Theor Appl Genet (2015) 128:453–464
DOI 10.1007/s00122-014-2444-2

ORIGINAL PAPER

Genetic characterization of the wheat association mapping initiative (WAMI) panel for dissection of complex traits in spring wheat

M. S. Lopes · S. Dreisigacker · R. J. Peña · S. Sukumaran · M. P. Reynolds

Received: 14 February 2014 / Accepted: 11 December 2014 / Published online: 25 December 2014

Abstract
Key message **The wheat association mapping initiative is appropriate for gene discovery without the confounding effects of phenology and plant height.**
Abstract The wheat association mapping initiative (WAMI) population is a set of 287 diverse advanced wheat lines with a narrow range of variation for days to heading (DH) and plant height (PH). This study aimed to characterize the WAMI and showed that this diverse panel has a favorable genetic background in which stress adaptive traits and their alleles contributing to final yield can be identified with reduced confounding major gene effects through genome-wide association studies (GWAS). Using single nucleotide polymorphism (SNP) markers, we observed lower gene diversity on the D genome, compared with the other genomes. Population structure was primarily related to the distribution of the 1B.1R rye translocation. The narrow range of variation for DH and PH in the WAMI population still entailed segregation for a few markers associated with the former traits, while *Rht* genes were associated with grain yield (GY). Genotype by environment (G × E) interaction for GY was primarily explained by *Rht-B1*, *Vrn-A1* and markers on chromosomes 2D and 3A when running GWAS with genotype scores from the G × E biplot. The use of PC scores from the G × E biplot seems a promising tool to determine genes and markers associated with complex interactions across environments. The WAMI panel lends itself to GWAS for complex trait dissection by avoiding the confounding effects of DH and PH which were reduced to a minimum (using *Rht-B1* and *Vrn-A1* scores as covariables), with significant associations with GY on chromosomes 2D, 3A and 3B.

Communicated by Peter Langridge.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-014-2444-2) contains supplementary material, which is available to authorized users.

M. S. Lopes
CIMMYT, PO Box 39, Emek, Ankara 06511, Turkey

S. Dreisigacker · R. J. Peña · S. Sukumaran · M. P. Reynolds (✉)
CIMMYT, Int. Apdo Postal 6-641, 06600 Mexico, DF, Mexico
e-mail: m.reynolds@cgiar.org

Introduction

Wheat productivity gains have been decreasing during the last 30 years (Olesen et al. 2011), a trend which has been linked to changing temperature and precipitation patterns (Lobell et al. 2011). Nonetheless, genetic gains in yield related to breeding of spring wheat are still being achieved (Lopes et al. 2012; Sharma et al. 2012) though at lower rates than in the past. To breed cultivars better adapted to a changing climate, positive alleles for improved performance under those changing conditions must be identified in current elite germplasm, as well as in new genetic pools.

Risch and Merikangas (1996) proposed that association mapping or genome-wide association studies (GWAS) may be a suitable complementary approach to bi-parental populations for identifying genetic markers associated with traits of interest. High-throughput genotyping technologies and statistical methods based on linkage disequilibrium (LD) that take population structure into account (Pritchard et al. 2000; Kang et al. 2008) enabled GWAS to be developed for analysis of complex traits. In wheat, association studies of yield, disease resistance, and bread making quality have already proved useful in studying complex traits

(Crossa et al. 2007; Charmet et al. 2009; Rehman Arif et al. 2012).

It has been difficult to identify consistent genetic markers associated with yield under stress environments (particularly drought) in either wheat or maize (Maccaferri et al. 2011; Dodig et al. 2012; Lu et al. 2012). In wheat especially, the identification of markers has been complicated by confounding effects of phenology; most stress adaptive traits interact with days to heading (DH) and plant height (PH), thus genes of major effect can mask the identification of minor effect genes (Reynolds et al. 2009), e.g., expression of cooler canopies varies with DH and PH with co-locations and pleiotropy being observed most of the times (Rebetzke et al. 2012). The discovery of markers associated with adaptive traits can only be accurately assessed if experimental populations have controlled phenology and PH, and are grown in reproducible field trials (Pinto et al. 2010; Lopes et al. 2013); however, this approach will not integrate genotype, or marker by environment interactions (G × E).

This study aimed to characterize the wheat association mapping initiative (WAMI) population, consisting of 287 advanced lines released by the International Maize and Wheat Improvement Center (CIMMYT) over the past 30 years. The 9K iSelect beadship assay (Cavanagh et al. 2013) was utilized and compared in terms of accuracy of marker-trait associations. Specifically, this study aimed to: (i) characterize population structure and linkage disequilibrium patterns in the WAMI; (ii) localize functional genes using marker scores for *Rht-B1*, *Rht-D1*, *Vrn-A1*, *Vrn-B1*, *Vrn-D1* in marker–marker associations to determine pleiotropy or co-location of grain yield (GY) with phenology and PH genes and their possible confounding effect; (iii) characterize minor marker-trait associations in the WAMI for DH, PH, and interactions with GY (if present) after removal of major *Vrn* and *Rht* effects; (iv) provide the wheat research community with a genetically well-characterized population of advanced elite lines, with a particular emphasis on the control of interactive effects from phenology and PH.

Materials and methods

Plant material, DNA extraction, and genotyping

The WAMI population was obtained from CIMMYT international nurseries that are annually distributed worldwide (26th, 27th, and 28th Elite Spring Wheat Yield Trial; 1st to 16th Semiarid Wheat Yield Trial; and 1st to 12th High Temperature Wheat Yield Trial). The population was carefully selected to obtain a reduced range of variation in phenology and height (9 days and 35 cm variation for DH and PH, respectively, averaged across 12 environments; Lopes et al. 2012). Seeds of all lines were obtained from the CIMMYT Genetic Resources Center and genomic DNA was extracted from five leaves using a CTAB procedure (Saghai-Maroof et al. 1984), modified according to CIMMYT laboratory protocols (http://repository.cimmyt.org/xmlui/handle/10883/3221).

Whole-genome marker polymorphisms were generated via genotyping with the 9K iSelect beadship assay, including gene-associated single nucleotide polymorphisms (SNPs), as described in Cavanagh et al. (2013). Markers with a minor allele frequency (MAF) <0.05 were excluded to avoid the high statistical power required to make meaningful statements about association with very rare alleles. A total number of 3,848 markers were located on the consensus map available in Cavanagh et al. (2013) and these were used for association analysis.

Ppd-1, *Vrn-1,* and *Rht-1* markers, and the 1B.1R rye translocation

We used primers described in Beales et al. (2007), Wilhelm et al. (2009), and Diaz et al. (2012) to detect alleles of the photoperiod loci *Ppd-A1*, *Ppd-B1*, and *Ppd-D1*, which confer sensitivity or insensitivity to day length, on chromosomes 2A, 2B, and 2D, respectively. PCR assays were applied to detect the 1,027 bp 'GS-100' type and 1,117 bp 'GS-105'-type deletions in the *Ppd-A1* gene; the junction between intact *Ppd-B1* copies in the 'Sonora64' allele; and the 2,089 bp deletion in the *Ppd-D1* gene. The two deletions in the *Ppd-A1* gene were identified exclusively in durum wheat, but are routinely transferred to CIMMYT bread wheat via the use of synthetic wheat. The photoperiod-insensitive allele was labeled as *Ppd-1A* in accordance with Beales et al. (2007). The alternative allele, which we have assumed to infer photoperiod sensitivity, was designated *Ppd-1B*.

