

Genetics of Fe, Zn, β -carotene, GPC and yield traits in bread wheat (*Triticum aestivum* L.) using multi-locus and multi-traits GWAS

Jitendra Kumar · Gautam Saripalli · Vijay Gahlaut · Neha Goel ·
Prabina Kumar Meher · Kaushlesh Kumar Mishra · Prafulla Chandra Mishra ·
Deepmala Sehgal · Prashant Vikram · Carolina Sansaloni · Sukhwinder Singh ·
Pradeep Kumar Sharma · Pushendra Kumar Gupta

Received: 8 August 2018 / Accepted: 13 October 2018 / Published online: 3 November 2018
© Springer Nature B.V. 2018

Abstract The present study was conducted to study the genetic architecture of grain micronutrients (Zn, Fe and β -carotene contents), grain protein content and four yield traits in a spring wheat reference set comprising 246 genotypes. Phenotypic data on these traits recorded at two locations and the genotyping data for 17,937 SNP markers (obtained through outsourcing) were used for genome wide association study, which gave following results after Bonferroni correction using four methods: (1) single locus single

trait analysis gave 136 marker-trait associations; (2) multi-locus mixed model gave 587 MTAs; (3) multi-trait mixed model gave 28 MTAs and (4) matrix-variate linear mixed model gave 33 MTAs. As many as 73 epistatic interactions were also detected. Keeping all the results in mind, nine most important MTAs were selected for biofortification. These markers were associated with three traits (GPC, GFeC and GYPP). These MTAs can be used in wheat improvement programs either using marker-assisted recurrent selection or pseudo-backcrossing method.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10681-018-2284-2>) contains supplementary material, which is available to authorized users.

J. Kumar · G. Saripalli · V. Gahlaut · N. Goel ·
P. K. Sharma · P. K. Gupta (✉)
Department of Genetics and Plant Breeding, Ch. Charan
Singh University, Meerut, India
e-mail: pkgupta36@gmail.com

D. Sehgal · P. Vikram · C. Sansaloni · S. Singh (✉)
International Maize and Wheat Improvement Center
(CIMMYT), Km. 45, Carretera Mex-Veracruz, El Batan,
CP 56237 Texcoco, Mexico
e-mail: Suk.Singh@cgiar.org

P. K. Meher
Centre for Agricultural Bioinformatics, ICAR-Indian
Agricultural Statistics Research Institute, New Delhi,
India

K. K. Mishra · P. C. Mishra
Zonal Agricultural Research Station, Powarkheda,
Jawaharlal Nehru Krishi Vishwa Vidyalaya, Jabalpur, India

Keywords Bread wheat (*Triticum aestivum* L.) · GWAS · SLST · MLMM · MTMM · mvLMM

Introduction

Bread wheat is the third most important food crop globally (after rice and maize) in terms of production and consumption. It is consumed by more than 40% world population as a staple food and is the primary source of calories for millions of people world-wide. However, the crop is deficient for major micronutrients like Fe, Zn and β -carotenoid, which are present only as minor constituents of wheat grain. Over three billion people, including one third of the children in developing countries suffer from micronutrient malnutrition or hidden hunger (Chattha et al. 2017). Deficiency of these micronutrients is also witnessed in the form of metabolic disorders like anaemia (due to Fe deficiency), night blindness and xerophthalma, cardiovascular diseases and a variety of cancers and neurological disorders (due to β -carotenoid deficiency; Colasuonno et al. 2017). There are also reports of poor pregnancy outcomes like impaired or stunted growth in children due to Zn deficiency. It is estimated that > 60% of the world population suffers from Fe deficiency and > 30% population suffers from Zn deficiency (White and Broadley 2009). Similarly, 190 million pre-school-children and 19.1 million pregnant women around the world suffer from β -carotenoid deficiency (WHO 2016).

Grain protein content (GPC) is another important trait, which has an impact on the nutritional value of the grain and also on the technological property of the flour. Protein content and essential amino acids also affect the key functions of the human body including development and maintenance of muscles. Protein energy malnutrition (PEM) has been noticed among 161 million children globally (<http://www.worldhunger.org/2015-world-hunger-and-poverty-facts-and-statistics/>).

Biofortification through genetic manipulations is known to be one of the best options for nutritional improvement of crops (Welch and Graham 2004; Ortiz-Monasterio et al. 2007). The most important work on genetic improvement for micronutrients contents (biofortification) in staple food crops (rice, wheat, maize, cassava, sweet-potato, pearl-millet and

bean) has been conducted under HarvestPlus project launched in 2004 by International Agricultural Research Consortium (IARC) of CGIAR. During HarvestPlus Phase I (2003–2008) involving screening of 3000 wheat accessions, contents of Zn and Fe were found to be in the range of 20–115 ppm and 23–88 ppm, respectively, with the highest levels found in landraces (<https://biofortconf.ifpri.info>). In recent years, biofortified varieties in some crops have also been released in ~ 30 countries including India in respect of various micronutrients (<http://www.harvestplus.org>).

The variability for micronutrients (including Fe, Zn and provitamin A) in wheat is largely genetic in nature with a complex polygenic control (Shi et al. 2008; Joshi et al. 2010; Velu et al. 2012; Srinivasa et al. 2014). This makes the improvement in these traits difficult through conventional breeding (Velu et al. 2012). Marker-assisted backcrossing (MABC) and marker-assisted recurrent selection (MARS) are good options. These methods would require determination of marker-trait associations through linkage-based interval mapping and LD-based genome wide association studies (GWAS). This will also help in understanding the details of genetic architecture of the traits (Tiwari et al. 2009). A number of interval mapping studies have already been conducted to identify QTLs for micronutrients in wheat (Pozniak et al. 2007; Peleg et al. 2009; Tiwari et al. 2009; Xu et al. 2012; Roshanzamir et al. 2013; Zhao et al. 2013; Tiwari et al. 2016; Sharma et al. 2018). GWAS has also been conducted for micronutrients in different crops like rice (Norton et al. 2014), pea (Diapari et al. 2015), maize (Suwarno et al. 2015), barley (Leplat et al. 2016), chickpea (Diapari et al. 2014) and cassava (Esuma et al. 2016; Rabbi et al. 2017). Some reports involving GWAS for Zn, Fe and carotenoid contents are also available in wheat (Gorafî et al. 2016; Manickavelu et al. 2017; Colasuonno et al. 2017). However, a majority of GWA studies conducted so far are based on the study of single locus and single trait, which gives information of limited utility. Often epistasis is also ignored during GWAS, although there are few GWA studies, where epistasis was examined (Jaiswal et al. 2016; Sehgal et al. 2017). Recently, MLMM and MTMM have become available for GWAS, which overcome the above limitations of genetic analysis (Segura et al. 2012; Korte et al. 2012; Jaiswal et al. 2016; Thoen et al. 2017). mvLMM has

also been used so that more than two correlated traits may be examined in multi-trait analysis (Zhou and Stephens 2012; Furlotte and Eskin 2015).

SNPs are the most abundant class of markers associated with sequence variability in the genome and thus have the potential to provide the highest map resolution. Therefore, SNPs have become the markers of choice for GWAS (Jones et al. 2007). In the present study, GWAS was conducted for micronutrients and other yield-related quantitative traits in a set of wheat genotypes. The study involved single locus single trait, MLMM, MTMM and mvLMM; where both main effects and epistatic MTAs were also identified. The results of this study should prove useful for developing wheat cultivars with improved nutritional value using MAS/MARS.

Materials and methods

Association mapping panel and genotyping

The association mapping panel comprised 246 wheat genotypes of a spring wheat reference set (SWRS) procured from CIMMYT gene bank, Mexico. A set of 17,937 SNP markers generated using DArT-seq, at Diversity Array Technology Pvt. Ltd. Australia under the “Seed for Discovery” project of CIMMYT Mexico, was used for genotyping of all the 246 accessions of bread wheat (Table S1). The markers were mapped on all the 21 chromosomes using DArT PL’s consensus map of wheat based on > 100 crosses. The map (version 4) has 110,000 markers including ~ 5000 original DArT markers, the remaining being DArTseq markers. Only 8637 SNPs could be placed on genetic map involving all the 21 chromosomes; 2973 belonged to A sub-genome, 4505 belonged to the B sub-genome and 1159 belonged to the D sub-genome.

Field trials and experimental data

The above association panel was raised in a simple lattice design with two replications at two different locations, during rabi season of 2013–2014 at Pow-erkheda (Location coordinates: 22°40′50.01N 77°44′59.18E), Madhya Pradesh, India and during 2014–2015 at Meerut (Location coordinates: 28.9845°N, 77.7064°E), Uttar Pradesh, India, using

normal field management practices (i.e., 200 kg/ha fertilizer; N:P:K = 8:8:8). Each genotype was raised in a plot of 3 rows of 1.5 m each, with a row to row distance of 0.25 m. Phenotypic data were recorded for four nutritional and four yield-related traits. The nutritional traits included the following: grain protein content (GPC) as per cent of grain weight, (2) grain β -carotenoid content (GBCC; $\mu\text{g/g}$), (3) grain iron content (GFeC; ppm), and (4) grain zinc content (GZnC; ppm). Similarly, the yield traits included the following: (1) tiller number per plant (TNPP), (2) grain number per spike (GNPS), (3) thousand grain weight (TGW) in g, (4) grain yield per plot (GYPP) in kg/ha.

