REVIEW ARTICLE

Genomewide association study for seeding emergence and tiller number using SNP markers in an elite winter wheat population

GUANG FENG CHEN[1,2], RU GANG WU[3], DONG MEI LI[2], HAI XIA YU[1], ZHIYING DENG[1] and JI CHUN TIAN[1]*

[1]*State Key Laboratory of Crop Biology, Group of Quality Wheat Breeding of Shandong Agricultural University, Tai'an 271018, People's Republic of China*
[2]*College of Ecology and Garden Architecture, Dezhou University, Dezhou 253023, People's Republic of China*
[3]*Dezhou Academy of Agricultural Sciences, Dezhou 253015, People's Republic of China*

Abstract

Seeding emergence and tiller number are the most important traits for wheat (*Triticum aestivum* L.) yield, but the inheritance of seeding emergence and tillering is poorly understood. We conducted a genomewide association study focussing on seeding emergence and tiller number at different growth stages with a panel of 205 elite winter wheat accessions. The population was genotyped with a high-density Illumina iSelect 90K SNPs assay. A total of 31 loci were found to be associated with seeding emergence rate (SER) and tiller number in different growth stages. Loci distributed among 12 chromosomes accounted for 5.35 to 11.33% of the observed phenotypic variation. With this information, 10 stable SNPs were identified for eventual development of cleaved amplified polymorphic sequence markers for SER and tiller number in different growth stages. Additionally, a set of elite alleles were identified, such as *Ra_c14761_1348-T*, which may increase SER by 13.35%, and *Excalibur_c11045_236-A* and *BobWhite_c8436_391-T*, which may increase the rate of available tillering by 14.78 and 8.47%, respectively. These results should provide valuable information for marker-assisted selection and parental selection in wheat breeding programmes.

[Chen G. F., Wu R. G., Li D. M., Yu H. X., Deng Z. and Tian J. C. 2017 Genomewide association study for seeding emergence and tiller number using SNP markers in an elite winter wheat population. *J. Genet.* **96**, 177–186]

Introduction

Wheat (*Triticum aestivum* L.) is one of the most important food crops worldwide, feeding about 40% of the world's population (Gupta *et al.* 2008). Accordingly, selection for high grain yield is an important goal of wheat breeding programmes. Seeding emergence rate (SER) and tillering are the most important agronomic traits affecting biomass and grain yield potential in wheat. SER is an important target for judging the quality of seeds, and it is a reflection of seedling vigour and growth that indirectly affects yields. Several studies on the relationship between high seed vigour and yield improvement have been reported in rice (Ismail *et al.* 2012), maize (Cervantes-Ortiz *et al.* 2007) and wheat (Botwright *et al.* 2002; Ludwig and Asseng 2010). Tillering is an important agronomic character in wheat and tiller number changes over the course of development. For winter wheat, tiller numbers increase throughout the autumn until tiller growth stops when the mean temperature reaches 0°C, during this point the maximum tillering of prewinter (MTW) is reached. In the following spring, again a great number of new tillers are formed when the mean temperature reaches 10°C. At the joint stage, the number of tillers overall reaches its maximum, which is called maximum tillering in spring (MTS). However, only a proportion of tillers survive to contribute to the ultimate number of tillers, i.e., the effective tillering at harvest (ETH) depends on initial tiller growth and the ultimate tiller survival (Hodges 1991; Evers *et al.* 2006). MTW, MTS and ETH are important traits that determined the shoot architecture across different wheat growth stages (Li *et al.* 2010). The architecture of the shoot system affects plant's light harvesting potential, the synchrony of flowering and seed set and, ultimately, the reproductive success of a plant (Yang *et al.* 2006; Kuraparthy *et al.* 2007). The rate of the available tillering (RAT) is a derived trait of ETH and MTS, which reflects potential wheat yields. Hence, the genetic elucidation of seedling emergence and tiller number at different stages of growth has become a focus of wheat genetics and breeding research programmes.

*For correspondence. E-mail: jctian9666@163.com.

Keywords. winter wheat; genomewide association study; single-nucleotide polymorphism; seeding emergence; tiller number.

At present, quantitative trait loci (QTL) mapping and genomewide association studies (GWAS) are the two main approaches to dissect the genetic bases of complex traits (Risch and Merikangas 1996). Many wheat QTL mapping studies have been performed on traits associated with seed vigour and tillering (Rebetzke *et al.* 2005; Spielmeyer *et al.* 2007; Landjeva *et al.* 2010; Li *et al.* 2010). However, in all these studies, although QTL mapping was successful in detecting QTL, the genetic variation explored in the mapping population is restricted to the genomes of only the two parents, and the genetic markers for detected QTL are not necessarily transferable to other populations.

As a complement to QTL mapping, GWAS is a high-resolution and cost-effective method for genetic mapping that use existing germplasm (such as landraces, elite cultivars and advanced breeding lines) and is based on linkage disequilibrium (LD) (Zhu *et al.* 2008; Waugh *et al.* 2009; Ingvarsson and Nathaniel 2011). As genomics have rapidly developed and genotyping costs have dramatically decreased, association mapping has rapidly become a promising approach for the genetic dissection of complex agronomic and crop quality traits in cereal crops such as rice, barley and maize (Yu and Buckler 2006; Cockram *et al.* 2010; Huang *et al.* 2012).

To achieve a high resolution association analyses, a large number of molecular marker sets are preferred for association mapping. Single-nucleotide polymorphisms (SNPs) are the most promising molecular markers, and they have two major advantages over other molecular markers: (i) SNPs are the most abundant form of genetic variation within genomes (Zhu *et al.* 2003); and (ii) a wide array of technologies have now been developed for high throughput genotyping (Fan *et al.* 2006). These factors combined make SNPs ideal for GWAS using large samples and high-density markers, which greatly improves the resolution of association mapping. At present, GWAS with SNPs have been widely used to characterize the genetics of many animals and plants, such as humans (Eric *et al.* 2011), *Drosophila* (Gerrard *et al.* 2013), rice (Huang *et al.* 2010) and maize (Wilson *et al.* 2004). However, GWAS using SNPs in wheat is a challenge because of wheat's complex genomic architecture and incomplete genome sequence (Sukumaran and Yu 2014). While few GWAS in wheat have identified SNP markers for grain yield (Sukumaran *et al.* 2014) and heading date (Zanke *et al.* 2014), a GWAS examining the association of SER and tiller number with SNP markers has not yet been reported.