Dominant spring alleles related to variation in the promoter region of the *Vrn-A1* locus were identified using the genome-specific primers VRN1AF and VRN1R, described by Yan et al. (2004). Deletion alleles affecting vernalization response in intron-1 of the *Vrn-B1* and *Vrn-D1* loci were detected as described in Fu et al. (2005).

Primers were amplified in single 10-µl PCRs, with final concentrations of: 1× Buffer with Green Dye (Promega Corp., US), 200 µM dNTPs, 1.2 mM $MgCl_2$, 0.25 µM of each primer, 1U DNA polymerase (Promega), and 50 ng DNA template. PCRs consisted of 94 °C for 2 min, followed by 30 cycles of 94 °C for 1 min then 54–60 °C for 2 min (dependent on the primer), then finally 72 °C for 2 min. Products were separated on 1.2 % agarose gels in TAE buffer.

Alleles of *Rht-B1* and *Rht-D1* (*a* or *b* for the dwarfing and wild-type allele, respectively) were classified using the

method described in Ellis et al. (2002). We scored polymorphisms using LGC Genomics KASP reagents (www.lgcgenomics.com) in reactions containing 2.5 µl water, 2.5 µl 2× KASPar reaction mix, 0.07 µl assay mix, and 50 ng dried DNA, with a PCR profile of 94 °C for 15 min (activation), followed by 20 cycles of 94 °C for 10 s, 57 °C for 5 s, 72 °C for 10 s, then 18 cycles of 94 °C for 10 s, 57 °C for 20 s, and 72 °C for 40 s. Fluorescence was measured as an end point reading at 25 °C. Primer combinations are described at http://www.cerealsdb.uk.net/CerealsDB/SNPs/Documents/MAS.php?URL.

The 1B.1R translocation was determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) of alcohol-soluble (gliadins) and alcohol-insoluble (glutenins) grain protein extracts, using the procedure described by Liu et al. (2005). The presence of *Sec-1* secalins in the gliadin gel and the *Glu-B3*j allele in the glutenin gel confirmed the presence of the 1B.1R translocation.

Gene diversity, polymorphism information content (PIC), population structure and linkage disequilibrium (LD)

Gene diversity, defined as the probability that two randomly chosen alleles from the population are different (Weir 1996), and polymorphism information content (PIC) values, defined by Botstein et al. (1980), were calculated using PowerMarker software V 3.25 (Liu and Muse 2005). LD was determined by calculating r^2 using allele frequencies of the markers with Graphical GenoTypes GGT2 (Berloo 2008) 2.0 software. The critical r^2 value, after which LD is due to true physical linkage, was calculated through the 95th percentile of normalized (square root) r^2 values. Population structure was analyzed applying principal component analysis (PCA) using markers at least 10 cM apart. The number of markers was 151, 151, and 45, for A, B, and D genomes, respectively. PCA was implemented in R (Everitt 2007). A neighbor joining (NJ) tree was determined with PowerMarker V3.25 (Liu and Muse 2005) to further support the PCA results.

Field trials, traits measured and genotype by environment interaction

Four field trials were conducted in the Norman E. Borlaug (CENEB) experimental station near Ciudad Obregon, Sonora, Mexico during the 2009–2010 season. Experiments were conducted at recommended sowing time (late November) under the following conditions: near optimal conditions, i.e., irrigated (IRRIGATED); drought (DROUGHT) with reduced irrigation; to test performance under heat stress, two experiments were sown late (February), one with irrigation (HEAT) and another one with reduced irrigation (HD) (for meteorological data see Lopes et al. 2012). Appropriate fertilization and weed, disease, and pest control were implemented to avoid yield limitations. Experimental design was a randomized lattice with two replications in 2-m-long and 0.8-m-wide plots consisting of one raised bed with 2 rows per bed at seed rates of 120 kg ha^{-1}. Best linear unbiased estimates (BLUPs) for all measured traits were used and procedures to determine GY, DH and PH have been described in Lopes et al. (2012). Prior analysis, averages across environments were calculated using normalized values (against the mean trial) to avoid giving more weight to high-yielding environments (where GY, DH and PH were all higher irrigated environments than in environments with either drought or heat). To obtain information on the (G × E) interactions, site regression analysis was used. Site regression analysis was performed according to Samonte et al. (2005) using normalized yield values in SAS (2004) and scores of genotypes and environments plotted in a biplot. Broad sense heritability (H^2) was estimated for each trait individually in each environment and across environments as shown in Lopes et al. (2012). Phenotypic correlations among traits in different environments were calculated with the PROC CORR procedure in SAS (2004).

Genome-wide association analysis (GWAS)

To determine the best model for GWAS allowing for false-positive control due to genetic relatedness a forward model selection using the Bayesian information criterion (BIC) was determined with the Genomic Association and Prediction Integrated Tool -GAPIT (Lipka et al. 2012 and http://www.maizegenetics.net/GAPIT) applied in R. Besides basic population structure control with kinship matrix and PCA scores, two sets of models were applied: (1) no correction was applied in the GWAS besides the kinship and PCA scores from population structure; (2) scores for *Rht* and *Vrn* functional genes included as covariables for correction of their effects in GWAS plus kinship and PCA scores; Furthermore, GWAS was conducted [using approaches (1) and (2) as explained above] on the following type of phenotypic data for each trait: (1) estimated means in each environment; (2) normalized means across all environments; (3) two first PC scores from site regression G × E biplot; (4) normalized average across stress environments (DROUGHT, HEAT and HD). The following mixed model was used to account for genetic relatedness among advanced lines and varieties in the genome-wide association mapping (Kang et al. 2008):

$$y = \mu + x\beta + u + e$$

where μ is the mean; x is the SNP marker effect; u is random effects due to genetic relatedness with Var $(u) = \sigma_g^2 K$ and Var $(e) = \sigma_e^2$; and K is the kinship matrix across all genotypes.

Table 1 Gene diversity, polymorphism information content (PIC), number of markers (no. Markers), chromosome size in cM (Chrom size), marker density (SNP Dens), and average linkage disequilibrium (AVG LD) for all chromosomes, genomes, and overall means; calculated using SNP markers in the wheat association mapping population (WAMI) of 287 advanced lines

Chr	Gene diversity	PIC	No. markers	Chrom size (cM)	SNPDens (no. locus cM^{-1})	AVG LD (r^2)
1A	0.29	0.23	321.00	278.10	1.15	0.10
1B	0.35	0.28	224.00	206.50	1.08	0.39
1D	0.34	0.28	63.00	153.80	0.41	0.21
2A	0.35	0.28	209.00	370.50	0.56	0.11
2B	0.38	0.30	419.00	395.40	1.06	0.16
2D	0.36	0.29	47.00	159.90	0.29	0.09
3A	0.36	0.29	240.00	279.30	0.86	0.11
3B	0.29	0.24	257.00	302.70	0.85	0.08
3D	0.31	0.25	10.00	10.45	0.39	0.26
4A	0.41	0.32	215.00	234.60	0.92	0.27
4B	0.32	0.26	92.00	157.50	0.58	0.09
4D	0.29	0.24	14.00	50.90	0.28	0.13
5A	0.38	0.30	291.00	240.80	1.21	0.05
5B	0.36	0.29	360.00	298.00	1.21	0.07
5D	0.35	0.28	10.67	31.23	0.24	0.24
6A	0.36	0.28	282.00	256.60	1.10	0.08
6B	0.41	0.32	277.00	243.30	1.14	0.15
6D	0.40	0.31	14.00	51.10	0.84	0.24
7A	0.30	0.24	256.00	281.20	0.91	0.08
7B	0.34	0.27	186.00	268.50	0.69	0.07
7D	0.35	0.28	7.67	17.97	0.84	0.35
Genome A	0.35	0.28	1,814	NA	0.96	0.11
Genome B	0.35	0.28	1,815	NA	0.95	0.14
Genome D	0.32	0.26	230	NA	0.43	0.25
Mean	0.34	0.27	NA	NA	0.78	0.17