For GPC and GBCC, 50–60 g of seed of each sample was used. GPC (%) was adjusted to 12% moisture content using the following formula and the value obtained was used for further analysis.

$$\text{Adjusted GPC (\%)} \text{ value} = \frac{100 - \text{Base M.C. (12\%)}}{100 - \text{Actual M.C. (\%)}} \times \text{Measured GPC (\%)} \text{ value}$$

where M.C. is moisture content.

For estimating Zn and Fe content, 10–12 g (90–100 kernels) seed of each genotype was used. GPC and GBCC data was recorded using Infratec (1241) Grain Analyzer available at CCSU, Meerut. Data on Zn and Fe contents were recorded using X-ray Fluorescence (EDXRF spectrometer X-Supreme 8000; Paltridge et al. 2012) available at Department of Genetics and Plant Breeding of Banaras Hindu University (BHU). The data on yield-related traits was recorded using the traditional methods. For each trait, the data was recorded for all replications. Only means of data over replications were used for further analysis.

Statistical analysis (Descriptive statistics, Pearson’s correlation coefficients, analysis of variance (ANOVA) and heritability)

The estimates of descriptive statistics including mean, range, standard error, coefficient of variation (CV as %), Pearson’s correlation coefficients were obtained using SPSS v. 17.0. ANOVA was conducted using additive main effects and multiplicative interactions (AMMI) model through Agricolae package of R program. Broad sense heritability (H^2) estimates were calculated from phenotypic variance (σ^2_p) and

the genotypic variance (σ_g^2) according to Allard (1999) using MS Excel 2010.

Population structure analysis

Model-based cluster analysis of association mapping panel was conducted to infer the level of population structure in the association panel using the software STRUCTURE version 2.2 (Pritchard et al. 2000), which was performed using 42 SNPs, one from each of the 42 chromosome arms.

The number of assumed sub-populations (K) was set from 2 to 20, and the process was repeated three to five times. For each run, burn-in and Markov Chain Monte Carlo (MCMC) iterations were set to 50,000 and 100,000, respectively and a model “without admixture and correlated allele frequencies” was used. The number of sub-populations was determined following delta K (ΔK) method (Evanno et al. 2005). Assignment of genotypes to sub-populations was done on the basis of their affiliation probabilities. A genotype was assigned to a specific sub-population, with which it has $\geq 80\%$ affiliation probability; genotypes with $\leq 80\%$ affiliation probability with each sub-population were treated as “admixtures”.

Linkage disequilibrium (LD) analysis

LD (in terms of r^2) analysis was performed for each of the 21 wheat chromosomes with associated mapped SNPs using window size 50 with the help of software TASSEL v. 5.0. Genome wide threshold LD was calculated using unlinked markers following Breseghello and Sorrells 2006.

Marker-trait associations (MTAs)

Phenotypic data for the micronutrient contents, GPC, and yield related traits for 246 genotypes of each location and the corresponding genotypic data of SNP markers were used for single locus single traits analysis to identify MTAs. TASSEL v. 3.0 was used for this purpose using both General Linear Model (GLM) and Mixed Linear Model (MLM). For GLM, population structure (the Q model) without familial relatedness (the K model) was used, whereas for MLM, both population structure and the familial relatedness (Q + K model) were used (Yu et al. 2006). Familial relatedness (kinship matrix) was

calculated using the genotypic data for all the 17,937 markers.

GWA analysis was also conducted using MLM, MTMM and mvLMM. The MLM, MTMM and mvLMM analyses were performed using relevant R packages (Segura et al. 2012; Korte et al. 2012; Furlotte and Eskin 2015). For MLM, background genome was considered as cofactors (as in composite interval mapping) using stepwise mixed-model regression with forward inclusion and backward elimination (Segura et al. 2012; Jaiswal et al. 2016). For MTMM, all pairs of phenotypic traits showing significant correlation were used. Similarly, for mvLMM, more than two phenotypic traits showing significant correlation with each other were used. In all cases, P value ≤ 0.001 was considered for identification of significant MTAs. Stringent criteria of $P \leq 0.001$ was used to deal with the problem of multiple testing. Each significant MTA was subjected to Bonferroni multiple correction (using additional $\alpha_{\text{GWAS}} \leq 0.05$, thus making the overall significant threshold to be $P \leq 0.00005$) for eliminating false positives in case of single locus single trait and MTMM. However, for MLM and mvLMM, Bonferroni correction was not needed, since the packages used already had a provision for Bonferroni correction.

Analysis of epistasis

Analysis for epistatic interactions was carried out using SNPAssoc package (González et al. 2007) of R program, where the function interactionPval was executed for computing the statistical significance (P value) of SNP–SNP interaction.

Results

Phenotypic data and correlations

The eight different phenotypic traits were placed in the following two groups for presentation of results: (1) nutritional traits, and (2) yield-related traits; Violin plots representing the frequency distributions of data for all the eight traits on each of the two locations are presented in Fig. 1a, b; the data for each trait gave a good fit to normal distribution. The results of descriptive statistics; the range of mean values were as follows: GPC, 9.99–18.87 (%); GBCC, 2.94–6.555

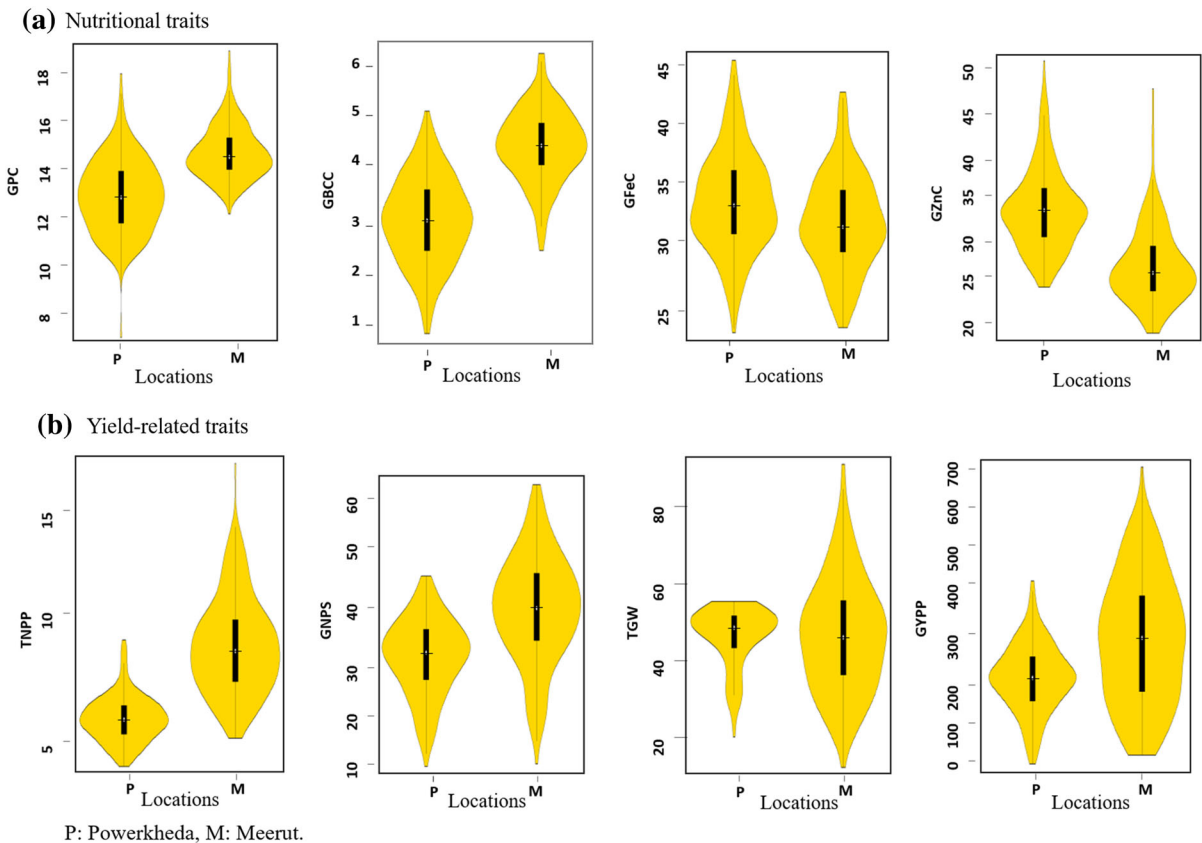


Fig. 1 a, b Violin plots showing the frequency distribution of 4 nutritional and 4 yield-related traits at two locations. Shaded regions of the violin plots represent the frequency distribution of data, in each case, the vertical solid bar indicates range of

(mg/kg); GFeC, 24.50–44.30 (ppm); GZnC, 17.75–49.70 (ppm); TNPP, 3.00–14.67; GNPS, 10.50–58.00; TGW, 12.43–59.36 g; GYPP, 42.50–520.00 kg/ha. The details of descriptive statistics and CV are presented in Table S2.