As such, the present study is a GWAS examining SER and tiller number in different growth stages and the association of these traits with 24,355 SNPs genotyped using the 90K Illumina iSelect array (Wang *et al.* 2014) in a diverse population of winter wheat. The objectives of this study were to identify molecular markers associated with SER and tiller number in different growth stages and explore a set of elite alleles influencing SER and RAT. The ultimate aim of this research was to facilitate molecular breeding and identify strategic combinations of traits in winter wheat.

Materials and methods

Plant material

To ensure a broadly representing sampling of wheat accessions, 182 elite varieties, developed largely from 1980s to 2010s, and 150 breeding lines including two varieties introduced from countries other than China were selected. A representative collection of 205 varieties and lines was chosen by discarding accessions with similar pedigrees and phenotypes related to yield. The remaining 205 varieties were used for the present GWAS. The accessions comprised 77 bred varieties, 55 founder parents and 73 breeder's lines from 10 provinces that are major winter wheat regions in China, as well as two additional lines, used as founder parents from Mexico and France (see table 1 of electronic supplementary material at http://www.ias.ac.in/jgenet/). This association mapping panel expressed substantial phenotypic variation in the main wheat agronomic characters including tiller number at different stages of growth.

Field trials and phenotypic trait evaluation

Field trials were conducted in 2013–2014 and 2014–2015 growing seasons at the experimental fields of Shandong Agricultural University, Tai'an (116°36′E, 36°57′N) and in 2014–2015 growing season at the Dezhou Institute of Agricultural Sciences, Dezhou (116°29′E, 37°45′N). The experiment followed a completely randomized block design with two replications at each location. Each plot consisted of three rows of wheat plants: two rows grown for the purpose of evaluating SER with a distance of 2.5 cm between plants, and one row for evaluating MTW, MTS and ETH with a distance of 10 cm between plants. Each row was 2 m in length, and the distance between adjacent rows was 25 cm. The field management followed the standard local practices. RAT was calculated as RAT = ETH/MTS. Trait measurements (i.e. MTW, MTS, ETH and RAT) were estimated from averages of five plants within each plot for subsequent statistical analyses.

Analysis of variance (ANOVA) and correlations among phenotypic traits were performed using the statistical software SPSS ver. 17.0 (SPSS, Chicago, USA). Heritability (h^2) was calculated as

$$h_B^2 = \sigma_g^2/(\sigma_g^2 + \sigma_e^2),$$

where σ_g^2 is between group variances and σ_e^2 is within group variances (Li *et al.* 2009). The estimates of σ_g^2 and σ_e^2 were obtained from variance estimates included in the ANOVA.

DNA extraction and SNP genotyping

DNA was extracted from young leaf tissues of each variety using the protocol recommended by Triticarte (http://www.triticarte.com.au). DNA quality was checked by electrophoresis on 0.8% agarose gels, and DNA concentration was determined with a NanoDropND-1000 UV-Vis spectrophotometer (NanoDrop Technologies, Wilmington, USA).

SNP genotyping was performed at the University of California, Davis Genome Center using the novel 90K iSelect wheat chip (Wang *et al.* 2014), which consists of 81,587 SNP loci distributed across all 21 wheat chromosomes. The SNP data were clustered and called automatically using the Illumina BeadStudio genotyping software (Illumina, San Diego, USA) and were filtered to remove alleles with a detection rate less than 0.8 and/or a minor allele frequency (MAF) less than 0.05 (Li *et al.* 2014a).

Genetic diversity and composite genetic map

Allele number, MAF and the polymorphism information content (PIC) value for each locus were calculated with the Power Marker program v3.25 (Liu and Muse 2005). A composite genetic map for the association panel was built to GWAS based on the published genetic map (Wang *et al.* 2014).

Population structure

Population structure analysis for the association panel was performed using the STRUCTURE v2.2 software (Pritchard *et al.* 2000) with 3297 SNP markers positioned at least 2 cM apart in the genome. STRUCTURE was run 10 times for each subpopulation value K, ranging from 1 to 15, using the admixture model with 100,000 iterations for burn-in and 50,000 iterations for the analysis itself. The K value was determined by using estimates of ln P (D) from the STRUCTURE output and an ad hoc statistic ΔK based on the rate of change in ln P (D) between successive K values (Evanno *et al.* 2005).

Genomewide association analysis

Significant marker-trait associations (MTAs) were identified using a general linear model (GLM) in TASSEL 3.0 (http://www.maizegenetics.net/). The population structure (Q matrix) was inferred using STRUCTUTE 2.2. The P value determined whether a QTL was associated with the marker or not and the R^2 value was used to evaluate the magnitude of the QTL effects. The false discovery rate (FDR) procedures were used to adjust for multiple testing from the number of SNP loci detected. The individual P values from GWAS were sorted from smallest to largest, and then were multiplied by different coefficients (the number of SNP loci tested in the GWAS / the rank values of individual P values), i.e. the corrected P values of the SNP loci detected. SNPs with corrected P values ≤ 0.05 were considered to be in significant association with phenotypic traits.

Results

Phenotypic performance

The measurements of the five traits occurring in the three growth stages for the association panel are listed in table 1. Extensive phenotypic variation for these five traits was observed in three environments (i.e. two growing seasons, the second of which was split between two locations) among 205 winter wheat accessions. Broad sense heritability of the five traits ranged from 61.78 (MTW) to 74.5% (SER), indicating that both genetic and environmental factors played roles in the expression of these measured traits. Both absolute values of skewness and kurtosis were less than 1.0, which indicates that these quantitative traits were determined by many alleles of small individual effects.

The correlations between SER and tiller number are shown in table 2. SER exhibited a significantly positive correlation with MTS, ETH and RAT. MTS exhibited a significantly positive correlation with ETH, but a significantly negative correlation with RAT. Additionally, ETH exhibited a significantly positive correlation with RAT.

Table 1. Phenotypic summary of the five traits in the association panel of three environments.

