Seqman genome assembler. DNASTAR Inc., Madison, WI.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol. Ecol. 14:2611–2620. doi:10.1111/j.1365-294X.2005.02553.x

Figueiredo, A.F., B. Sine, J. Chantereau, C. Mestres, G. Fliedel, J.F. Rami, J.C. Glaszmann, M. Deu, and B. Courtois. 2010. Variability of grain quality in sorghum: Association with polymorphism in Sh2, Bt2, SssI, Ae1, Wx and O2. Theor. Appl. Genet. 121:1171–1185. doi:10.1007/s00122-010-1380-z

Flint-Garcia, S.A., A.C. Thuillet, J. Yu, G. Pressoir, S.M. Romero, S.E. Mitchell, J. Doebley, S. Kresovich, M.M. Goodman, and E.S. Buckler. 2005. Maize association population: A high-resolution platform for quantitative trait locus dissection. Plant J. 44:1054–1064. doi:10.1111/j.1365-313X.2005.02591.x

Fusari, C., V. Lia, E. Hopp, R. Heinz, and N. Paniego. 2008. Identification of single nucleotide polymorphisms and analysis of linkage disequilibrium in sunflower elite inbred lines using the candidate gene approach. BMC Plant Biol. 8:7. doi:10.1186/1471-2229-8-7

Giordani, T., M. Buti, L. Natali, C. Pugliesi, F. Cattonaro, M. Morgante, and A. Cavallini. 2011. An analysis of sequence variability in eight genes putatively involved in drought response in sunflower (*Helianthus annuus* L.). Theor. Appl. Genet. 122:1039–1049. doi:10.1007/s00122-010-1509-0

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

Haudry, A., A. Cenci, C. Ravel, T. Bataillon, D. Brunel, C. Poncet, I. Hochu, S. Poirier, S. Santoni, S. Glémin, and J. David. 2007. Grinding up wheat: A massive loss of nucleotide diversity since domestication. Mol. Biol. Evol. 24:1506–1517. doi:10.1093/molbev/msm077

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

Kobayashi, F., M. Ishibashi, and S. Takumi. 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

Langridge, P., and D. Fleury. 2011. Making the most of 'omics' for crop breeding. Trends Biotechnol. 29:33–40. doi:10.1016/j.tibtech.2010.09.006

Latini, A., C. Rasi, M. Sperandei, C. Cantale, M. Lannetta, M. Dettori, K. Ammar, and P. Galeffi. 2007. Identification of a DREB-related gene in *Triticum durum* and its expression under stress conditions. Ann. Appl. Biol. 150:187–195. doi:10.1111/j.1744-7348.2007.00128.x

Lopes, M.S., and M.P. Reynolds. 2012. Stay-green in spring wheat can be determined by spectral reflectance measurements (normalized difference vegetation index) independently from phenology. J. Exp. Bot. 63:3789–3798. doi:10.1093/Jxb/Ers071

Lopes, M., M.P. Reynolds, M. Jalal-Kamali, M. Mousa, Y. Feltaous, I. Tahir, N. Barma, and M. Vargas. 2012. The yield correlations of selectable physiological traits in a population of advanced spring wheat lines grown in warm and drought environments. Field Crops Res. 128:129–136. doi:10.1016/j.fcr.2011.12.017

Lothier, J., B. Lasseur, K. Le Roy, A. Van Laere, M. Prud'homme, P. Barre, W. Van Den Ende, and A. Morvan-Bertrand. 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. J. Exp. Bot. 58:1969–1983. doi:10.1093/jxb/erm053

Manmathan, H., D. Shaner, J. Snelling, N. Tisserat, and N. Lapitan. 2013. Virus-induced gene silencing of *Arabidopsis thaliana* gene homologues in wheat identifies genes conferring improved drought tolerance. J. Exp. Bot. 64:1381–1392. doi:10.1093/jxb/ert003

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

McCouch, S., K. Zhao, M. Wright, C. Tung, K. Ebana, M. Thomson, A. Reynolds, D. Wang, G. DeClerck, M. Ali, A. McClung, G. Eizenga, and C. Bustamante. 2010. Development of genome-wide SNP assays for rice. Breed. Sci. 60:524–535. doi:10.1270/jsbbs.60.524

Mochida, K., Y. Yamazaki, and Y. Ogihara. 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. doi:10.1007/s00438-003-0939-7

Nei, M. 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York, NY.