Correlations in majority of the pairs of traits at both the locations were highly significant at $P \leq 0.01$ (Table 1). As many as 14 of the possible 28 pairs of traits, each having a correlation above 0.25, were selected for MTMM analysis; four of these correlations were consistent at both the locations. Of the remaining ten correlations, seven were available in data from Meerut and only three were available in data from Powerkheda. Correlations involving more than two traits were also available (e.g. GZnC and GYPP); in such cases, mvLMM approach was also used for determining MTAs involving more than two correlated traits.

average values, and median is shown as a minute white circle within the solid bar, with a horizontal bar, depicting the lower, medium and upper quartile

ANOVA and heritability

The combined ANOVA revealed highly significant variation for each trait with the following sources of variation: genotypes, environments and $G \times E$ interaction. Estimates of broad sense heritability (H^2) of nutritional traits ranged from 6.65% (GPC) to 62.05% (GFeC) and that for yield traits ranged from 52% (TNPP) to 89.66% (GYPP; Table 2).

Population structure and LD analysis

Model-based cluster analysis revealed that the AM panel used in the present study is structured and comprised four subpopulations viz. G1, G2 G3 and G4. The four sub-populations included 30 (G1), 42 (G2), 49 (G3) and 125 (G4; admixture) genotypes, respectively (Fig. 2). The information generated by

Table 1 Estimates of Pearson correlation coefficients for different traits of the association panel at two locations

Traits	Nutritional traits				Yield traits			
	GPC	GBCC	GFeC	GZnC	TNPP	GNPS	TGW	GYPP
GPC	1	-0.38**	0.24**	0.42**	0.09	-0.35**	-0.30**	-0.46**
GBCC	0.03	1	0.07	-0.15*	-0.03	0.26**	0.18**	0.24**
GFeC	0.61**	-0.07	1	0.19**	-0.1	0	0.12	0.1
GZnC	0.12	0.11	0.08	1	0.12	-0.37**	-0.24**	-0.47**
TNPP	0.05	-0.03	0.06	0.11	1	-0.15*	-0.1	-0.14*
GNPS	0.02	-0.11	0.09	-0.33**	-0.09	1	0.47**	0.51**
TGW	-0.08	-0.25**	0.09	-0.26**	-0.02	0.31**	1	0.30**
GYPP	0.02	-0.08	0.06	-0.42**	-0.13*	0.33**	0.19**	1

*, **correlations are significant at the 0.05 and 0.01 level, respectively

■ Upper panel indicates correlation values for different traits recorded in 2014–15 at Meerut

■ Lower panel indicates correlation values for different traits recorded in 2013–14 at Powerkheda

Table 2 Analysis of variance (ANOVA) and heritability (H^2) of yield traits, GPC and micronutrient contents in SWRS genotypes tested for 2 years/environments and four replications (two in each environment)

Source of variations	df	Mean square							
		Nutritional traits				Yield traits			
		GPC	GBCC	GFeC	GZnC	TNPP	GNPS	TGW	GYPP
Env	1	772.87*	611.18**	204	9082.1*	3554.1***	21382.5**	941.53	1511148**
Rep within Env	2	21.03	1.77**	12.68*	269.1	1.3	26.6	134.56*	9901***
Genotypes	245	11.80*	1.57***	30.12***	153.6***	8***	211.1***	390.89***	26243***
Env × Gen	245	10.93	1.38***	22.88***	139.1**	6.7***	131.1***	237.72***	21083***
Heritability %		6.65	43.91	62.05	11.63	52	66.83	71.9	89.66

***0.1% level of significance; **1% level of significance; *5% level of significance

population structure was used for analysing marker-trait associations to reduce the number of false positives.

LD between pairs of markers was estimated for each of the 21 chromosomes. The measures of LD decay for all the 21 chromosomes are summarized in Figs. S1. The number of SNP pairs with significant LD ($P \leq 0.01$) on individual chromosome ranged from 328 (on chromosome 4D) to 6955 (on chromosome 1A). Genome-wide threshold LD was $r^2 = 0.21$ and the genome-wide mean genetic distance showing no LD decay was 3.0 cM. The genetic distance showing LD ranged from a minimum of 2 cM in certain regions of chromosomes 2B, 3A, 3B and 6B to a maximum of 20 cM in a region on chromosome 3D (Fig. S1).

MTAs identified using single locus single trait and MLM

MTAs for two groups of traits including nutritional traits and yield-related traits are described separately. Results of MTMM and mvLMM are described for both the groups together, since the analyses involved two and more than two traits, sometime belonging to both the groups. The most important MTAs are depicted in Fig. 3.

Nutritional traits

Using SLST, a total of 584 significant MTAs for four nutritional traits at two locations were identified; only

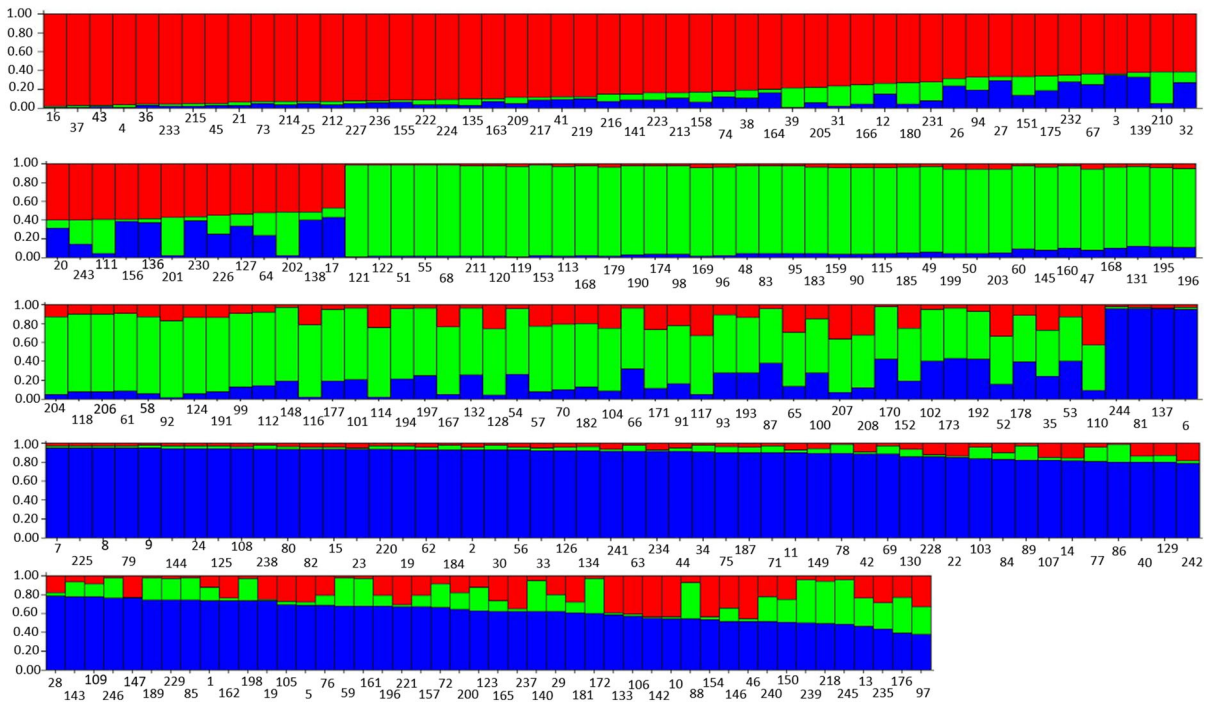


Fig. 2 AM panel showing structuring of four subpopulations in different colours, viz. G1 (red colour, 30 genotypes), G2 (green colour, 42 genotypes), G3 (blue colour, 49 genotypes) and G4 (a single bar with two or more colours, admixture of 125 genotypes)

10 of these MTAs passed Bonferroni correction, of which one belonged to sub-genome A, six belonged to sub-genome B, and none belonged to D genome; the associated markers for the remaining three MTAs were not assigned to chromosomes (Table S3, S4 and Fig. S2). The above 10 MTAs all belonged to Meerut location only.

Using MLM, 271 MTAs (after Bonferroni correction) were identified for the same four nutritional traits. The 164 MTAs belonged to Powerkheda, 106 belonged to Meerut and only one belonged to both the locations (Table S3). Of these only 144 MTAs were mapped on chromosomes, with 58 for sub-genome A, 66 for sub-genome B and 20 for sub-genome D (Table S3, S5 and Fig. S2).

Yield-related traits

For the four yield-related traits, SLST gave a total of 3251 MTAs involving both the locations ($P < 0.001$). But only 126 of these MTAs passed Bonferroni correction, of which 96 MTAs were detected at Powerkheda, 8 were detected at Meerut and 22 at both the locations (Table S3, S4 and Fig. S2).

Similarly, using MLM, 316 MTAs (Bonferroni correction passed) were identified for these traits (149 MTAs at Powerkheda, 161 at Meerut and 6 at both the locations). Of the 316 associated markers, only 209 markers were mapped on different chromosomes, with 101 on sub-genome A, 92 for sub-genome B and 16 for sub-genome D (Table S3, S5 and Fig. S2).