A restricted maximum likelihood (REML) estimate of σ_g^2 and σ_e^2 were computed using EMMA and the association mapping conducted based on the estimated variance components, with a standard F test to test $\beta \neq 0$. In genome-wide association mapping studies with n molecular markers, the same statistical test is performed n times at the significance level α. The threshold for defining a marker to be significant was taken at $p \leq 0.0005$ considering the small number of markers and the deviation of the observed quantile–quantile (Q–Q) plots. The Q–Q plots were constructed by plotting the observed $-\log(p$ value) of the markers against the expected $-\log10(p$ value), under the null hypothesis that there is no association between marker and phenotype.

To verify if marker-trait associations were co-localized to any of the used functional markers or closely linked SNPs associated with DH (*Vrn* and *Ppd*) and PH (*Rht*), marker–marker associations were determined using the scores of the functional markers in GWAS. This was also useful to localize the functional markers within chromosomes and to ensure that the WAMI, markers, map, and methods were adequate for GWAS.

Results

Polymorphism, gene diversity, marker density, and LD

LD was higher and decay was faster in the D genome, compared to the A and B genomes (Table 1; Fig. 1). For the ca. 9K SNP markers, 4,059 displayed polymorphic SNP loci, of which 3,848 were located on the consensus map available in Cavanagh et al. (2013). The SNP markers showed different levels of gene diversity; values ranged from 0.26 to 0.41, with an average of 0.34 (Table 1). PIC generated with SNP markers ranged from 0.22 to 0.32 with a mean value of 0.27 (Table 1). Decreased gene diversity and PIC were depicted in genome D when compared to genomes A and B (Table 1).

Population structure

Population structure of the WAMI was examined with PCA, using markers more than 10 cM apart (Fig. 2). Across genomes PCA analysis separated the WAMI lines in two

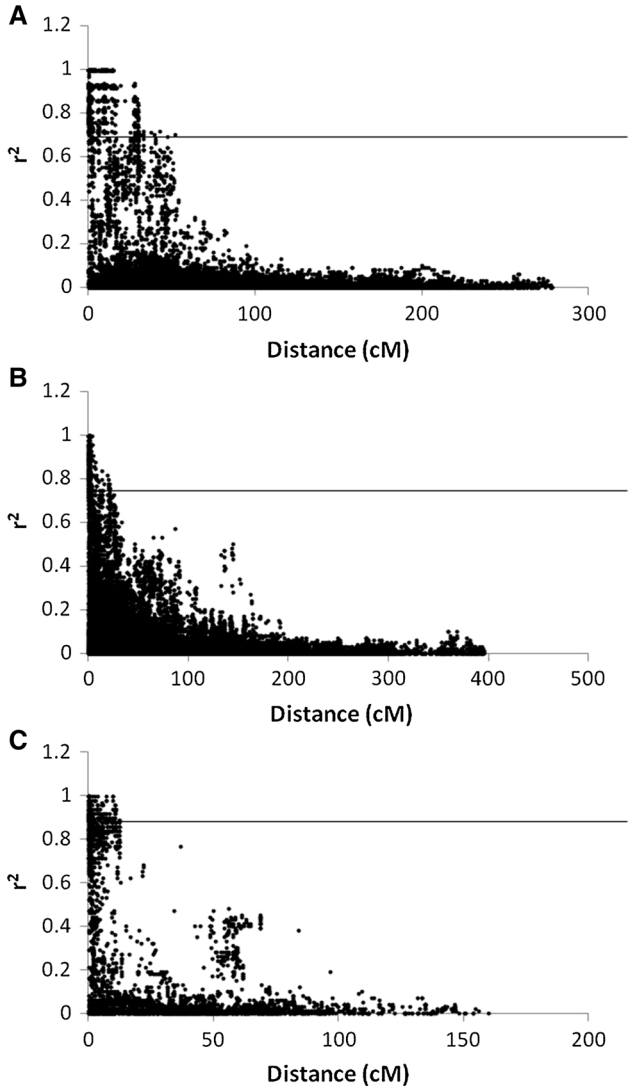

Fig. 1 Linkage disequilibrium (LD) using the squared allele-frequency correlation (r^2) against marker distance for genomes A, B and D (**a**, **b** and **c**, respectively). The 95th percentile of r^2 for each genome is represented by a *dark horizontal line*

groups, with and without the 1B.1R translocation (Fig. 2a). Using SNP markers localized in the A genome, 'Pastor' and 'Babax' derivative elite lines were separated (Fig. 2b). SNP markers from B genome grouped according to the 1B.1R translocation (Fig. 2c) and markers localized in D genome mainly separated 'Kauz' derivative advanced lines (Fig. 2d).

Marker–marker associations of *Rht-B1*, *Rht-D1*, *Vrn-A1*, *Vrn-B1*, and *Vrn-D1* functional genes

Nearly all lines (except five) carried the photoperiod-insensitive *Ppd-D1a* allele. Similarly, the frequency of the 1,027 bp 'GS-100' type and 1,117 bp 'GS-105'-type deletions in the *Ppd-A1* gene was low and did not fall under the MAF >0.05 threshold. Both alleles (*Ppd-D1b and Ppd-A1a*) were, therefore, excluded from further analyses and only *Vrn* and *Rht* genes were considered. The marker–marker associations for *Rht-B1* and *Rht-D1* localized the genes on chromosomes 4B and 4D at 76.58 and 0.50 cM, respectively (Supplementary Table 1). *Vrn-A1* was localized on chromosome 5A at 82.82 and 83.37 cM, whereas *Vrn-B1* was localized on chromosome 5B at 96.68 and 97.98 cM (Supplementary Table 1). *Vrn-D1* was localized on chromosome 5D at 16.86 cM and on chromosome 5A at 83.37 cM (Supplementary Table 1).

Genotype by environment (G × E) interaction for GY, DH and PH: correlations between traits and heritability

Site regression analysis is shown in Fig. 3. For GY, all stressed environments including HEAT, HD, DROUGHT were clustered together, whereas the near optimal environment IRRIGATED was separated from all other environments and causing the main G × E (Fig. 3). For DH and PH G × E was also significant. DH showed the same trend (as observed for GY) with all stressed environments having similar patterns and only IRRIGATED was separated from all other environments (though similar to DROUGHT). PH measured in HD and HEAT environments, gave very similar responses whereas IRRIGATED and DROUGHT were separated by PC2 (Fig. 3). Heritability of all traits per environment and across environments is shown in Table 2. Phenotypic correlations between GY and DH were all significantly negative, but stronger under DROUGHT, HEAT and HD conditions (Table 2). Correlations between GY and PH were all significantly positive except under IRRIGATED conditions where correlations were negative (Table 2).