Trait[a]	Env.[b]	Min.	Max.	Range	Mean	SD	Skewness	Kurtosis	H^2 (%)
SER	E1	0.16	0.90	0.74	0.53	0.15	0.08	−0.56	74.50
	E2	0.46	0.97	0.51	0.67	0.09	0.44	0.01	
	E3	0.28	0.93	0.65	0.63	0.11	−0.04	−0.13	
MTW	E1	2.31	8.76	6.45	5.25	1.27	0.42	0.32	61.78
	E2	3.13	9.69	6.65	4.98	1.45	0.17	0.12	
	E3	3.82	13.21	9.39	5.47	3.15	0.15	0.07	
MTS	E1	3.50	15.30	11.80	8.26	2.21	0.53	0.03	70.56
	E2	2.71	13.41	10.70	8.53	1.88	0.30	0.15	
	E3	5.45	15.81	10.36	10.30	2.33	0.17	−0.32	
ETH	E1	1.90	8.40	6.50	4.71	1.31	0.19	0.33	73.79
	E2	1.78	6.74	4.96	4.13	1.02	0.14	−0.46	
	E3	2.75	8.01	5.26	5.06	1.12	0.19	−0.12	
RAT	E1	0.31	0.84	0.53	0.57	0.09	0.21	0.14	72.51
	E2	0.25	0.88	0.63	0.49	0.09	0.42	0.33	
	E3	0.30	0.95	0.65	0.51	0.08	0.17	0.38	

[a]SER, seeding emergence rate; MTW, maximum tillering of pre-winter; MTS, maximum tillering in spring; ETH, effective tillering in harvest; RAT, rate of the available tillering.
[b]E1, 2013 Tai'an; E2, 2014 Tai'an; E3, 2014 Dezhou.

Genetic diversity and the composite genetic map

A total of 32,432 of 81,587 SNPs assayed with the functional iSelect bead chip were polymorphic among the 205 winter wheat accessions, and the polymorphism rate was 39.75%. After SNPs with calling detection frequencies exceeding 0.8 and MAF less than 0.05 were excluded, 24,355 SNPs that mapped to the 21 wheat chromosomes were used for the MTA analysis.

A total of 48,710 alleles were detected (i.e. two alleles per SNP locus). About 59.9% of the SNPs (14,597/24,355) exhibited MAF exceeding 0.2, which were then selected as markers with normal allele frequencies, and 9.60% of the SNPs (2337/24,355) exhibited almost equal allele frequencies (i.e. MAFs close to 0.5). The overall SNP diversity was expressed by a PIC value. The PIC values estimated for 24,355 SNP markers ranged from 0.05 to 0.38 with a mean of 0.27, and about 69% of SNPs (16,757/24,355) had PIC values over 0.20.

The number of markers and map lengths of the composite genetic map are shown in table 3. The total length of the map was 3674.16 cM with a mean genetic distance of 0.15 cM between markers. Chromosome 1B contained the highest number of markers ($n = 2390$), followed by 5B ($n = 2187$), and chromosome 4D had a very few loci ($n = 78$). Among A, B and D genomes, B genome contained the most loci ($n = 12,321$) with a total length of 1150.47 cM, followed by A genome ($n = 9523$) with a length of 1252.51 cM, and D genome ($n = 2511$) with a length of 1271.18 cM.

Population structure

The potential existence of population structure among the 205 winter wheat accessions was detected using STRUCTURE program. The ΔK value was plotted against the number of hypothetical subgroups K with the highest ΔK observed when $K = 4$ (figure 1), indicating that four subgroups exist in the association panel used in this study (figure 2). Q values of 205 accessions in four groups are shown in table 2 of electronic supplementary material. Using the maximum membership probability in STRUCTURE, subgroup 1 (43 varieties), subgroup 2 (32 varieties) and subgroup 3 (105 varieties) were dominated by varieties from Henan province, Hebei province and Shandong province, respectively, and subgroup 4 (25 varieties) was dominated by those from three provinces (Shanxi, Jiangsu and Ningxia provinces). The clustering analysis based on the genotypic data generally corresponded to known pedigrees, and the varieties that are closely related in the pedigree usually clustered together.

Table 2. Correlation analysis between SER and tiller number.

Trait[a]	SER	MTS	ETH
MTS	0.296**		
ETH	0.388**	0.663**	
RAT	0.338**	−0.314**	0.476**

[a]Trait abbreviations are same as for table 1.
**Correlation is significant at the 0.01 level (2-tailed).

MTA

In this study, 24,355 SNPs that mapped to the 21 wheat chromosomes were used for the MTA analysis. After FDR corrections for multiple testing (with $P = 0.05$), 31 significant MTAs were detected for the five traits in the three environments. These SNPs were located on 12 chromosomes and accounted for 5.35–11.33% of PVE. Detailed information on significant markers for each trait are shown in table 4. Manhattan plots of the GWAS results are shown in figure 3.

Four MTAs for SER were detected on chromosomes 5A, 7A and 7B and explained 5.51–8.18% of PVE. Of which, the marker *Kukri_c28160_2017* on chromosome 7B had the

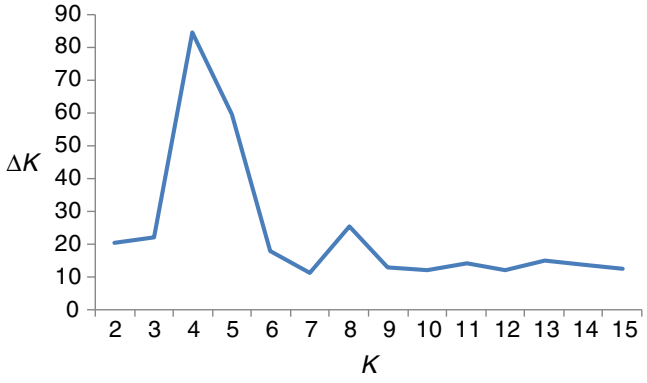

Figure 1. Rate of change in the log probability of data between successive K values (ΔK).

Table 3. Information of SNP in the composite linkage map.