MTAs/QTLs identified using MTMM and mvLMM

Using MTMM, 1253 MTAs were identified for 14 pairs of correlated traits at both the locations (no MTA was available for GPC/GBCC, TGW/GPC and TGW/GBCC), but only 28 of these MTAs (6 for GPC/GFeC, Powerkheda and Meerut; 1 for GYPP/TGW, Meerut; 21 for GPC/GZnC, Meerut) for three pairs of traits qualified after Bonferroni correction. Only 18 MTAs could be mapped on A and B sub-genomes (3 belong to sub-genome A; 15 belong to sub-genome B), none was found on sub-genome D (Table S6 and S7; Fig. 2).

Using mvLMM, 33 MTAs were identified at both the locations for eight pairs of correlated traits. Seven MTAs for three pairs of traits were identified for

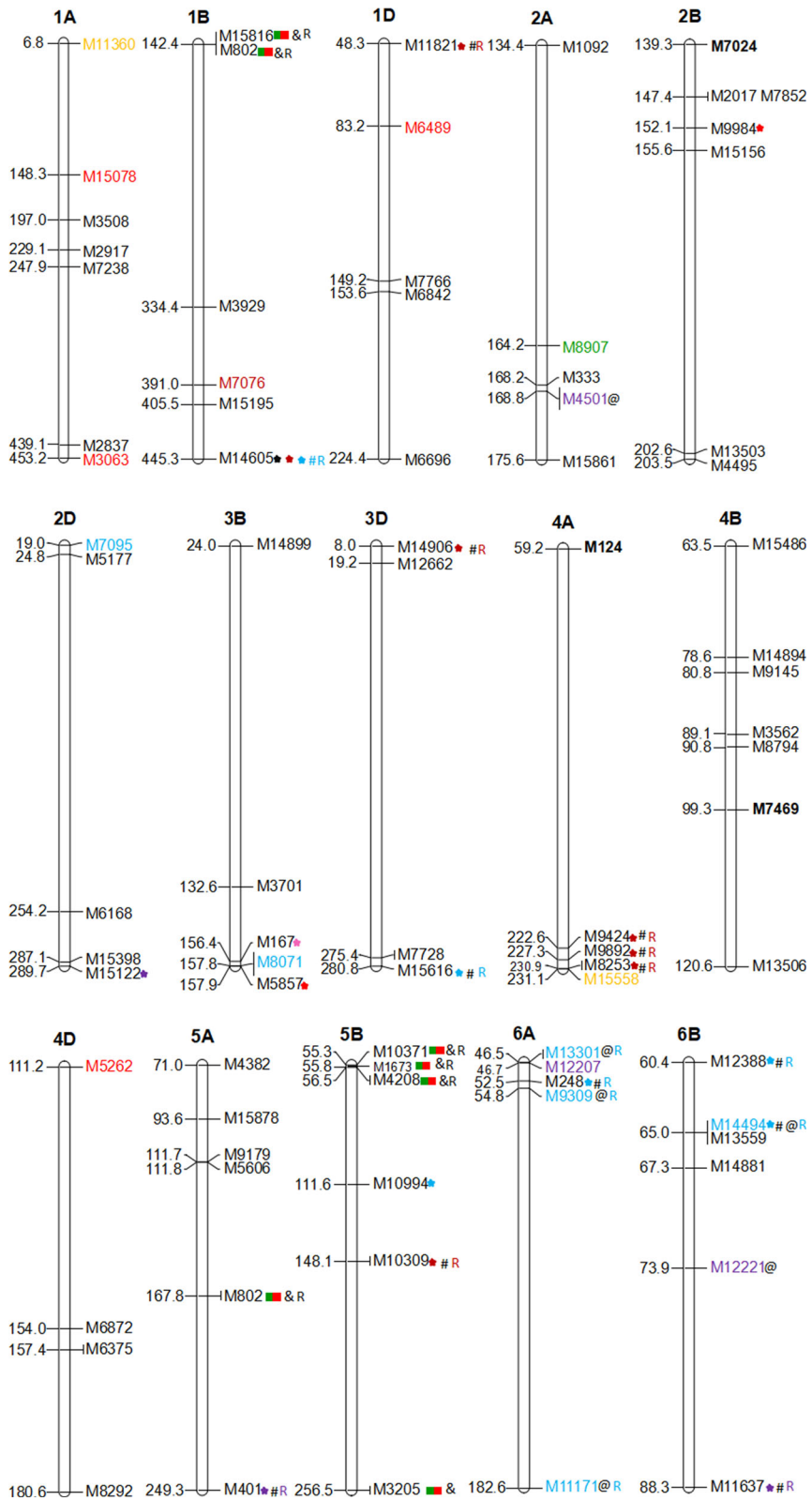


Fig. 3 Distribution of significant MTAs on different chromosomes identified using SLST, MLMM and MTMM. Different colour shaded star ★ symbols indicate different traits associated with SNPs identified using SLST (GNPS-★, GYPP-★, GPC-★, GBCC-★, GZnC-★). Similarly, different coloured marker names indicate the MTAs identified using MLMM (TNPP, TGW, GPC, and GBCC) and different combinations of coloured squares, like ■ indicate MTAs identified using MTMM (GPC/GFeC-■). MTAs indicated with #, @ and & indicate those detected in both the environments using SLST, MLMM and MTMM, respectively. MTAs indicated with R are those which were reported earlier

Powerkheda location; five of these seven MTAs were mapped, one on sub-genome A (3A) and four on sub-genome B (one each on 2B, 3B and two on 6B). The remaining 26 MTAs involving five pairs of traits were identified at Meerut location; of these, 4 MTAs belonged to sub-genome A; 11 belonged to sub-genome B and only 2 belonged to sub-genome D. The remaining nine MTAs involved SNPs that were not mapped (Table 3).

There were also MTAs that were identified each by more than one approaches. There were also MTAs, which were common for more than one traits. There were 22 MTAs, involving three yield traits, each identified by both SLST and MLMM (13 MTAs for GYPP, 6 for GNPS and 3 for TGW). There were another 12 MTAs involving three pairs of correlated traits (with 5 traits) that were detected by both MTMM

and MLMM, thus placing a higher level of confidence in these markers (Table S8). Similarly, mvLMM detected one MTA (involving marker M5415) that was associated with a combination of four traits (GYPP/GNPS/TGW/GZnC); this marker was also detected by MLMM for GYPP. Thus, there were more several individual MTAs, each involving more than one trait; these MTAs may or may not represent pleiotropic genes/QTLs. These MTAs included the following: (1) 1 MTA detected using SLST was common between two traits; (2) 32 MTAs detected using MLMM involved 8 traits (Table S8). All these markers are depicted in Fig. 3.

MTAs involved in epistatic interactions

For six of the eight traits (excluding GFeC and TNPP), 73 epistatic interactions involving 146 markers {spread over 18 of the 21 chromosomes (except 4D, 5D and 6D)} were identified (43 at Powerkheda and 30 at Meerut). The interactions for individual traits ranged from only one interaction for TGW to a maximum of 27 interactions for GYPP (Table S9).

Relative importance of MTAs

The MTAs identified by all the methods were subjected to scrutiny in order to identify the most important MTAs, which could be recommended for

Fig. 3 continued

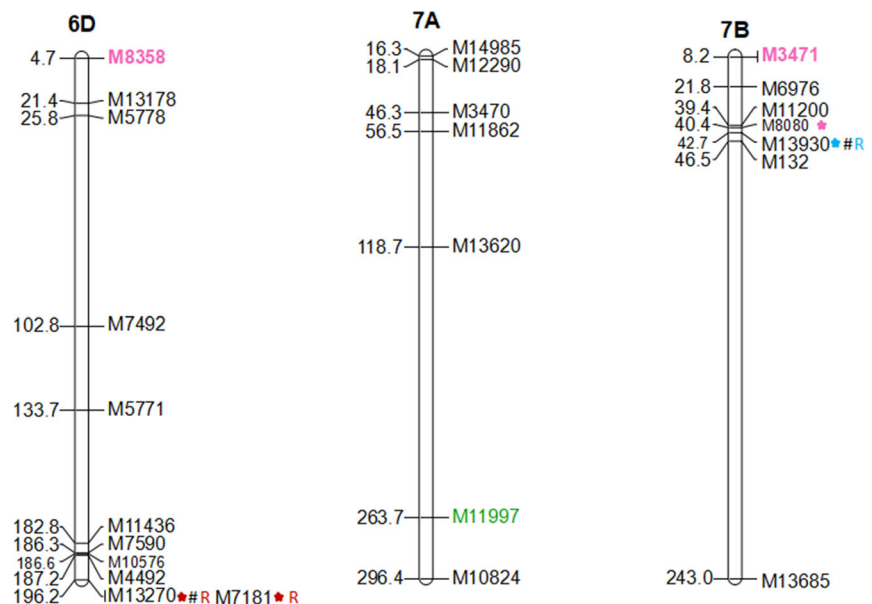


Table 3 A summary of significant ($P \leq 0.001$) MTAs detected using mvLMM at Powerkheda and Meerut