Genetic architecture of GY, DH and PH determined by GWAS in different environments

The type of model used for GWAS for each trait and environment was determined by forward model selection using the BIC to determine the optimal number of parameters (kinship, PCs from population structure analysis and *Rht-B1* and *Vrn-A1* as covariables) that best control for population structure, and results are shown in Supplementary Table 2 (with and without correction of *Rht-B1* and *Vrn-A1* covariables). GWAS results are shown in detail in Supplementary Tables 3, 4, 5, for GY, DH and PH, respectively. GWAS was conducted for each trait or phenotype, using estimated means in each environment, normalized means across environments and PC scores from site regression biplot (G × E). Q–Q plots are shown in Supplementary Fig. 1 for the traits and environments where significant associations were observed.

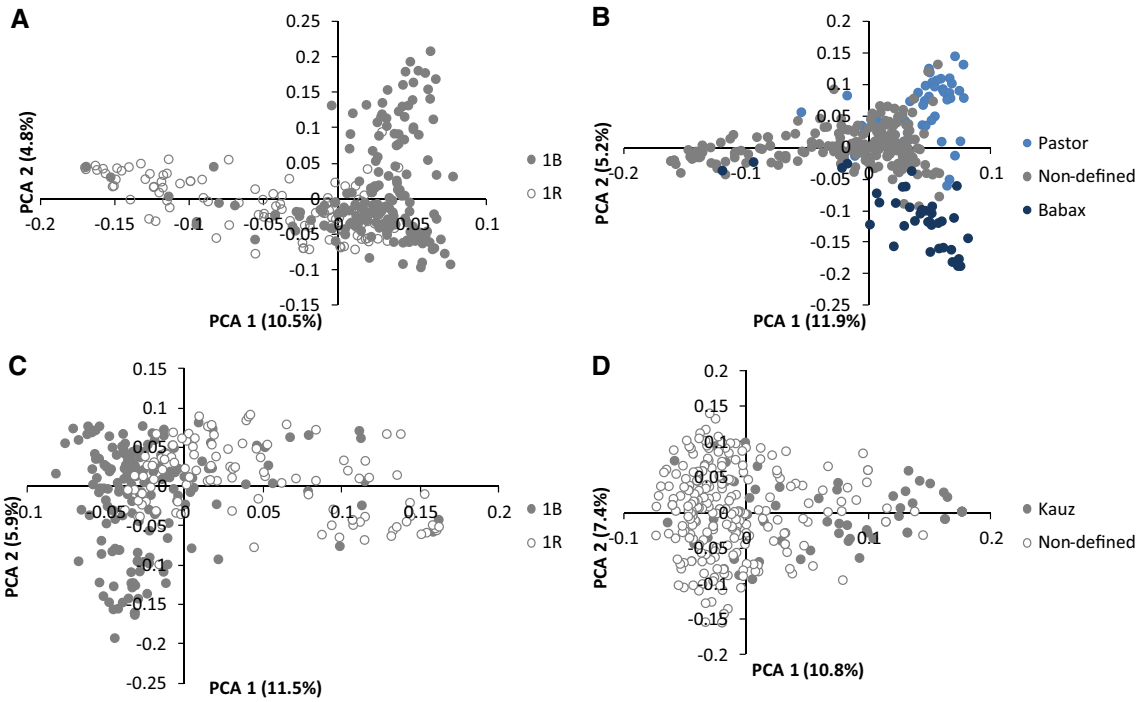

Fig. 2 Principal component analysis of the WAMI based on SNP markers (using markers >10 cM apart). **a** Including markers from all genomes; **b** Including markers present in genome **a** only; **c** Including markers present in genome **b** only; **d** Including markers present in genome **d** only. Non-defined, no clear common pedigree or cluster pattern

Approach 1-GWAS without *Vrn* and *Rht* marker scores as covariables

Markers associated with GY across all environments (normalized average) were detected on chromosomes 2D at 94.63 cM, 3A at 30.82 cM, 3B at 278.7 cM (with lower allelic effects) and *Rht-B1* on chromosome 4B (Table 3 and Supplementary Table 3). The same SNP marker on chromosome 3B at 278.7 cM was associated with GY averaged across stress environments (DROUGHT, HEAT and HD, see Table 3 and Supplementary Table 3). SNP markers on chromosomes 2D at 94.63 cM, 3A at 30.82 cM and *Rht-B1* were also associated with PC1 scores from the G × E biplot, whereas a few SNP markers on chromosome 5A (near *Vrn-A1*, see Supplementary Table 1 for *Vrn* localization) were associated with PC2 (Table 3 and Supplementary Table 3).

For DH, the most significant associations across all environments were observed on chromosome 5A and the *Vrn-A1* functional gene between 74.68 and 86.37 cM (Table 3 and Supplementary Table 4).

Major effects for PH were associated with *Rht-B1* and *Rht-D1* and several SNP markers on chromosome 6A in 'IRRIGATED' and 'DROUGHT' environments (Table 3 and Supplementary Table 5).

Approach 2-GWAS using *Vrn* and *Rht* as covariables

With this approach it was possible to remove associations of *Vrn-A1* or *Rht-B1* (Table 4 and Supplementary Table 3) with GY. All SNPs associated with GY with the first approach on chromosomes 2D and 3A (Table 3) were also detected in GWAS with *Vrn-A1* and *Rht-B1* as covariables, however, allelic effects decreased with the correction (comparing allelic effects of markers in Table 3 with Table 4). Using Approach 2, DH was not associated with any marker under any of the conditions tested (Supplementary Table 4) and PH was associated with *Rht-D1* under IRRIGATED conditions (Supplementary Table 5).

The most significant markers identified by GWAS on chromosomes 2D, 3A and 3B associated with GY were summarized for each line in the WAMI to determine their direct effects on GY, PH and DH (Fig. 4). It was shown that when the favorable allele 'A' is present in the three SNPs, the highest GY was obtained with no changes in DH (Fig. 4). PH remained constant for almost all allele combinations but it was higher when the unfavorable allele 'B' was present in all SNPs (Fig. 4). Furthermore, Fig. 4b also shows that the impact of *Rht-B1* on yield was smaller than the effect caused by the new SNPs identified as associated with GY in the WAMI.