Chr.	Marker number	Length of linkage group (cM)	Chr.	Marker number	Length of linkage group (cM)	Chr.	Marker number	Length of linkage group (cM)
1A	1506	161.35	1B	2390	174.1	1D	629	196.97
2A	1462	185.46	2B	1977	180.33	2D	769	151.92
3A	1154	184.56	3B	1628	150.97	3D	331	156.06
4A	1145	164.12	4B	882	118.91	4D	78	161.1
5A	1243	144.15	5B	2187	219.77	5D	240	207.32
6A	1463	180.74	6B	1786	127.54	6D	234	156.53
7A	1550	232.13	7B	1471	178.85	7D	230	241.28
Genome A	9523	1252.51	Genome B	12321	1150.47	Genome D	2511	1271.18
Total	24355	3674.16						

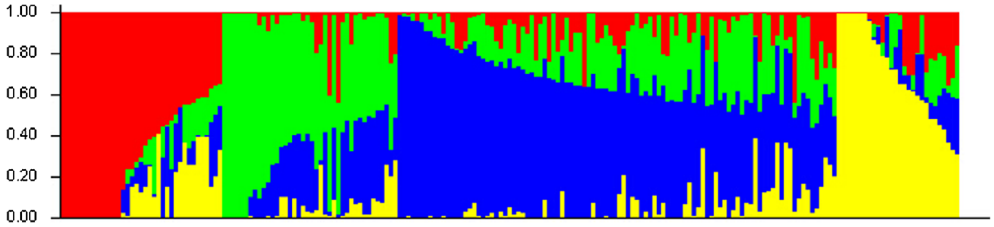

Figure 2. Population structure of the association panel based on SNP marker.

Table 4. Summary of marker–trait associations detected for the five traits in three environments.

Trait[a]	Marker	Alleles	Chr.	Position (cM)	*P* value[c]			R^2 (%)		
					E1[b]	E2[b]	E3[b]	E1[b]	E2[b]	E3[b]
SER	*wsnp_Ex_c4666_8349206*	AA/GG	5A	45		4.23E-07(0.010)			7.64	
	wsnp_Ex_rep_c107017_90850230	GG/TT	5A	111			5.06E-07(0.012)			5.51
	Ra_c14761_1348	GG/TT	7A	194			1.34E-06(0.016)			6.19
	Kukri_c28160_2017	AA/GG	7B	147		3.25E-06(0.039)	1.40E-06(0.011)		8.18	6.03
MTW	*Tdurum_contig13879_919*	CC/TT	1B	160	2.35E-07(0.005)			9.19		
	GENE-0592_268	AA/CC	2B	57	4.15E-07(0.005)		2.48E-06(0.049)	8.15		9.02
	wsnp_Ex_c12341_19693090	AA/GG	3A	123		6.35E-06(0.049)			7.92	
	BS00097939_51	AA/GG	3A	173		1.49E-06(0.036)			9.65	
	CAP7_rep_c12537_81	AA/GG	3A	177		1.64E-06(0.020)			9.31	
	CAP8_rep_c9749_188	GG/TT	3D	143	2.78E-06(0.011)			9.15		
	Tdurum_contig51422_127	AA/GG	5B	129	5.73E-06(0.020)			8.59		
	Excalibur_c25522_755	CC/TT	5B	144	5.32E-07(0.004)			10.14		
	Excalibur_c25522_791	AA/GG	5B	144	5.32E-07(0.003)			10.14		
	Excalibur_c60554_394	GG/TT	5B	144	5.32E-07(0.003)			10.14		
MTS	*Excalibur_c40993_1129*	CC/TT	1B	78	2.64E-06(0.021)	4.15E-06(0.046)		5.35	5.68	
	Tdurum_contig13879_919	CC/TT	1B	160	6.78E-07(0.017)			8.38		
	Tdurum_contig45965_563	CC/TT	1B	160	7.84E-07(0.010)			8.07		
	Kukri_c47259_416	AA/GG	2B	114		4.65E-07(0.011)	8.33E-06(0.101)		8.24	5.82
	RAC875_c14309_317	AA/GG	6B	76		6.14E-06(0.049)	5.53E-06(0.134)		5.74	5.61
ETH	*Kukri_c17540_394*	AA/GG	1B	63			5.77E-07(0.014)			8.84
	Tdurum_contig13879_919	CC/TT	1B	160			9.17E-07(0.007)			7.93
	BS00011478_51	GG/TT	2A	139			9.00E-07(0.011)			7.92
	Jagger_c6853_60	AA/GG	2B	103	1.53E-06(0.037)		3.36E-06(0.020)	7.29		6.71
	BS00097939_51	AA/GG	3A	173		4.18E-06(0.047)			6.21	
	Tdurum_contig4974_355	CC/TT	4B	61		4.35E-07(0.011)			9.08	
	wsnp_Ex_c14654_22713386	AA/GG	7A	42	5.76E-06(0.070)	6.56E-06(0.049)		5.94	8.36	
RAT	*wsnp_Ex_rep_c69766_68723140*	CC/TT	1B	79	3.76E-06(0.046)		3.18E-06(0.026)	7.14		7.09
	BobWhite_c8436_391	CC/TT	4A	153			5.18E-07(0.013)			11.33
	Excalibur_c11045_236	AA/GG	5B	72	6.95E-07(0.017)			9.41		
	BS00073989_51	CC/TT	7A	211		3.83E-06(0.041)	2.56E-06(0.032)		7.52	7.29
	IAAV3713	AA/GG	7B	145	4.16E-06(0.034)	2.57E-06(0.049)		6.17	6.15	

[a]Trait abbreviations are same as for table 1.
[b]Same as for table 1.
[c]The data in brackets are corrected *P* values.

highest R^2 (8.18%) and was detected as significant in two environments (figure 3a).

We detected 10 MTAs for MTW on chromosomes 1B, 2B, 3A, 3D and 5B, explaining 7.92–10.14% of PVE. The marker *GENE-0592_268* on chromosome 2B exhibited a significant MTA in two environments. Three markers, *Excalibur_c25522_755*, *Excalibur_c25522_791* and *Excalibur_c60554_394* located at effectively the same position of chromosome 5B (144 cM), had the highest R^2 (10.14%) and the smallest *P* value (figure 3b).

Five MTAs on chromosomes 1B, 2B and 6B were identified for MTS, which explained phenotypic variance ranging from 5.35 to 8.38%. The markers *Excalibur_c40993_1129* on chromosomes 1B, *Kukri_c47259_416* on chromosome 2B and *RAC875_c14309_317* on chromosome 6B were inferred to have significant MTA in two environments (figure 3c).