Locations	Trait combinations and associated markers	SNP	SNP ID	Ch.	Pos. (cM)	<i>P</i> value
Powerkheda (2013–14)	GYPP/GNPS/TGW/GZnC					
	M10256	G/A	1138343	UM	UM	1.42E–03
	M5415	G/A	978845	6B	60.43	1.50E–03
	M5016	T/C	1093566	3B	94.94	1.50E–03
	GNPS/TGW/GZnC					
	M13584	A/C	1062637	2B	95.62	1.30E–03
	M351	A/G	980758	3A	226.1	1.98E–03
	TGW/GBCC/GZnC					
	M15649	A/G	1009430	UM	UM	6.49E–04
	M11982	A/T	1084107	6B	67.32	7.30E–04
	Meerut (2014–15)	TNPP/GYPP/GNPS				
M11281		T/C	1139497	5D	247.21	1.84E–03
M11637		G/A	1116608	6B	88.27	3.50E–04
M11833		T/C	1055031	UM	UM	1.05E–04
M12137		G/A	1094422	2B	174.97	5.73E–04
M13595		C/G	1026516	UM	UM	1.22E–03
M14642		T/C	1188947	1B	18.27	7.74E–04
M14899		T/A	1068875	3B	23.96	1.59E–04
M3882		G/A	1126196	2B	132.6	6.50E–04
M4208		A/G	987806	5B	56.46	9.74E–04
M6254		T/C	1003907	UM	UM	2.78E–04
M7061		G/C	3034003	UM	UM	7.10E–05
GYPP/GNPS/TGW/GPC/GBCC/GZnC						
M5558		T/C	1079853	UM	UM	9.01E–04
GNPS/TGW/GPC/GBCC/GZnC						
M11941		A/G	1107464	3B	116.09	1.92E–03
M12005		G/A	996862	UM	UM	1.85E–03
M14642		T/C	1188947	1B	18.27	1.06E–03
TGW/GPC/GBCC/GZnC						
M11941		A/G	1107464	3B	116.09	7.62E–04
M12005		G/A	996862	UM	UM	9.15E–04
M15741		T/G	1109943	6A	189.77	1.16E–03
GPC/GBCC/GFeC/GZnC						
M10312		A/G	1203021	2B	148.07	1.62E–03
M11941		A/G	1107464	3B	116.09	1.05E–03
M12005		G/A	996862	UM	UM	1.79E–03
M13996		G/A	1228227	6A	46.4	1.79E–03
M15741		T/G	1109943	6A	189.77	5.21E–04
M184		A/G	1206653	7A	285.84	3.36E–04
M2678		C/G	1211630	7D	51.51	4.77E–04
M9672		A/G	1092948	UM	UM	1.46E–04

Table 4 Summary of most important QTLs

Traits	QTLs	SNP ID	SLST	MLMM	MTMM	Ch.	Pos. (cM)	Bonf. <i>P</i> value	References
GYPP	M15616	1080778	+	–	–	3D	280.77	0.004	Gupta et al. (2007)
	M14494	1694762	+	+	–	6B	65.03	0.01	Marza et al. (2006)
GPC/GFeC	M3205	1126883	–	+	+	5B	256.49	0.01	Tadesse et al. (2015)
	M802	1001832	–	+	+	5A	167.78	0.02	Wang et al. (2012)
	M10371	1001902	–	+	+	5B	55.28	0.04	Tadesse et al. (2015)
	M1673	1012837	–	+	+	5B	55.76	0.02	Tadesse et al. (2015)
	M4208	987806	–	+	+	5B	56.46	0.01	Tadesse et al. (2015)
	M15816	1135339	–	+	+	1B	142.36	0.01	–

Ch, chromosome; Pos, position; Bonf, Bonferroni; +, present; –, not present

marker-assisted selection (MAS). The criteria used for this purpose included the following: (1) lowest *P* value, (2) credibility (if identified by more than one method) and (3) whether or not the same marker was detected in earlier studies (including both interval mapping and GWAS) on the same chromosome. The MTAs and QTL detected at only one of the two locations were due to environmental effect ($G \times E$) and were treated as location specific (see Table S3, S4, S5 and S6). The following 6 MTAs fulfilled the last two criteria: M3205, M802, M10371, M1673 & M4208 (all five for Fe and protein content) and M14494 (for grain yield). One MTA (M15816; for Fe and protein content) was found to be novel, being reported for the first time and fulfilled the criteria (ii) and the remaining last MTA (M15616; for GYPP) fulfilled the first criteria. There were also markers, which were highly significant (lowest *P* value), some of these confirming the markers identified in earlier studies (Table 4 and Fig. 3). A solitary MTA (M5415 for GYPP using MLMM and for GYPP/GNPS/TGW/GZnC using mvLMM; Table 3) fulfilled the only single (ii) criterion, this MTA is not included in Table 4.

Discussion

An improvement in the content of major micronutrients in cereal grains, popularly described as biofortification, is an important area of research in order to address the problem of nutritional security (Neeraja et al. 2017). This requires an understanding of the genetic architecture that is associated with the content of each micronutrient and protein. The same is true for

yield-related traits. Also, since screening a large segregating population for contents of individual micronutrients in a breeding program can be labour intensive and cost-ineffective, marker-assisted selection (MAS) is a desirable option, if information about marker-trait associations (MTAs) is available (Loladze 2014; Myers et al. 2014).

Results of the present study support earlier reports that adequate variability for different micronutrients (Zn, Fe and β -carotene), protein content and yield-traits is available in the global wheat germplasm (Tiwari et al. 2009; Srinivasa et al. 2014; Li et al. 2016). However, these traits are quantitative in nature, making generation of knowledge about genetic architecture for improvement of these traits relatively difficult (Shi et al. 2008; Srinivasa et al. 2014). The variability in the concentration of Fe (24.50–44.30 ppm), Zn (17.75–49 ppm) and beta-carotene (0.5–6.5 mg/kg) in the material used in the present study is apparently adequate. However, there were clear location effects, with plants accumulating more Fe and Zn in wheat grain at Powerkheda than at Meerut (Table S2), although the experiments were conducted with uniform nutrient management practices at both the locations. In contrast, the micronutrient concentration in some land races of wheat is relatively high particularly in Afghan landraces (Fe: 55.14–122.2 ppm and Zn: 15.56–87.29 ppm; White and Broadley 2009; Ortiz-Monasterio et al. 2007; Manickavelu et al. 2017). A diverse collection of 132 wheat cultivars that are available at CIMMYT also has a wide range of micronutrient concentration (Fe: 28.8–56.50 ppm; Zn: 25.2–53.3 ppm; Graham et al. 1999; Xu et al. 2011). The variability of concentration of β -carotene was also adequate in the association

panel used, although concentration up to 14 mg/kg have been reported in einkorn wheat, which can be used as a good source for improvement of β -carotene concentration in the current high yielding wheat varieties (Hentschel et al. 2002; Hidalgo et al. 2006; Zhou et al. 2005; Leenhardt et al. 2006).

The present study is an effort firstly to evaluate the different methods available for GWAS, secondly to supplement the knowledge about the genetic architecture of micronutrient contents and some yield-related traits and finally to provide MTAs for these traits for MAS. The study was based on an association panel of 246 genotypes with large range of variation, in contrast to limited variation sampled in biparental mapping populations, which have been more frequently used for QTL analysis.

For conducting GWAS analysis, mixed linear model is generally used, which involves use of information about population structure (Q matrix) and familial relatedness/kinship (K matrix). In the present study, population structure was examined using 42 markers, located on 42 different arms of 21 wheat chromosomes. Such markers showing no linkage are generally used in GWAS studies to avoid the confounding due to linkage between markers. A similar strategy was also followed in two earlier independent studies, where 40 markers (Ogbonnaya et al. 2017) and 49 markers (Mulki et al. 2013) showing no linkage and covering the entire wheat genome were used for population structure. In order to find out the adequacy of 42 markers for population structure, we also conducted population structure analysis using 84 markers; the results of population structure based on 84 markers were no different from the results obtained using 42 markers suggesting that 42 markers were adequate for determining the population structure.

Similarly, familial relatedness or kinship analysis accounts for relationship among genotypes that are used for GWAS, since the results of GWAS may be confounded if familial relationship that may arise due to selection and/or genetic drift is not taken into account (Bernardo et al. 1996; Yu et al. 2006; Zhang et al. 2007). It has been shown that the use of kinship estimates in mixed models markedly reduces the number of false positive associations (Korte and Farlow 2013). In the present study, all available 17,937 SNP markers were used for estimating familial relatedness.

MTAs using different approaches (SLST, MLMM, MTMM & mvLMM)

Although single locus single trait (SLST) analysis has several limitations that have been widely discussed (Gupta et al. 2014; Jaiswal et al. 2016), this analysis was conducted in the present study mainly for the purpose of comparing its results with those of several other recently developed and improved approaches that were used in parallel during the present study. These new approaches included MLMM, MTMM and mvLMM that were proposed during recent years (Segura et al. 2012; Korte et al. 2012; Zhou and Stephens 2014; Furlotte and Eskin 2015; Jaiswal et al. 2016; Thoen et al. 2017) and were used in several recent studies including the present study. These three newer approaches for GWAS used in the present study take into consideration the genetic background and epistatic interaction (addressed in MLMM) and pleiotropy involving QTL/genes, which may each influence more than one traits (addressed in MTMM and mvLMM).