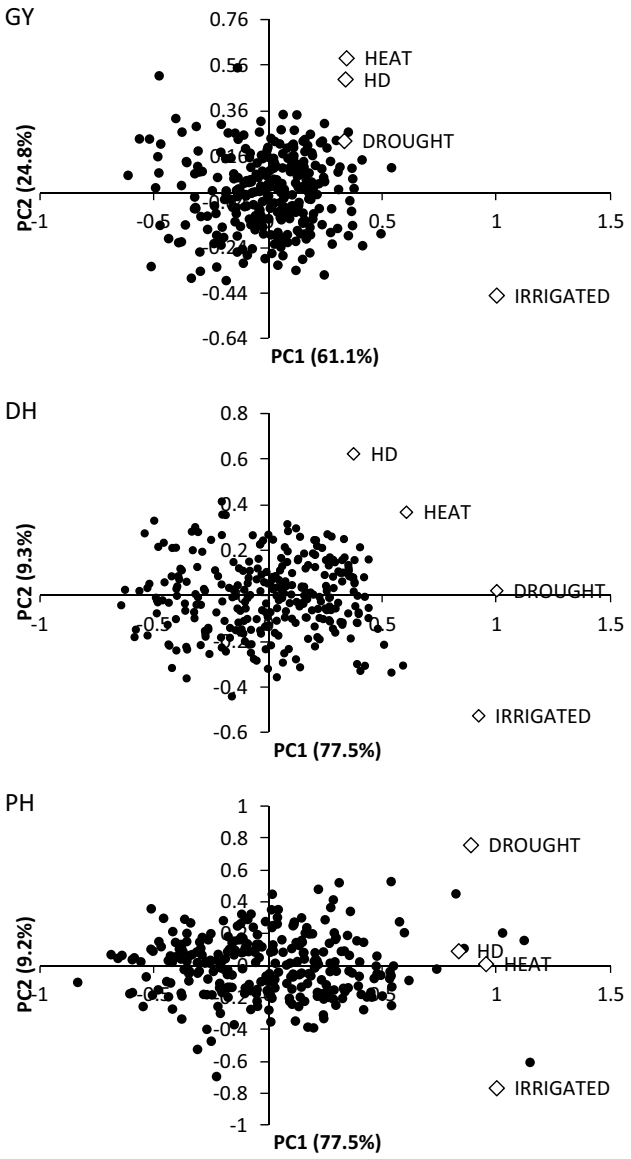

Fig. 3 Site regression analysis of the WAMI grown in four environments in Mexico (DROUGHT, HEAT, heat combined with drought (HD) and IRRIGATED) for three variables: grain yield (GY), days to heading (DH) and plant height (PH) as shown in the *left corner* of each panel. Genotypes are indicated with *dark circles* and environments with *open diamonds*. Percentages of variation for genotype by environment interaction are shown for the first two principal components (PC1 and PC2)

Discussion

Gene diversity, LD, and population structure

Gene diversity is an important factor for crop improvement and is, therefore, essential for the effective use of genetic resources in breeding programs (Chao et al. 2007) as long as this diversity integrates positive and profitable genes. Gene diversity varied across wheat genomes in the WAMI population; specifically, gene diversity was high and similar in genomes A and B, and lower in genome D (with both marker platforms). This lower gene diversity in the D genome was observed despite the fact that around 15 % of elite lines included in the WAMI population were derived synthetic wheats; the result corroborates previous findings observed in a Chinese winter wheat collection (Chen et al. 2012). It is worth noting that gene diversity in the D genome among synthetic-derived lines in the WAMI population was higher (0.332) when compared with non-synthetic advanced lines of the population (0.328). This further confirms that synthetic-derived lines are a valuable source for broadening the genetic base of elite wheat breeding germplasm (Zhang et al. 2005). The gene diversity and PIC values observed in the WAMI population were higher than other wheat elite germplasm collections (Chao et al. 2010), though similar trends were observed in both studies in that the D genome had lower PIC values compared to the other genomes.

LD declines rapidly with distance and knowledge of LD extent is important for determining the adequate marker density required for proper marker-trait association (for a review of LD, see Gaut and Long 2003). LD in the WAMI population was highest in the D genome (highest average r^2) when compared to the other genomes, a result that is in accordance with findings published by Chao et al. (2010). Higher LD in the wheat D genome has been linked to episodes of recent introgression and population bottlenecks accompanying the origin of hexaploid wheat (Chao et al. 2010). Nevertheless, CIMMYT materials have been reported to show the lowest rate of LD decay among wheat populations within 30 cM (Chao et al. 2010; Dreisigacker et al. 2012). In this study, a faster LD decay in the D genome (similar results were observed by Edae et al. 2014) indicates the need for higher marker density to locate genes in the D genome; the marker platforms used here are still poor in terms of D genome cover. For the A and B genomes, LD decay was slower and thus higher marker density may or may not improve localization.

Population structure or stratification is particularly important in genome-wide studies because it can cause spurious marker-trait associations (Pritchard et al. 2000; Patterson et al. 2006; Kang et al. 2008). In the present study, analysis of population structure has been used for a general overview and knowledge of biological background explaining the WAMI structure. Particularly we were interested to understand the broad relatedness across individuals and used PCA. In this study, the WAMI population was mostly stratified by the presence or absence of the 1B.1R translocation (this was confirmed by both marker platforms). Population stratification in the A genome was mainly associated with pedigree (separating 'Pastor'-, 'Kauz'-, and 'Babax'-derived lines) whereas B genome stratification

Table 2 Pearson correlation coefficients (r) and associated probabilities ($P<$) of grain yield (GY) with days to heading (DH and correlations with GY are indicated by GY_DH) and plant height (PH and correlations with GY are indicated by rGY_PH) in the WAMI grown under IRRIGATED, DROUGHT, HEAT and heat combined with drought (HD) conditions plus the average of all environments (AVG)

	IRRIGATED	DROUGHT	HEAT	HD	AVG
rGY_DH	−0.12	−0.29	−0.34	−0.22	−0.24
$P<$	0.05	0.0001	0.0001	0.001	0.0001
rGY_PH	−0.15	0.14	0.13834	0.19	0.17
$P<$	0.05	0.0001	0.0001	0.01	0.001
h^2 GY	0.74	0.71	0.76	0.73	0.67
h^2 DH	0.85	0.93	0.86	0.86	0.87
h^2 PH	0.7	0.77	0.81	0.6	0.7

Heritability (h^2) of GY, DH and PH are shown for each environment and calculated across environments (AVG)

was related with the 1B.1R rye translocation, as expected. The 1B.1R translocation has been previously identified as an important cause of population stratification, as shown by Zhang et al. (2011) in Chinese wheat germplasm.

Removal of major confounding effects of DH and PH in the WAMI for the genetic dissection of complex traits like GY

Complex traits like GY and most physiological traits are strongly influenced by the time of flowering and height of the crop in wheat (Lopes et al. 2014). These are two baseline traits that if well synchronized in the crop will maximize yield and have been basically fixed by breeders (Richards et al. 2010). However, many other potential physiological responses and traits can effectively contribute

Table 3 Main SNP markers associated with grain yield (GY), days to heading (DH) and plant height (PH)

TRAIT	ENV	SNP	Chr	Pos (cM)	*P* value	Maf	All Eff
GY	AVG	wsnp_Ex_c8303_14001708	2D	94.63	7.10E−05	0.070	−0.163
		wsnp_Ex_c55051_57706127	3A	30.82	3.06E−05	0.070	−0.183
		wsnp_CAP12_c2297_1121142	3B	278.7	9.00E−05	0.382	−0.105
		rhtB1	4B	76.58[a]	0.0005	0.214	0.097
	AVG STRESS	wsnp_CAP12_c2297_1121142	3B	278.7	5.28E−05	0.382	−0.027
		wsnp_Ex_c55777_58153636	5A	74.68	0.000167	0.225	0.033
	PC1	wsnp_Ex_c8303_14001708	2D	94.63	0.000128	0.070	−0.028
		wsnp_Ex_c55051_57706127	3A	30.82	5.51E−06	0.070	−0.036
	PC2	rhtB1	4B	76.58[a]	0.000128	0.214	0.017
		wsnp_Ex_c55777_58153636	5A	74.68	0.000119	0.225	0.019
		wsnp_Ex_c621_1230852	5A	78	0.000299	0.267	0.017
DH	AVG	wsnp_AJ612027A_Ta_2_1	5A	82.82	1.83E−05	0.186	−0.013
		vrnA1	5A	82.82–83.37[a]	6.55E−05	0.189	0.012
		VrnA1a	5A	82.82–83.37[a]	0.000171	0.221	−0.010
	PC1	NA		NA	NA	NA	NA
	PC2	NA		NA	NA	NA	NA
PH	DROUGHT[b]	RhtB1	4B	0	0.000297	0.214	−0.349
		wsnp_Ex_rep_c107564_91144523	4D	0.5	4.47E−05	0.200	0.816
	PC1	NA	NA	NA	NA	NA	NA
	PC2	NA	NA	NA	NA	NA	NA