For ETH, seven MTAs on chromosomes 1B, 2A, 2B, 3A, 4B and 7A were found to account for 5.94 to 9.08% of PVE, of which the markers *Jagger_c6853_60* on chromosome 2B and *wsnp_Ex_c14654_22713386* on chromosome

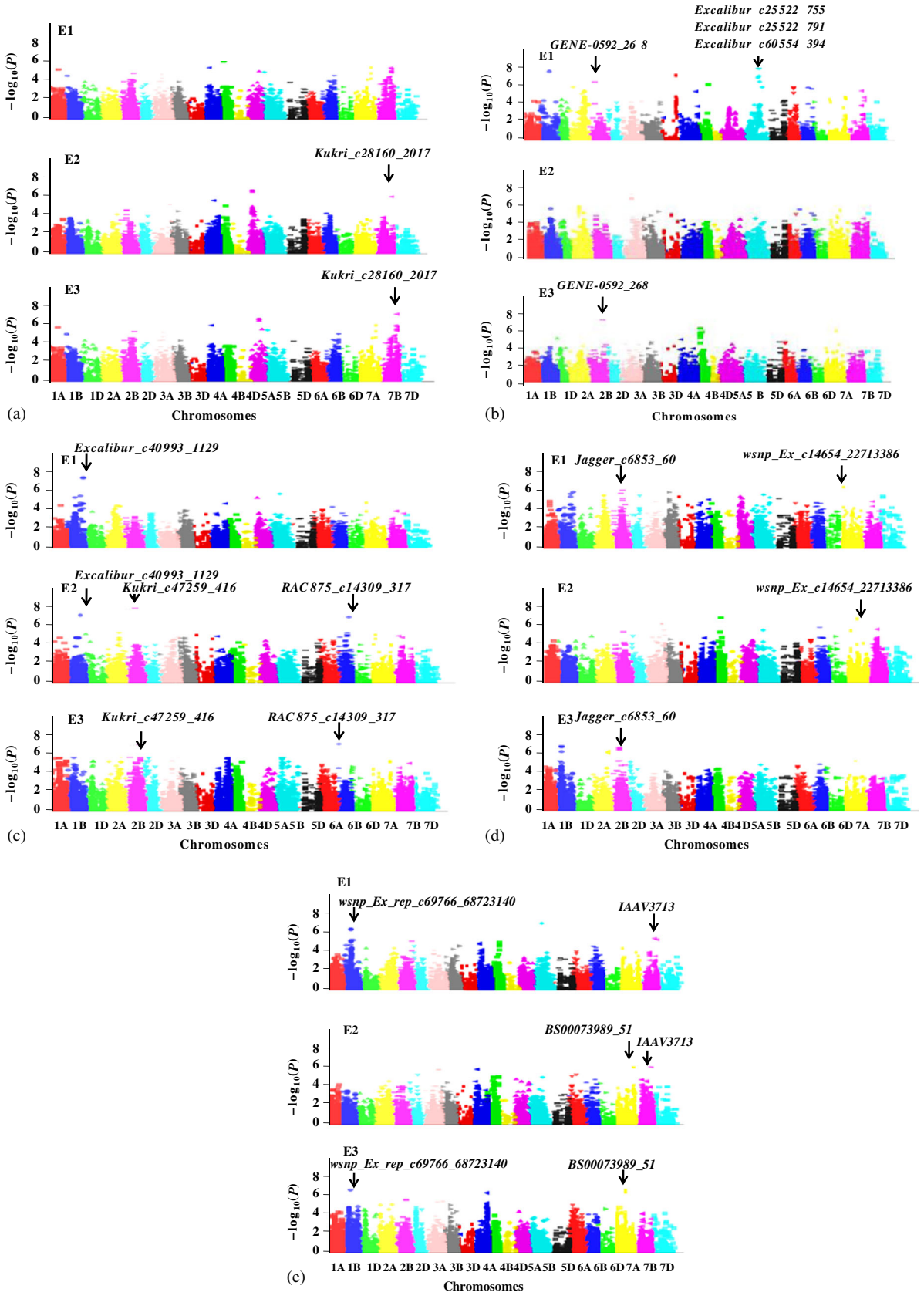

Figure 3. Manhattan plot of GWAS for five traits with general linear model. (a), (b), (c), (d) and (e) indicate manhattan plots for SER, MTW, MTS, ETH and RAT, respectively; E1, E2 and E3 indicate 2013 Tai'an, 2014 Tai'an and 2014 Dezhou, respectively).

7A exhibited significant MTAs in two environments (figure 3d).

For RAT, five MTAs on chromosomes 1B, 4A, 5B, 7A and 7B were detected, which explained 6.15–11.33% of PVE. The markers *wsnp_Ex_rep_c69766_68723140* on chromosome 1B, *BS00073989_51* on chromosome 7A and *IAAV3713* on chromosome 7B exhibited significant MTAs in two environments (figure 3e).

Overall, 10 MTAs were consistently detected in two environments, which were therefore regarded as relatively stable loci controlling these traits. The highest number of stable loci were found for MTS and RAT, with three MTAs each, followed by ETH with two MTAs. Only one stable MTA was detected for SER and MTW. Further, most of the MTAs were detected in just a single environment, indicating SER and tiller-related traits were substantially affected by environmental factors.

Elite allele exploration

The genetic bases of high SER and RAT values have always been a research focus, but it has been a difficult area for breeders. In this study, to explore the elite alleles of these traits, MTAs for SER and RAT were further analysed through a comparison of phenotypic effects between alleles at each locus. Phenotypic effects of alleles for the loci influencing SER and RAT are given in table 5.

Three of four MTAs for SER, exhibited phenotypic differences between alleles that reached a significant level ($P < 0.01$). Of these, alleles T and G of the marker *Ra_c14761_1348* were associated with the largest phenotypic differences (13.35%). The phenotypic value of SER with *Ra_c14761_1348-T* was significantly higher than that of *Ra_c14761_1348-G* across all three environments, which indicats *Ra_c14761_1348-T* is an elite allele when compared with *Ra_c14761_1348-G*. *wsnp_Ex_c4666_8349206-G* and *wsnp_Ex_rep_ c107017_ 90850230-T* alleles were two other elite alleles explored in this study.