For identification of significant MTAs using different approaches, one of the major problem is multiple testing, which results in a large number of false positives. In order to overcome this problem, we used two measures, one is the use of a stringent P value ≤ 0.001 (instead of P value of 0.05 or 0.01), and the other was application of Bonferroni correction, which has been specially designed for this purpose. These criteria (including a P value ≤ 0.001) were also used in several earlier studies to identify significant MTAs in wheat (Ain et al. 2015; Wang et al. 2017) and other crops (Mogga et al. 2018). The problem of ≤ 0.001 has been further addressed through Bonferroni correction ($\alpha_{\text{GWAS}} \leq 0.05$), which is the most conservative approach (overall $P \leq 0.00005$) to deal with the problem of multiple testing, so that one does not expect any false positives after application of Bonferroni correction, although this may lead to some false negatives, which is not as serious a problem as the problem of false positives. In MLMM and mvLMM, which are the other two methods used in the present study Bonferroni correction is built-in within the software used, so that no Bonferroni correction is applied on MTAs obtained.

In the present study, 3835 MTAs were obtained using SLST and 990 MTAs were obtained using MTMM. However, only 136 MTAs from SLST and 28

MTAs from MTMM qualified after Bonferroni correction. MLM and mvLMM gave 587 and 33 MTAs respectively (Bonferroni corrections were inherent in these approaches). However, it is widely recognized that Bonferroni correction is a method, which leads to overcorrection giving many false negatives and is therefore a trade-off; consequently, many true and genuine MTAs must have been lost due to Bonferroni correction, in each of the approaches used during the present study. However, additional MTAs identified using epistatic interactions and a comparison with the results of earlier studies may partially address the problem of false negatives. In SLST results, there were a number of MTAs, which did not qualify Bonferroni, but were perhaps genuine MTAs, because they actually confirmed earlier reports (5 for GZnC to 7 for GPC), thus vindicating our conclusion that Bonferroni correction really leads to overcorrection, resulting in many false negatives. For instance, after application of Bonferroni correction to the SLST results, no MTAs were available for two traits (TNPP, GFeC), although six MTAs for TNPP were reported in two earlier studies, and eight MTAs for GFeC were reported in three earlier studies. The identification of 35 MTAs for TNPP and 33 for GFeC through MLM also suggests that SLST results are confounded due to genetic background that was ignored. Similar results with respect to Bonferroni were also available using MTMM involving 11 pairs of correlated traits, where Bonferroni correction reduced the number of MTAs from 1253 to 28 MTAs and that too for only 3 of the 11 pairs of correlated traits, once again suggesting overcorrection leading to false negatives due to Bonferroni correction. Similarly, a total of 97 MTAs were detected using epistatic interaction for 6 traits (GNPS; 18, TGW; 2, GYPP; 31, GPC; 16, GBCC; 17, GZnC; 13); these MTAs were not detected using MLM (Table S5 and S9). Similarly, 666 MTAs were also detected using epistasis for eight trait combinations (GYPP/GNPS/TGW/GZnC; 126, GNPS; 56, TGW/GBCC/GZnC; 42, GYPP/GNPS/TGW/GPC/GBCC/GZnC; 146, GNPS/TGW/GPC/GBCC/GZnC; 92, TGW/GPC/GBCC/GZnC; 58 and GPC/GBCC/GFeC/GZnC; 58). These MTAs were also not detected using mvLMM for the same traits combinations. Similarly, 666 additional MTAs which were also detected using epistasis but for different trait combinations were also perhaps false negatives and hence were not detected for the same trait combinations

when analysed using mvLMM. We can thus conclude that the commonly used SLST approach with or without Bonferroni correction is a very inefficient approach for GWAS. In SLST, MTMM and mvLMM, the results are also confounded due to the genetic background. This issue is addressed in MLM and therefore, MLM is certainly an improvement over SLST.

MTAs each controlling two or more than two correlated traits also deserve attention, although it is not easy to find out whether such multi-trait MTAs are due to pleiotropism or due to close linkage. Nevertheless, identification of such MTAs will help not only in the study of genetic architecture of correlated traits, but also in simultaneous improvement of more than two correlated traits using these MTAs for correlated traits for MAS.

Although MTMM gives useful information involving identification of common MTAs for pairs of correlated traits, its major limitation is that MTAs can be identified for only two correlated traits at a time, so that mvLMM approach was developed to overcome this limitation and to allow multi-trait analysis involving more than two correlated traits. MTAs detected using mvLMM involved 33 SNPs associated with eight combinations of traits (Table 3). The results of MTMM and mvLMM together suggest wide occurrence of pleiotropy or close linkage. The MTMM approach for identification of multi-trait MTAs using GWAS has also been used in two earlier studies in plants, one in wheat (Jaiswal et al. 2016) and the other in Arabidopsis (Thoen et al. 2017). It has, however, been utilized in several studies in animal systems including humans and cattle (Heid and Winkler 2017; Bolormaa et al. 2014; Crispim et al. 2015; Pausch et al. 2016). For mvLMM, the present study is perhaps the first study in plants, although it was earlier used in mouse (Zhou and Stephens 2014) and humans (Furlotte and Eskin 2015).

MTAs identified using more than one approaches, i.e. SLST and MLM or MLM and MTMM or SLST and MTMM can be accepted with a higher level of confidence. During the present study, 22 MTAs involving three traits (GNPS, TGW and GYPP) were common between SLST and MLM, 12 MTAs involving three pairs of traits (TGW/GNPS, GPC/GZnC and GYPP/TGW) were common between MLM and MTMM (Table S8). These MTAs can be used for MAS with a higher level of confidence.

Some novel MTAs

A large number of MTAs identified using three different methods of analysis during the present study were novel. As many as 222 novel MTAs were identified using SLST and MLM taken together; of these 70 MTAs were for nutritional traits identified using MLM only (none through SLST) and 152 for yield-related traits were identified using SLST and MLM (Tables S4, S5). This suggests that many MTAs remain to be identified. These MTAs are new additions to the already available MTAs for a variety of traits in wheat (Tables S4, S5 and S7).

Epistasis is often overlooked in GWAS

Knowledge of epistasis is also necessary to understand the genetic architecture of a trait (Boone et al. 2007; Phillips 2008; Jaiswal et al. 2016). If this information is not available, it may lead to under-utilization of genomic information for crop improvement (Wang et al. 2011; Jaiswal et al. 2016). However, such epistatic interactions have been seldom examined during GWAS despite their wide-spread occurrence (Reif et al. 2011; Kao et al. 1999; Langer et al. 2014; Jaiswal et al. 2016). In the present study, 73 important epistatic interactions were detected for six traits. Six of these epistatic interactions (one for each of 6 different traits) were highly significant (on the basis of lowest *P* value) and included the following: (1) M3424-7A × M15107-3B) for GNPs, (2) (M519-6B × M71-6B) for TGW, (3) (M10613-2A × M12763-3B) for GYPP, (4) (M61-5A × M3743-7B) for GPC, (5) (M2908-2A × M5863-3B) for GBCC and (6) GZnC (M8512-7A × M9308-1D) for GZnC (Table S9).

Epistatic interactions in wheat using GWAS have been reported in several earlier studies for flower time (Reif et al. 2011; Langer et al. 2014) and stem rust resistance (Yu et al. 2011). In a recent study from our own laboratory also, 63 epistatic interactions involving 13 different traits in wheat were reported using GWAS (Jaiswal et al. 2016). Interval mapping involving biparental populations have also been used for detection of epistatic interactions (Li et al. 2011; Xu et al. 2012). Pairs of markers involved in each of the reported epistatic interactions including those detected in the present study may be useful for molecular breeding in wheat.

G × E interactions

In analysis of variance (ANOVA) G × E interactions were found to be significant for all traits except GPC (Table 2). However, specific QTL × E interaction effects were not estimated. However, the location and year effects were apparent from the availability of QTL that were location/year-specific. Such location-specific QTLs can be used for breeding varieties for specific location. Also since one of the two locations (*i.e.* Powerkheda) used in the present study is known to experience drought and high temperature, the QTL identified at Powerkheda (absent at Meerut) may be useful, while breeding for tolerance against abiotic stresses (drought and high temperature).

Markers for marker-assisted selection (MAS): biofortification

For the purpose of selection of superior individual plants for enhanced micronutrient content in a segregating population is time-consuming and labour-intensive. Keeping this in view, the goal of the present study was to find markers to be used for MAS or MARS for improvement of not only the nutrition traits but also the yield-related traits. Although a large number of MTAs were detected in the present study, nine major MTAs for three traits were considered to be the most important (using criteria mentioned in Results). Three of these nine markers can be used for GYPP, and other six markers can be used each for selecting two traits together (GPC and GFeC; see Tables 3, 4 for details). For other traits, M14605 is important for GZnC; M167 and M3471 for GBCC; M401 and M4501 for TGW; M13270 for GNPS and M15558 for TNPP (Table S4, S5 and Fig. 3).