For each marker, chromosome (Chr) position (Pos), *P* values, minor allele frequency (maf) and allelic effects (All Eff) are shown. GWAS was conducted in normalized averages across four environments (AVG), average of stress environments (AVG STRESS) and in each environment (ENV): IRRIGATED, DROUGHT, HEAT and heat combined with drought (HD). Principal components (PC1 and PC2) from G × E interaction biplot were used in GWAS to identify markers associated with G × E

NA non available

[a] See Supplementary Table 1 for localization and position of functional markers

[b] Only available under Drought

Table 4 Main SNPs associated with grain yield (GY)

ENV	SNP	Chr	Pos (cM)	*P* value	Maf	All Eff
AVG	wsnp_Ex_c8303_14001708	2D	94.63	0.000109	0.070	−0.155
	wsnp_Ex_c55051_57706127	3A	30.82	6.03E–05	0.070	−0.173
	wsnp_CAP12_c2297_1121142	3B	278.7	7.03E–05	0.382	−0.105
AVG STRESS	wsnp_CAP12_c2297_1121142	3B	278.7	2.23E–05	0.382	−0.028
PC1	wsnp_Ex_c55051_57706127	3A	30.82	1.73E–05	0.070	−0.033
PC2	NA	NA	NA	NA	NA	NA
DROUGHT	wsnp_Ex_c8303_14001708	2D	94.63	0.00039	0.070	−0.151
	wsnp_Ex_c790_1554988	5A	66.48	0.000301	0.379	−0.097
HD	NA	NA	NA	NA	NA	NA
HEAT	wsnp_CAP12_c2297_1121142	3B	278.7	0.000138	0.382	−0.116
IRRIGATED	wsnp_Ex_c55051_57706127	3A	30.82	1.59E–05	0.070	−0.400

For each marker, chromosome (Chr) position (Pos), *P* values, minor allele frequency (maf) and allelic effects (All Eff) are shown. GWAS was conducted after model correction with *Rht-B1* and *Vrn-A1* (used as covariables). GWAS was conducted in normalized averages across four environments (AVG), average of stress environments (AVG STRESS) and in each environment (ENV): IRRIGATED, DROUGHT, HEAT and heat combined with drought (HD). Principal components (PC1 and PC2) from G × E interaction biplot were used in GWAS to identify markers associated with G × E

NA non available

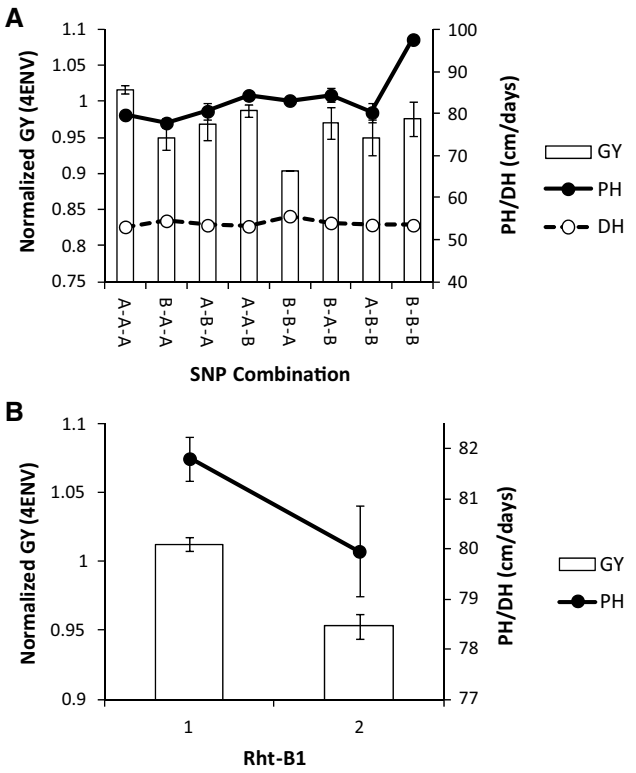

Fig. 4 Average normalized grain yield (GY), average days to heading (DH) and plant height (PH) across four environments of WAMI lines having eight different combinations of SNP markers identified as associated with GY: *wsnp_Ex_c8303_14001708*, *wsnp_Ex_c55051_57706127* and *wsnp_CAP12_c2297_1121142* on chromosomes 2D, 3A and 3B, respectively. Combination of alleles (**a**) and (**b**) are shown in the *XX* axes

to maximize GY but associated genomic regions cannot easily be identified due to the confounding effects of DH and PH. These confounding effects have been observed in almost all QTL discovered so far, particularly for GY and stress adaptive traits that frequently correlate with DH and PH (e.g., Rebetzke et al. 2012; Bennett et al. 2012a, b). Some efforts to optimize DH and PH in experimental populations have been recently presented in the literature with the Seri × Babax population (Pinto et al. 2010; Lopes et al. 2013). The WAMI population was designed for a narrow range of DH and PH variation, but some variation was still observed (9 days and 35 cm on average, respectively, see Lopes et al. 2012). Furthermore, phenotypic correlations were significant for DH and PH with GY and particularly significant under DROUGHT, HEAT and HD. Even under these circumstances (DH and PH correlated with GY), it is worth noting that neither DH nor PH explained more than 12 % of phenotypic yield variation in the WAMI, thus indicating that GY can be further dissected into other minor associated genetic and phenotypic components. Using the WAMI population, despite its narrow range of variation for DH and PH, effects were still observed for *Vrn-A1* and *Rht-B1* on chromosomes 5A and 4B, respectively. It has been reported that *Vrn-1* is dominant epistatic and has the strongest effect in reducing vernalization requirements relative to other *Vrn* genes (Trevaskis et al. 2003) and our results support this. While *Ppd-1* gene was monomorphic, *Vrn-A1* and *Rht-B1*, were responsible for the major G × E interaction for GY in the four different environments where the WAMI was tested. This was shown by significant

associations of the functional markers with PC1 and PC2 scores from GY site regression biplot. Genotype by environment interactions of GY were primarily explained by SNP markers on chromosomes 2D, 3A and the *Rht-B1* functional gene (associated with PC1 of site regression analysis), followed by *Vrn-A1* and nearby markers (associated with PC2 of site regression analysis). This result primarily explained the G × E observed between the stressed environments (DROUGHT, HEAT and HD) where earliness becomes more important to increase GY and the IRRIGATED environment, where increased height may reduce yields due to increased chance of lodging. Yan and Hunt (2001) have inferred this concept when correlating the PCs from site regression with several winter wheat traits and showed that G × E interactions were explained by different traits according to years: some years winter survival scores, other years could relate with heading dates, plant height and lodging. Yan and Hunt (2001) suggested that an increase or decrease in the levels of expression of these traits would, therefore, improve the specific adaptation of the genotypes to certain environments, but it is unlikely to lead to improved overall cultivar performance. The concept suggested by Yan and Hunt (2001) was to target trait optimization rather than maximizing or minimizing a specific trait which will lead only to specific adaptation. The use of PC scores from the G × E biplot seems a promising tool to determine genes and markers associated with complex interactions across environments. Furthermore, Lopes et al. (2014) has shown how these scores are useful to sort markers identified by GWAS to find patterns of important SNPs contributing to yield in clusters of environments.