Pei, Z.M., M. Ghassemian, C.M. Kwak, P. McCourt, and J.I. Schroeder. 1998. Role of farnesyltransferase in ABA regulation of guard cell anion channels and plant water loss. Science 282:287–290. doi:10.1126/science.282.5387.287

Pellegrineschi, A., M. Reynolds, M. Pacheco, R.M. Brito, R. Ameraya, K.Y. Shinozaki, and D. Hoisington. 2004. Stress-induced expression in wheat of the *A. thaliana DREB1A* gene delays water stress symptoms under greenhouse conditions. Genomics 47:493–500.

Pritchard, J., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Rafalski, A. 2002. Novel genetic mapping tools in plants: SNPs and LD-based approaches. Plant Sci. 162:329–333. doi:10.1016/S0168-9452(01)00587-8

Ravel, C., S. Praud, A. Canaguier, P. Dufour, S. Giancola, F. Balfourier, B. Chalhoub, D. Brunel, L. Linossier, M. Dardevet, M. Beckert, M. Rousset, A. Murigneux, and G. Charmet. 2007. DNA sequence polymorphisms and their application to bread wheat quality. Euphytica 158:331–336. doi:10.1007/s10681-006-9288-z

Reif, J.C., P. Zhang, S. Dreisigacker, M.L. Warburton, M. van Ginkel, D. Hoisington, M. Bohn, and A.E. Melchinger. 2005. Wheat genetic diversity trends during domestication and breeding. Theor. Appl. Genet. 110:859–864. doi:10.1007/s00122-004-1881-8

Rozas, J., P. Librado, J.C. Sánchez-DelBarrio, X. Messeguer, and R. Rozas. 1999. DnaSP version 5: An integrated program for population genetics and molecular evolution analysis. Bioinformatics 15:174–175. doi:10.1093/bioinformatics/15.2.174

Rozen, S., and H. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. In: S. Krawetz and S. Misener, editors, Bioinformatics methods and protocols. Humana Press, Totowa, NJ. p. 365–386.

Saint Pierre, C., J. Crossa, D. Bonnet, K. Yamaguchi-Shinozaki, and M. Reynolds. 2012. Phenotyping transgenic wheat for drought resistance. J. Exp. Bot. 63:1799–1808. doi:10.1093/jxb/err385

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.

Shinozaki, K., and K. Yamaguchi-Shinozaki. 2007. Gene networks involved in drought stress response and tolerance. J. Exp. Bot. 58:221–227. doi:10.1093/jxb/erl164

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

Vadez, V., J.S. Rao, P. Bhatnagar-Mathur, and K.K. Sharma. 2013. DREB1A promotes root development in deep soil layers and increases water extraction under water stress in groundnut. Plant Biol. 15:45–52. doi:10.1111/j.1438-8677.2012.00588.x

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

VSN International. 2005. CycDesigN 3.0. VSN International Ltd., Hemel Hempstead, UK.

Wei, B., R. Jing, C. Wang, J. Chen, X. Mao, X. Chang, and J. Jia. 2009. Dreb1 genes in wheat (*Triticum aestivum* L.): Development of functional markers and gene mapping based on SNPs. Mol. Breed. 23:13–22. doi:10.1007/s11032-008-9209-z

Yang, D., R. Jing, X. Chang, and W. Li. 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. doi:10.1534/genetics.106.068361

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

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

Ziegelhoffer, E., J. Leonard, and M. Elliot. 2000. Cloning of the Arabidopsis WIGGUM gene identifies a role for farnesylation in meristem development. Proc. Natl. Acad. Sci. USA 97:7633–7638. doi:10.1073/pnas.130189397