Conclusions

Molecular breeding involving MAS has now become a component of conventional plant breeding, particularly for improvement of complex quantitative traits (QTs). A prerequisite for molecular breeding is the availability of markers associated with the targeted QTs. The two approaches, which have been extensively used for this purpose, include interval mapping and association mapping (GWAS), both having their merits and demerits. While markers detected through

interval mapping have already been put to use in many crops including wheat, we don't have many documented examples of the use of the results of GWAS for this purpose. One classical example is the improvement of provitamin A in maize (Suwarno et al. 2015). However, interval mapping has a serious limitation of using limited genetic variability, available between two parents of a biparental population. This limitation is addressed in GWAS through the utilization of almost complete genetic variability, although it does have other limitations, which are being addressed (Gupta et al. 2014). The present GWA study made use of an association panel including diverse global spring wheat germplasm, with a particular emphasis on the study of genetics of the contents of micronutrients like Fe, Zn and β -carotene. A large number of MTAs were detected, which can be used by wheat breeders, after due prioritization and validation. The present study is one of the first few studies, where four different approaches were used for GWAS. The analysis has been undertaken, resulting in more reliable and useful information. The results of the present study also demonstrated that GWAS is a powerful tool for genetic dissection of complex traits, if newer approaches are utilized. The results obtained in this study should prove useful not only in molecular breeding, but also for further studies thus giving direction to future research in the field of association mapping.

Acknowledgements JK received financial support in the form of a Senior Research Fellowship (SRF) from Indian Council of Medical Research (ICMR). Head, Genetics and Plant Breeding, CCS University, Meerut provided the facilities and Prof. V.K. Mishra, Banaras Hindu University Varanasi, India provided facility for zinc and iron estimations. CIMMYT gene bank is sincerely acknowledged for providing seed material and genotypic data for the present research.

References

Ain QU, Rasheed A, Anwar A, Mahmood T, Imtiaz M, Mahmood T et al (2015) Genome-wide association for grain yield under rainfed conditions in historical wheat cultivars from Pakistan. *Front Plant Sci* 6:743

Allard RW (1999) Principles of plant breeding. Wiley, New York

Bolormaa S, Pryce JE, Reverter A et al (2014) A multi-trait, meta-analysis for detecting pleiotropic polymorphisms for stature, fatness and reproduction in beef cattle. *PLoS Genet*

10:e1004198. <https://doi.org/10.1371/journal.pgen.1004198>

Boone C, Bussey H, Andrews BJ (2007) Exploring genetic interactions and networks with yeast. *Nat Rev Genet* 8:437–449

Breseghele F, Sorrells ME (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172:1165–1177. <https://doi.org/10.1534/genetics.105.044586>

Chattha MU, Hassan MU, Khan I et al (2017) Biofortification of wheat cultivars to combat zinc deficiency. *Front Plant Sci* 8:281

Colasuonno P, Lozito ML, Marcotuli I et al (2017) The carotenoid biosynthetic and catabolic genes in wheat and their association with yellow pigments. *BMC Genomics* 18:122. <https://doi.org/10.1186/s12864-016-3395-6>

Crispim AC, Kelly MJ, Guimarães SEF et al (2015) Multi-trait GWAS and new candidate genes annotation for growth curve parameters in brahman cattle. *PLoS One* 10:e0139906. <https://doi.org/10.1371/journal.pone.0139906>

Diapari A, Bett K, Deokar A, Warkentin TD, Tar'an BM (2014) Genetic diversity and association mapping of iron and zinc concentrations in chickpea (*Cicer arietinum* L.). *Genome* 57:459–468. <https://doi.org/10.1139/gen-2014-0108>

Diapari M, Sindhu A, Warkentin TD et al (2015) Population structure and marker-trait association studies of iron, zinc and selenium concentrations in seed of field pea (*Pisum sativum* L.). *Mol Breed* 35:30. <https://doi.org/10.1007/s11032-015-0252-2>

Esuma W, Herselman L, Labuschagne MT et al (2016) Genome-wide association mapping of provitamin A carotenoid content in cassava. *Euphytica* 212:97–110. <https://doi.org/10.1007/s10681-016-1772-5>

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>

Furlotte NA, Eskin E (2015) Efficient multiple trait association and estimation of genetic correlation using the matrix-variate linear mixed-model. *Genetics* 200:59–68. <https://doi.org/10.1534/genetics.114.171447>

González JR, Armengol L, Solé X et al (2007) SNPpass: an R package to perform whole genome association studies. *Bioinformatics* 23:644–645. <https://doi.org/10.1093/bioinformatics/btm025>

Gorafi YSA, Ishii T, Kim JS et al (2016) Genetic variation and association mapping of grain iron and zinc contents in synthetic hexaploid wheat germplasm. *Plant Genet Resour* 1:9. <https://doi.org/10.1017/S1479262116000265>

Graham R, Senadhira D, Beebe S et al (1999) Breeding for micronutrient density in edible portions of staple food crops: conventional approaches. *Field Crop Res* 60:57–80. [https://doi.org/10.1016/S0378-4290\(98\)00133-6](https://doi.org/10.1016/S0378-4290(98)00133-6)

Gupta PK, Harindra SB, Pawan LK et al (2007) QTL analysis for some quantitative traits in bread wheat. *J Zhejiang Univ Sci B* 8:807–814. <https://doi.org/10.1631/jzus.2007.B0807>

Gupta PK, Kulwal PL, Jaiswal V (2014) Association mapping in crop plants: opportunities and challenges. *Adv Genet*. <https://doi.org/10.1016/b978-0-12-800271-1.00002-0>

- Heid IM, Winkler TW (2017) A multitrait GWAS sheds light on insulin resistance. *Nat Genet* 49:7–8
- Hentschel V, Kranl K, Hollmann J et al (2002) Spectrophotometric determination of yellow pigment content and evaluation of carotenoids by high-performance liquid chromatography in durum wheat grain. *J Agric Food Chem* 50:6663–6668. <https://doi.org/10.1021/jf025701p>
- Hidalgo A, Brandolini A, Pompei C, Piscozzi R (2006) Carotenoids and tocots of einkorn wheat (*Triticum monococcum* ssp. *monococcum* L.). *J Cereal Sci* 44:182–193. <https://doi.org/10.1016/j.jcs.2006.06.002>
- Jaiswal V, Gahlaut V, Meher PK et al (2016) Genome wide single locus single trait, multi-locus and multi-trait association mapping for some important agronomic traits in common wheat (*T. aestivum* L.). *PLoS One* 11:e0159343. <https://doi.org/10.1371/journal.pone.0159343>
- Jones ES, Sullivan H, Bhatramakki D, Smith JSC (2007) A comparison of simple sequence repeat and single nucleotide polymorphism marker technologies for the genotypic analysis of maize (*Zea mays* L.). *Theor Appl Genet* 115:361–371. <https://doi.org/10.1007/s00122-007-0570-9>
- Joshi AK, Crossa J, Arun B et al (2010) Genotype × environment interaction for zinc and iron concentration of wheat grain in eastern Gangetic plains of India. *Field Crop Res* 116:268–277. <https://doi.org/10.1016/j.fcr.2010.01.004>
- Kao CH, Zeng ZB, Teasdale RD (1999) Multiple interval mapping for quantitative trait loci. *Genetics* 152:1203–1216
- Korte A, Vilhjálmsson BJ, Segura V et al (2012) A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nat Genet* 44:1066–1071. <https://doi.org/10.1038/ng.2376>
- Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant methods* 9:29
- Langer SM, Longin CFH, Würschum T (2014) Flowering time control in European winter wheat. *Front Plant Sci* 5:537. <https://doi.org/10.3389/fpls.2014.00537>
- Leenhardt F, Lyan B, Rock E et al (2006) Genetic variability of carotenoid concentration, and lipoxygenase and peroxidase activities among cultivated wheat species and bread wheat varieties. *Eur J Agron* 25:170–176. <https://doi.org/10.1016/j.eja.2006.04.010>
- Leplat F, Pedas PR, Rasmussen SK, Husted S (2016) Identification of manganese efficiency candidate genes in winter barley (*Hordeum vulgare*) using genome wide association mapping. *BMC Genomics* 17:805. <https://doi.org/10.1186/s12864-016-3129-9>
- Li WH, Wei LIU, Li LIU, You MS, Liu GT, Li BY (2011) QTL mapping for wheat flour color with additive, epistatic, and QTL × environmental interaction effects. *Agric Sci China* 10:651–660. [https://doi.org/10.1016/s1671-2927\(11\)60047-3](https://doi.org/10.1016/s1671-2927(11)60047-3)
- Li G, Xu X, Bai G, Carver BF, Hunger R, Bonman JM et al (2016) Genome-wide association mapping reveals novel qtl for seedling leaf rust resistance in a worldwide collection of winter wheat. *Plant Genome* 9:0
- Loladze I (2014) Hidden shift of the ionome of plants exposed to elevated CO2 depletes minerals at the base of human nutrition. *Elife* 3:e02245. <https://doi.org/10.7554/eLife.02245>
- Manickavelu A, Hattori T, Yamaoka S (2017) Genetic nature of elemental contents in wheat grains and its genomic prediction: toward the effective use of wheat landraces from Afghanistan. *PLoS One* 12:e0169416
- Marza F, Bai GH, Carver BF, Zhou WC (2006) Quantitative trait loci for yield and related traits in the wheat population Ning7840 × Clark. *Theor Appl Genet* 112:688–698. <https://doi.org/10.1007/s00122-005-0172-3>
- Mogga M, Sibiyi J, Shimelis H, Lamo J, Yao N (2018) Diversity analysis and genomewide association studies of grain shape and eating quality traits in rice (*Oryza sativa* L.) using DArT markers. *PLoS One* 13:e0198012
- Mulki MA, Jighly A, Ye GY, Emebiri LC, Moody D, Ansari O, Ogbonnaya FC (2013) Association mapping for soilborne pathogen resistance in synthetic hexaploid wheat. *Mol Breed* 31:299–311
- Myers SS, Zanobetti A, Kloog I et al (2014) Increasing CO₂ threatens human nutrition. *Nature* 510:139–142. <https://doi.org/10.1038/nature13179>
- Neeraja CN, Babu VR, Ram S, Hossain F et al (2017) Biofortification in cereals: progress and prospects. *Curr Sci* 113:6
- Norton GJ, Douglas A, Lahner B et al (2014) Genome wide association mapping of grain arsenic, copper, molybdenum and zinc in rice (*Oryza sativa* L.) grown at four international field sites. *PLoS One* 9:e89685. <https://doi.org/10.1371/journal.pone.0089685>
- Ogbonnaya FC, Rasheed A, Okechukwu EC, Jighly A et al (2017) Genome-wide association study for agronomic and physiological traits in spring wheat evaluated in a range of heat prone environments. *Theor Appl Genet* 130:1819–1835
- Ortiz-Monasterio JJ, Palacios-Rojas N, Meng E et al (2007) Enhancing the mineral and vitamin content of wheat and maize through plant breeding. *J Cereal Sci* 46:293–307. <https://doi.org/10.1016/j.jcs.2007.06.005>
- Paltridge NG, Palmer LJ, Milham PJ et al (2012) Energy-dispersive X-ray fluorescence analysis of zinc and iron concentration in rice and pearl millet grain. *Plant Soil* 361:251–260. <https://doi.org/10.1007/s11104-011-1104-4>
- Pausch H, Emmerling R, Schwarzenbacher H, Fries R (2016) A multi-trait meta-analysis with imputed sequence variants reveals twelve QTL for mammary gland morphology in Fleckvieh cattle. *Genet Sel Evol* 48:14. <https://doi.org/10.1186/s12711-016-0190-4>
- Peleg Z, Cakmak I, Ozturk L et al (2009) Quantitative trait loci conferring grain mineral nutrient concentrations in durum wheat × wild emmer wheat RIL population. *Theor Appl Genet* 119:353–369. <https://doi.org/10.1007/s00122-009-1044-z>
- Phillips PC (2008) Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet* 9:855–867. <https://doi.org/10.1038/nrg2452>
- Pozniak CJ, Knox RE, Clarke FR, Clarke JM (2007) Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. *Theor Appl Genet* 114:525–537. <https://doi.org/10.1007/s00122-006-0453-5>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959. <https://doi.org/10.1111/j.1471-8286.2007.01758.x>