Four of five MTAs for RAT, exhibited significant phenotypic differences among alleles ($P < 0.01$). Four alleles, *wsnp_Ex_rep_c69766_68723140-T*, *BobWhite_c8436_391-T*, *Excalibur_c11045_236-A* and *IAAV3713-G* were elite relative to their alternative alleles at each locus, which have increased RAT from 4.23 to 14.78%. *Excalibur_c11045_236-A* had the most significant effect, increasing RAT by 14.78%, followed by *BobWhite_c8436_391-T* (8.47%) and *IAAV3713-G* (4.23%). The other marker, *BS00073989_51*, did not exhibit a significant phenotypic difference among its alleles.

Discussion

Rapid advancements in sequencing and genotyping technologies over the last decade have enabled the development of SNP arrays even for polyploid crops such as wheat (Edwards *et al.* 2012). GWAS using SNPs is a promising avenue for future research that combines routinely generated phenotypic data with high-density markers to elucidate the genetic architecture of agronomic traits. Here, we used the currently available 90K SNP array to genotype a mapping population

Table 5. Phenotypic effect of alleles for the loci of SER and RAT.

Locus	Chr.	Allele	Material number	Env.[b] E1	Env.[b] E2	Env.[b] E3	Average	Difference[c]
SER[a] (%)								
wsnp_Ex_c4666_8349206	5A	G	154	61.38 A	67.10 A	65.00 A	64.49 A	10.48
		A	51	49.24 B	55.80 B	57.00 B	54.01 B	
wsnp_Ex_rep_c107017_90850230	5A	T	174	57.89 A	64.12 A	66.12 A	62.71 A	8.01
		G	31	51.36 B	63.25 A	49.48 B	54.70 B	
Ra_c14761_1348	7A	T	179	59.48 A	69.87 A	64.99 A	64.78 A	13.35
		G	26	43.27 B	53.89 B	57.13 B	51.43 B	
Kukri_c28160_2017	7B	A	120	55.35 A	69.86 A	63.22 A	62.81 A	2.10
		G	85	51.23 A	64.78 A	66.12 A	60.71 A	
RAT[a] (%)								
wsnp_Ex_rep_c69766_68723140	1B	T	143	62.78 A	54.32 A	50.45 A	55.85 A	7.48
		C	62	52.17 B	45.34 B	47.60 A	48.37 B	
BobWhite_c8436_391	4A	T	33	61.45 A	52.37 A	53.74 A	55.85 A	8.47
		C	172	54.43 B	44.65 B	43.06 B	47.38 B	
Excalibur_c11045_236	5B	A	189	63.22 A	59.14 A	57.12 A	59.83 A	14.78
		G	16	50.36 B	41.78 B	43.01 B	45.05 B	
BS00073989_51	7A	C	156	58.23 A	53.34 A	50.39 A	53.99 A	1.19
		T	49	56.34 A	49.78 A	52.28 A	52.80 A	
IAAV3713	7B	G	185	59.36 A	54.78 A	51.47 A	55.20 A	4.23
		A	20	55.23 B	45.49 B	52.19 B	50.97 B	

[a]Trait abbreviations are same as for table 1.
[b]Same as for table 1.
[c]Difference between alleles. Different capital letters after data indicate significant difference between alleles on one locus at $P \leq 0.01$.

panel, from which we performed GWAS for SER and tiller number and investigated the advantages of this approach.

SNP markers for GWAS in wheat

High-density marker coverage of the genome and reliable genotype data are the preconditions for GWAS. SNP markers are the most common type of sequence variation in the genome (Rafalski 2002), and they are consequently well-suited for genomics approaches that require a high number of markers such as association mapping (Myles *et al.* 2009). High-throughput SNP genotyping platforms have been available for diploid crops like rice (Huang *et al.* 2010) or maize (Yang *et al.* 2006) for some time, but not for polyploid species. For wheat, SNP discovery and detection have both progressed much more slowly for several reasons: (i) the lack of a reference genome sequence, (ii) availability of gene sequences from only a few genotypes, and (iii) the highly repetitive and duplicated nature of the wheat allopolyploid genome (Edwards and Batley 2010). Consequently, genotypic data in wheat has so far been obtained with traditional markers like AFLPs and SSRs (Zhang *et al.* 2013), which are laborious and time consuming to use. SNP arrays constitute a major breakthrough in wheat genotyping, enabling the cost-effective, large-scale acquisition of genotypic data in hexaploid wheat (Akhunov *et al.* 2009). Recently, the wheat 90K iSelect SNP genotyping array has made GWAS possible in wheat. In the present study, we used the currently available 90K SNP array developed by Wang *et al.* (2014) to genotype a panel of 205 diverse Chinese winter wheat lines. The very high number of SNPs on the 90K array revealed clear genetic differences among the 205 wheat accessions, specifically, 32,432 SNPs. Of these 32,432 SNPs, a total of 24,355 SNPs mapped to the 21 wheat chromosomes and were used for MTA analysis. The size of our dataset in marker number (24,355 SNPs) for GWAS was larger than that of similar association mapping studies that have been carried out in wheat so far (e.g., Somers *et al.* 2007; Zhang *et al.* 2011; Le *et al.* 2012). A high number of markers are required for genomic studies like GWAS to increase mapping resolution. In the near future, when the complete genome sequence is available for wheat, GWA mapping will be facilitated by SNP markers with known physical locations in the genome, and this can facilitate the discovery of major genes that control the traits under investigation as it has already occurred in rice (Famoso *et al.* 2011) and maize (Setter *et al.* 2011).

Population structure analysis

Different levels of population structure in Chinese winter wheat have been reported (Chen *et al.* 2012). In our study, using the maximum membership probability in STRUCTURE, the association population was classified into four groups (comprised of 43, 32, 105 and 25 varieties, respectively) (figure 2). Individuals were unequally distributed among the four subgroups, which indicate that there was a high extent of population structure in our mapping population. Some elite lines typically are used in many crosses and thus, genetically contribute more to the association panel population than other lines that have been used in breeding populations (Würschum *et al.* 2011). The causes are also likely to be important for the subpopulation differentiation observed in our mapping population because our mapping panel included many bred varieties and breeder's lines, which shared one or more founder parents in wheat breeding programmes. Consequently, a correction for population structure might be required to enhance analyses of this dataset. Moreover, the population structure analysis was used to explore the relationships among varieties used in the association mapping panel, which is important not only for parental selection but also for breeding system design.