By removing the effects of major genes *Vrn-A1* and *Rht-B1* (Approach 2, using *Vrn-A1* and *Rht-B1* as covariables in GWAS) in the WAMI, it was possible to identify SNPs exclusively associated with GY, on chromosomes 2D and 3A and with lower allelic effects on chromosome 3B. Moreover, lines with 'A' calls for the SNPs associated with GY on chromosomes 2D, 3A and 3B showed the highest GY while no major changes in PH and DH were observed. Markers on chromosomes 2D are most probably not associated with *Ppd* genes, as almost all (except five) WAMI wheat lines were monomorphic for *Ppd*, plus these markers were not associated with DH or PH in any of the tested environments. Consistent GY associations with markers on chromosome 3B support previous results reported in the literature for QTL identified in the same chromosome (Bonneau et al. 2013).

Concluding remarks

The wheat association mapping initiative (WAMI) has been systematically characterized in terms of population structure, linkage disequilibrium and confounding effects by phenology and plant height The WAMI panel lends itself to GWAS for complex trait dissection by avoiding the confounding effects of DH and PH which were reduced to a minimum, explaining around just 10 % of the phenotypic variation of complex traits like GY. This small confounding effect was further removed using *Rht-B1* and *Vrn-A1* as covariables in GWAS and, therefore, markers associated with GY on chromosomes 2D, 3A and 3B indentified in this study were shown to be exclusively associated with GY. As such the WAMI will be the first model population for marker discovery of complex traits, particularly for adaptation to environmental conditions where the confounding effects of DH and PH are large. It is suggested that the future development of new populations without the confounding effects of DH and PH must be primarily based on: (i) prior phenotypic evaluation in the field of at least 1,000 genotypes for maturity class and height and selection of around 600 genotypes showing the lowest range of variation; (ii) screening for major functional genes like *Rht*, *Vrn* and *Ppd* in the subset selected in (i); (iii) GWAS conducted for DH and PH and major associated genes identified on the subset selected in (i); (iv) selection of genotypes monomorphic for major genes identified in (iii).

Author contribution statement MR.: designed research; M.S.L, J.P., S.D.: performed research; M.S.L., J.P., S.S.: analyzed data; M.S.L.: wrote the manuscript.

Acknowledgments Authors would like to thank Mayra Barcelo, Araceli Torres and Eugenio Perez for assistance with data and trial management. Editing assistance was received from Emma Quilligan. The German Federal Ministry for Economy Cooperation and Development (BMZ) and the Australian Grains Research and Development Corporation (GRDC) are acknowledged for their financial support.

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standards This experiment complies with the current laws of Mexican authorities.

References

Beales J, Turner A, Griffiths S, Snape JW, Laurie DA (2007) A pseudo-response regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (Triticum aestivum L.). Theor Appl Genet 115:721–733. doi:10.1007/s00122-007-0603-4

Bennett D, Izanloo A, Edwards J, Kuchel H, Chalmers K, Tester M, Reynolds M, Schnurbusch T, Langridge P (2012a) Identification of novel quantitative trait loci for days to ear emergence and flag leaf glaucousness in a bread wheat (*Triticum aestivum* L.) population adapted to southern Australian conditions. Theor Appl Genet 124:697–711. doi:10.1007/s00122-011-1740-3

Bennett D, Reynolds M, Mullan D, Izanloo A, Kuchel H, Langridge P, Schnurbusch T (2012b) Detection of two major grain yield QTL

in bread wheat (*Triticum aestivum* L.) under heat, drought and high yield potential environments. Theor Appl Genet 125:1473–1485. doi:10.1007/s00122-012-1927-2

Berloo RV (2008) Computer note: GGT 2.0: versatile software for visualization and analysis of genetic data. J Hered 99:232–236. doi:10.1093/jhered/esm109

Bonneau J, Taylor J, Parent B, Bennett D, Reynolds M, Feuillet C, Langridge P, Mather D (2013) Multi-environment analysis and improved mapping of a yield-related QTL on chromosome 3B of wheat. Theor Appl Genet 126:747–761. doi:10.1007/s00122-012-2015-3

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Cavanagh CR, Chao S, Wang HBE, Stephen S, Kiani S, Forrest K, Saintenac C, Brown-Guedira GL, Akhunova A, See D, Bai G, Pumphrey M, Tomar L, Wong D, Kong S, Matthew Reynolds M, Lopez MS, Bockelman H, Talbert L, Anderson JA, Dreisigacker S, Baenziger S, Carter A, Korzun V, Morrell PL, Dubcovsky J, Morell MK, Sorrells ME, Hayden MJ, Akhunov E (2013) Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. Proc Nat Acad Sci 110:8062–8067

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

Chao S, Dubcovsky J, Dvorak J, Luo MC, Baezinger SP, Matnyazov R, Clark DR, Talbert LE, Anderson JA, Dreisigacker S, Glover K, Chen J, Campbell K, Bruckner PL, Rudd JC, Haley S, Carver BF, Perry S, Sorrells ME, Akhunov ED (2010) Population and genome specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (*Triticum aestivum* L.). BMC Genom 11:727. doi:10.1186/1471-2164-11-727

Charmet G, Masood-Quraishi U, Ravel C, Romeuf I, Balfourier F, Perretant MR, Joseph JL, Rakszegi M, Guillon F, Sado PE, Bedo Z, Saulnier L (2009) Genetics of dietary fibre in bread wheat. Euphytica 170:155–168. doi:10.1007/s10681-009-0019-0

Chen X, Min D, Yasir TA, Hu YG (2012) Genetic diversity, population structure and linkage disequilibrium in elite Chinese winter wheat investigated with SSR markers. PLoS ONE 7:e44510. doi:10.1371/journal.pone.0044510

Crossa J, Burgueno J, Dreisigacker S, Vargas M, Herrera-Foessel SA, Lillemo M, Singh RP, Trethowan R, Warburton M, Franco J, Reynolds M, Crouch JH, Ortiz R (2007) Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. Genetics 177:1889–1913. doi:10.1534/genetics.107.078659

Diaz A, Zikhali M, Turner AS, Isaac P, Laurie DA (2012) Copy number variation affecting the photoperiod-B1 and Vernalization-A1 genes is associated with altered flowering time in wheat (Triticum aestivum). PLoS ONE 7(3):e33234. doi:10.1371/journal.pone.0033234

Dodig D, Zoric M, Kobiljski B, Savic J, Kandic V, Quarrie S, Barnes J (2012) Genetic association mapping study of wheat agronomic traits under contrasting water regimes. Int J Mol Sci 13:6167–6188. doi:10.3390/ijms13056167

Dreisigacker S, Shewayrga H, Crossa J, Arief VN, DeLacy IH, Singh RP, Dieters MJ, Braun HJ (2012) Genetic structures of the CIMMYT international yield trial targeted to irrigation environments. Mol Breed 29:529–541. doi:10.1007/s11032-011-9569-7