- Rabbi IY, Udoh LI, Wolfe M et al (2017) Genome-wide association mapping of correlated traits in cassava: dry matter and total carotenoid content. *Plant Genome* 03:18. <https://doi.org/10.3835/plantgenome2016.09.0094>
- Reif JC, Maurer HP, Korzun V et al (2011) Mapping QTLs with main and epistatic effects underlying grain yield and heading time in soft winter wheat. *Theor Appl Genet* 123:283–292. <https://doi.org/10.1007/s00122-011-1583-y>
- Roshanzamir H, Kordenaeej A, Bostani A (2013) Mapping QTLs related to Zn and Fe concentrations in bread wheat (*Triticum aestivum* L.) grain using microsatellite markers. *IJGPB* 2:10–17
- Segura V, Vilhjálmsson BJ, Platt A et al (2012) An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. *Nat Genet* 44:825–830. <https://doi.org/10.1038/ng.2314>
- Sehgal D, Autrique E, Singh R et al (2017) Identification of genomic regions for grain yield and yield stability and their epistatic interactions. *Sci Rep* 7:41578. <https://doi.org/10.1038/srep41578>
- Sharma P, Sheikh I, Kumar S, Verma SK et al (2018) Precise transfers of genes for high grain iron and zinc from wheat-*Aegilops* substitution lines into wheat through pollen irradiation. *Mol Breed* 38:81
- Shi R, Li H, Tong Y et al (2008) Identification of quantitative trait locus of zinc and phosphorus density in wheat (*Triticum aestivum* L.) grain. *Plant Soil* 306:95–104
- Srinivasa J, Arun B, Mishra VK et al (2014) Zinc and iron concentration QTL mapped in a *Triticum spelta* × *T. Aestivum* cross. *Theor Appl Genet* 127:1643–1651. <https://doi.org/10.1007/s00122-014-2327-6>
- Suwarno WB, Pixley KV, Palacios-Rojas N et al (2015) Genome-wide association analysis reveals new targets for carotenoid biofortification in maize. *Theor Appl Genet* 128:851–864. <https://doi.org/10.1007/s00122-015-2475-3>
- Tadesse W, Ogbonnaya FC, Jighly A et al (2015) Genome-wide association mapping of yield and grain quality traits in winter wheat genotypes. *PLoS One* 10:e0141339. <https://doi.org/10.1371/journal.pone.0141339>
- Thoen MPM, Davila Olivas NH, Kloth KJ et al (2017) Genetic architecture of plant stress resistance: multi-trait genome-wide association mapping. *New Phytol* 213:1346–1362. <https://doi.org/10.1111/nph.14220>
- Tiwari VK, Rawat N, Chhuneja P et al (2009) Mapping of quantitative trait loci for grain iron and zinc concentration in diploid a genome wheat. *J Hered* 100:771–776. <https://doi.org/10.1093/jhered/esp030>
- Tiwari C, Wallwork H, Arun B et al (2016) Molecular mapping of quantitative trait loci for zinc, iron and protein content in the grains of hexaploid wheat. *Euphytica* 207:563–570. <https://doi.org/10.1007/s10681-015-1544-7>
- Velu G, Singh RP, Huerta-Espino J et al (2012) Performance of biofortified spring wheat genotypes in target environments for grain zinc and iron concentrations. *Field Crops Res* 137:261–267. <https://doi.org/10.1016/j.fcr.2012.07.018>
- Wang D, Eskridge KM, Crossa J (2011) Identifying QTLs and epistasis in structured plant populations using adaptive mixed LASSO. *J Agric Biol Environ Stat* 16:170–184. <https://doi.org/10.1007/s13253-010-0046-2>
- Wang L, Cui F, Wang J et al (2012) Conditional QTL mapping of protein content in wheat with respect to grain yield and its components. *J Genet* 91:303–312. <https://doi.org/10.1007/s12041-012-0190-2>
- Wang X, Pang Y, Zhang J, Wu Z, Chen K, Ali J, Ye G, Xu J, Li Z (2017) Genome-wide and gene-based association mapping for rice eating and cooking characteristics and protein content. *Sci Rep* 7:17203
- Welch RM, Graham RD (2004) Breeding for micronutrients in staple food crops from a human nutrition perspective. *J Exp Bot* 55:353–364
- White PJ, Broadley MR (2009) Biofortification of crops with seven mineral elements often lacking in human diets - Iron, zinc, copper, calcium, magnesium, selenium and iodine. *New Phytol* 182:49–84
- World Health Organization (2016) World health statistics—monitoring health for the SDGs. *World Heal Organ*. <https://doi.org/10.1017/cbo9781107415324.004>
- Xu Y, An D, Li H, Xu H (2011) Review: breeding wheat for enhanced micronutrients. *Can J Plant Sci* 91:231–237. <https://doi.org/10.4141/CJPS10117>
- Xu Y, An D, Liu D et al (2012) Molecular mapping of QTLs for grain zinc, iron and protein concentration of wheat across two environments. *Field Crops Res* 138:57–62. <https://doi.org/10.1016/j.fcr.2012.09.017>
- Yu J, Pressoir G, Briggs WH et al (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* 38:203–208. <https://doi.org/10.1038/ng1702>
- Yu LX, Lorenz A, Rutkoski J et al (2011) Association mapping and gene-gene interaction for stem rust resistance in CIMMYT spring wheat germplasm. *Theor Appl Genet* 123:1257–1268. <https://doi.org/10.1007/s00122-011-1664-y>
- Zhang Y, Ni Z, Yao Y, Nie X, Sun Q (2007) Gibberellins and heterosis of plant height in wheat (*Triticum aestivum* L.). *BMC Genet* 8:40
- Zhao Y, Sun HY, Wang YY et al (2013) QTL mapping for the color, carotenoids and polyphenol oxidase activity of flour in recombinant inbred lines of wheat. *Aust J Crop Sci* 7:328–337
- Zhou X, Stephens M (2012) Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* 44:821–824. <https://doi.org/10.1038/ng.2310>
- Zhou X, Stephens M (2014) Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat Methods* 11:407–409. <https://doi.org/10.1038/nmeth.2848>
- Zhou K, Yin JJ, Yu L (2005) Phenolic acid, tocopherol and carotenoid compositions, and antioxidant functions of hard red winter wheat bran. *J Agric Food Chem* 53:3916–3922. <https://doi.org/10.1021/jf050117c>