MTA analysis

Seedling emergence is one of the most important indices for seed vigour in agricultural production. Earlier studies have reported QTLs for seed vigour on at least 13 chromosomes, including 1A, 1B, 1D, 2A, 2D, 3A, 3B, 4B, 5A, 5B, 5D, 7A and 7D (Rebetzke *et al.* 2001; Landjeva *et al.* 2010; Liu *et al.* 2013; Xiao *et al.* 2013; Li *et al.* 2014b). In the present study, four MTAs were detected for SER on chromosomes 5A, 7A and 7B. By comparing these chromosomes, some important loci controlling seed vigour on chromosomes 5A and 7A were detected.

Earlier studies have detected QTLs for tillering on almost all wheat chromosomes using the traditional QTL mapping approach (Kato *et al.* 2000; Kuraparthy *et al.* 2007; Li *et al.* 2010; Deng *et al.* 2011). Compared with previous reports, the marker *Kukri_c47259_416* on chromosome 2B associated with MTS was located in a similar position as a significant SSR marker reported by Li *et al.* (2010), who used a biparental population. Marker *BS00097939_51* on chromosome 3A in this study was a pleiotropic loci, explaining about 9.60 and 6.01% of PVE in MTW and ETH, respectively. Kato *et al.* (2000) and Kuraparthy *et al.* (2007) also detected a QTL for tiller number on chromosome 3A with significant effects. Deng *et al.* (2011) detected a QTL for ETH on chromosome 4B. Yang *et al.* (2013) also detected a pleiotropic QTL on chromosome 4B associated with both ETH and MTS. At a similar location, the marker *Tdurum_contig4974_355* was associated with ETH in this study. This indicates that an association mapping approach is a viable alternative to QTL mapping to identify some quantitative traits such as those corresponding to tillering-related traits, which are influenced by environmental effects. Further, we found some QTLs with significant effects on tiller number that were not reported by previous studies. For example, the major marker, *Tdurum_contig13879_919* on chromosome 1B was detected in different growth stages and explained 7.98–9.69% of PVE, thus playing an important role in wheat tiller development. Whether these QTLs correspond to a new locus detected using GWAS remain to be further confirmed.

The expression of a stable MTA is less affected by the environment. In general, a MTA that is consistent across environments is of great value for marker-assisted selection in breeding varieties adapted to various ecological environments. Therefore, the three stable MTAs for MTS, three for RAT, two for ETH, and one each for SER and MTW detected in two environments in the present study should be used to develop cleaved amplified polymorphic sequence (CAPS) markers for marker-assisted selection and to begin mining SER and tillering genes through bioinformatic analyses.

Elite allele analysis

Identifying useful alleles by association analysis has become a practical strategy for plant genomics research over recent years. In the present study, we established a link between genotypes and SER and RAT phenotypes by analying differences in phenotypic values between different alleles and identified elite alleles for SER and RAT. For example, the allele *Ra_c14761_1348-T* could increase SER by 13.35%, while alleles *Excalibur_c11045_236-A* and *BobWhite_c8436_391-T* could increase RAT by 14.78 and 8.47%, respectively. Accordingly, the lines that carry these elite alleles should be used as genetic stock for the improvement of these traits through breeding.

In conclusion, the association mapping is a powerful tool for identifying molecular markers associated with agronomic and quality traits in wheat. Thirty-one significant loci were determined to underlie variation in SER and tiller number in different growth stages in this population using GWAS. Of these loci, markers for SER and tillering traits at 10 loci were consistently detected in two environments and should therefore be used to develop CAPS markers. A set of elite alleles for SER and RAT were identified for the purpose of improving these traits through breeding procedures.

Acknowledgements

This research was supported by the Shandong Provincial Agriculture Liangzhong Project Foundation of China (2014 no. 96) and National Natural Science Foundation of China (no. 31171554).

References

Akhunov E., Nicolet C. and Dvorak J. 2009 Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. *Theor. Appl. Genet.* **119**, 507–517.

Botwright T. L., Condon A. G., Rebetzke G. J. and Richards R. A. 2002 Field evaluation of early vigor for genetic improvement of grain yield in wheat. *Aust. J. Agric. Res.* **53**, 1137–1145.

Cervantes-Ortiz F., García-De G., Carballo-Carballo A., Bergvinson D., Crossa J. L., Mendoza-Elos M. *et al.* 2007 Seedling vigor inheritance and its relationship to adult plant traits in inbred tropical maize lines. *Agrociencia* **41**, 425–433.

Chen X. J., Min D. H., Tauqeer A. Y. and Hu Y. G. 2012 Genetic diversity, population structure and linkage disequilibrium in elite Chinese winter wheat investigated with SSR markers. *PLoS One* **7**, e44510.

Cockram J., White J., Zuluaga D. L., Smith D., Comadran J., Macaulay M. *et al.* 2010 Genomewide association mapping to candidate polymorphism resolution in the unsequenced barley genome. *Proc. Natl. Acad. Sci. USA* **107**, 21611–21616.

Deng S., Wu X., Wu Y., Zhou R., Wang H., Jia J. and Liu S. 2011 Characterization and precise mapping of a QTL increasing spike number with pleiotropic effects in wheat. *Theor. Appl. Genet.* **122**, 281–289.

Edwards D. and Batley J. 2010 Plant genome sequencing: applications for crop improvement. *Plant Biotechnol.* **8**, 2–9.

Edwards D., Batley J. and Snowdon R. J. 2012 Accessing complex crop genomes with next-generation sequencing. *Theor. Appl. Genet.* **126**, 1–11.

Eric M., James J. M., Frank A. M., Benjamin N. and Stephen V. F. 2011 Genome-wide association study of blood pressure response to methylphenidate treatment of attention-deficit/hyperactivity disorder. *Prog. Neuro-Psychopharm. Biol. Psychiat.* **35**, 466–472.

Evanno G., Regnaut S. and Goudet J. 2005 Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611–2620.

Evers J. B., Vos J., Andrieu B. and Struik P. C. 2006 Cessation of tillering in spring wheat in relation to light interception and red: far-red ratio. *Ann. Bot.* **97**, 649–658.