Edae EA, Byrne PF, Haley SD, Lopes MS, Reynolds MP (2014) Genome-wide association mapping of yield and yield components of spring wheat under contrasting moisture regimes. Theor Appl Genet in press doi:10.1007/s00122-013-2257-8

Ellis MH, Spielmeyer W, Gale KR, Rebetzke GJ, Richards RA (2002) Perfect markers for the Rht-B1b and Rht-D1b dwarfing genes in wheat. Theor Appl Genet 105:1038–1042. doi:10.1007/s00122-002-1048-4

Everitt B (2007) Principal component analysis. In: An R and S-Plus companion to multivariate analysis. London, Springer-Verlag pp 41–62

Fu DL, Szücs P, Yan LL, Helguera M, Skinner JS, Zitzewitz JV, Hayes PM, Dubcovsky J (2005) Large deletions within the first intron in *VRN*-1 are associated with spring growth habit in barley and wheat. Mol Gen Genomics 273:54–65. doi:10.1007/s00438-004-1095-4

Gaut BS, Long AD (2003) The lowdown on linkage disequilibrium. Plant Cell 15:1502–1506. doi:10.1105/tpc.150730

Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, Eskin E (2008) Efficient control of population structure in model organism association mapping. Genetics 178:1709–1723. doi:10.1534/genetics.107.080101

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z. (2012) GAPIT: genome association and prediction integrated tool. Bioinformatics pp 1–2. doi:10.1093/bioinformatics/bts444

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129. doi:10.1093/bioinformatics/bti282

Liu L, He ZH, Yan J, Zhang Y, Xia X, Peña RJ (2005) Allelic variation at the Glu-1 and Glu-3 loci, presence of the 1B.1R translocation, and their effects on mixographic properties in Chinese bread wheats. Euphytica 142:197–204. doi:10.1007/s10681-005-1682-4

Lobell DB, Schlenker W, Costa-Roberts J (2011) Climate trends and global crop production since 1980. Science 333:616–620. doi:10.1126/science.1204531

Lopes MS, Reynolds MP, Jalal-Kamali MR, Moussa M, Feltaous Y, Tahir ISA, 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 Res 128:129–136. doi:10.1016/j.fcr.2011.12.017

Lopes MS, Reynolds MP, McIntyre CL, Mathews KL, Jalal Kamali MR, Mossad M, Feltaous Y, Tahir ISA, Chatrath R, Ogbonnaya F, Baum M (2013) QTL for yield and associated traits in the Seri/Babax population grown across several environments in Mexico, in the West Asia, North Africa, and South Asia regions. Theor Appl Genet 126:971–984. doi:10.1007/s00122-012-2030-4

Lopes MS, Rebetzke G, Reynolds M (2014) Integration of phenotyping and genetic platforms for a better understanding of wheat performance under drought. J Exp Bot in press doi:10.1093/jxb/eru384

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

Maccaferri M, Samguineti MC, Demontis A, El-Ahmed A, del Moral LG, 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. J Exp Bot 62:409–438. doi:10.1093/jxb/erq287

Olesen JE, Trnka M, Kersebaum KC, Skjelvåg AO, Seguin B, Peltonen-Sainio P, Rossi F, Kozyra J, Micale F (2011) Impacts and adaptation of European crop production systems to climate change. Eur J Agron 34:96–112. doi:10.1016/j.eja.2010.11.003

Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. PLoS Genet 2:e190. doi:10.1371/journal.pgen.0020190

Pinto RS, Reynolds MP, Mathews KL, McIntyre CL, Olivares-Villegas JJ, Chapman SC (2010) Heat and drought adaptive QTL

in a wheat population designed to minimize confounding. Theor Appl Genet 121:1001–1021. doi:10.1007/s00122-010-1351-4

Pritchard JK, Stephens M, Rosenberg NA, Donnelly P (2000) Association mapping in structured populations. Am J Hum Genet 67:170–181

Rebetzke GJ, Rattey AR, Farquahar GD, Richards RA, Condon AG (2012) Genomic regions of canopy temperature and their genetic association with stomatal conductance and grain yield in wheat. Funct Plant Biol 40:14–33

Rehman Arif MA, Nagel N, Newmann K, Kobiljski B, Lohwasser U, Borner A (2012) Genetic studies of seed longevity in hexaploid wheat using segregation and association mapping approaches. Euphytica 1869:1–13. doi:10.1007/s10681-011-0471-5

Reynolds MP, Manes Y, Izanloo A, Langridge P (2009) Phenotyping for physiological breeding and gene discovery in wheat. Ann Appl Biol 155:309–320. doi:10.1111/j.1744-7348.2009.00351.x

Richards RA, Rebetzke GJ, Watt M, Condon AG, Spielmeyer W, Dolferus R (2010) Breeding for improved water–productivity in temperate cereals: phenotyping, quantitative trait loci, markers and the selection environment. Funct Plant Biol 37:1–13

Risch N, Merikangas K (1996) The future of genetic studies of complex human diseases. Science 273:1516–1517. doi:10.1126/science.273.5281.1516

Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. Proc Nat Acad Sci 81:8014–8018

Samonte SOPB, Wilson LT, McClung AM, Medley JC (2005) Targeting cultivars onto rice growing environments using AMMI and SREG GGE biplot analysis. Crop Sci 45:2414–2424. doi:10.2135/cropsci2004.0627

Schwarz EG (1978) Estimating the dimension of a model. Ann Stat 6:461–464. doi:10.1214/aos/1176344136

Sharma RC, Crossa J, Velu G, Huerta-Espino J, Vargas M, Payne TS, Singh RP (2012) Genetic gains for grain yield in CIMMYT spring bread wheat across international environments. Crop Sci 52:1522–1533. doi:10.2135/cropsci2011.12.0634

Trevaskis B, Bagnall DJ, Ellis MH, Peacock WJ, Dennis ES (2003) MADS box genes control vernalization-induced flowering in cereals. Proc Natl Acad Sci USA 100:130–13104. doi:10.1073/pnas.1635053100

Weir BS (1996) Genetic data analysis II. Sinauer Associates, Sunderland

Wilhelm EP, Turner AS, Laurie DA (2009) Photoperiod insensitive Ppd-A1a mutations in tetraploid wheat (Triticum durum Desf.). Theor Appl Genet 118:285–294. doi:10.1007/s00122-008-0898-9

Yan W, Hunt LA (2001) Interpretation of genotype x environment interaction for winter wheat yield in Ontario. Crop Sci 41:19–25. doi:10.2135/cropsci2001.41119x

Yan LL, Helguera M, Kato K, Fukuyama S, Sherman J, Dubcovsky J (2004) Allelic variation at the *VRN*-1 promoter region in polyploid wheat. Theor Appl Genet 109:1677–1686. doi:10.1007/s00122-004-1796-4

Zhang P, Dreisigacker S, Melchinger AE, Reif JC, Mujeeb Kazi A, Van Ginkel M, Hoisington D, Warburton ML (2005) Quantifying novel sequence variation and selective advantage in synthetic hexaploid wheats and their backcross-derived lines using SSR markers. Mol Breed 15:1–10. doi:10.1007/s00122-004-1796-4

Zhang LY, Liu DC, Guo XL, Yang WL, Sun JZ, Wang DW (2011) Investigation of genetic diversity and population structure of common wheat cultivars in northern China using DArT markers. BMC Genet 12:42. doi:10.1186/1471-2156-12-42