Famoso A. N., Zhao K., Clark R. T., Tung C. W. and Wright M. H. 2011 Genetic architecture of aluminum tolerance in rice (*Oryza sativa*) determined through genome-wide association analysis and QTL mapping. *PLoS Genet.* **7**, e1002221.

Fan J. B., Gunderson K. L., Bibikova M., Yeakley J. M., Chen J., Wickham Garcia E. *et al.* 2006 Illumina universal bead arrays. *Methods Enzymol.* **410**, 57–73.

Gerrard D. T., Fricke C., Edward D. A., Edwards D. R. and Chapman T. 2013 Genome-wide responses of female fruit flies subjected to divergent mating regimes. *PLoS One* **8**, e68136.

Gupta P. K., Rustgi S. and Mir R. R. 2008 Array-based high-throughput DNA markers for crop improvement. *Heredity* **101**, 5–18.

Hodges T. 1991 *Predicting crop phenology*. CRC Press, Boca Raton, USA.

Huang X. H., Wei X. H., Sang T., Zhao Q., Feng Q., Zhao Y. *et al.* 2010 Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Gene* **42**, 961–967.

Huang X. H., Zhao Y., Wei X. H., Li C. Y. and Wang A. H. 2012 Genome-wide association study of flowering time and grain yield traits in a worldwide collection of rice germplasm. *Nat. Genet.* **44**, 32–39.

Ingvarsson P. K. and Nathaniel R. S. 2011 Association genetics of complex traits in plants. *New Phytol.* **189**, 909–922.

Ismail A. M., Johnson D. E., Ella E. S., Vergara G. V. and Baltazar A. M. 2012 Adaptation to flooding during emergence and seedling growth in rice and weeds, and implications for crop establishment. *AoB Plant.* **pls019**, 1–18.

Kato K., Miura H. and Sawada S. 2000 Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. *Theor. Appl. Genet.* **101**, 1114–1121.

Kuraparthy V., Sood S., Dhaliwal H. S., Chhuneja P. and Gill B. S. 2007 Identification and mapping of a tiller inhibition gene (tin3) in wheat. *Theor. Appl. Genet.* **114**, 285–294.

Landjeva S., Lohwasser U. and Borner A. 2010 Genetic mapping within the wheat D genome reveals QTL for germination, seed vigor and longevity, and early seedling growth. *Euphytica* **171**, 129–143.

Le G. J., Bordes J., Ravel C., Heumez E. and Faure S. 2012 Genome-wide association analysis to identify chromosomal regions determining components of earliness in wheat. *Theor. Appl. Genet.* **124**, 597–611.

Li F., Chen B. Y., Xu K. and Wu X. M. 2014a Genome-wide association study dissects the genetic architecture of seed weight and seed quality in Rapeseed (*Brassica napus* L.). *DNA Res.* **21**, 355–367.

Li X. M., Chen X. M., Xiao Y. G., Xia X. C., Wang D. S., He Z. H. *et al.* 2014b Identification of QTLs for seedling vigor in winter wheat. *Euphytica* **198**, 199–209.

Li Y. X., Wang Y., Shi Y. S., Song Y. C., Wang T. Y. and Li Y. 2009 Correlation analysis and QTL mapping for traits of kernel structure and yield components in maize. *Scientia Agric. Sin.* **2**, 408–418.

Li Z., Peng T., Xie Q., Han S. and Tian J. 2010 Mapping of QTL for tiller number at different stages of growth in wheat using double haploid and immortalized F_2 populations. *J. Genet.* **89**, 405–415.

Liu K. and Muse S. V. 2005 PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* **21**, 2128–2129.

Liu X., Li R., Chang X. and Jing R. 2013 Mapping QTLs for seedling root traits in a doubled haploid wheat population under different water regimes. *Euphytica* **189**, 51–66.

Ludwig F. and Asseng S. 2010 Potential benefits of early vigor and changes in phenology in wheat to adapt to warmer and drier climates. *Agric. Syst.* **103**, 127–136.

Myles S., Peiffer J., Brown P. J., Ersoz E. S., Zhang Z., Costich D. E. *et al.* 2009 Association mapping: critical considerations shift from genotyping to experimental design. *Plant Cell* **21**, 2194–2202.

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

Rafalski A. 2002 Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* **5**, 94–100.

Rebetzke G. J., Appels R., Morrison A., Richards R. A., McDonald G., Ellis M. H. *et al.* 2001 Quantitative trait loci on chromosome 4B for coleoptile length and early vigor in wheat (*Triticum aestivum* L.). *Aust. J. Agric. Res.* **52**, 1221–1234.

Rebetzke G. J., Bruce S. E. and Kirkegaard J. A. 2005 Longer coleoptiles improve emergence through crop residues to increase seedling number and biomass in wheat. *Plant Soil* **272**, 87–100.

Risch N. and Merikangas K. 1996 The future of genetic studies of complex human diseases. *Science* **80**, 1516–1517.

Setter T. L., Yan J. B., Warburton M., Ribaut J. M. and Xu Y. B. 2011 Genetic association mapping identifies single nucleotide polymorphisms in genes that affect abscisic acid levels in maize floral tissues during drought. *J. Exp. Bot.* **62**, 701–716.

Somers D. J., Banks T., De Pauw R., Fox S., Clarke J., Pozniak C. and McCartney C. 2007 Genome-wide linkage disequilibrium analysis in bread wheat and durum wheat. *Genome* **50**, 557–567.

Spielmeyer W., Hyles J., Joaquim P., Azanza F., Bonnett D., Ellis M. E. *et al.* 2007 A QTL on chromosome 6A in bread wheat (*Triticum aestivum*) is associated with longer coleoptiles, greater seedling vigour and final plant height. *Theor. Appl. Genet.* **115**, 59–66.

Sukumaran S. and Yu J. 2014 Association mapping of genetic resources: achievements and future perspectives. In *Genomics of plant genetic resources*, pp. 207–235. Springer, Dordrecht, The Netherlands.

Sukumaran S., Dreisigacker S., Lopes M., Chavez P. and Reynolds M. P. 2014 Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. *Theor. Appl. Genet.* (doi: 10.1007/s00122-014-2435-3).

Wang S., Wong D., Forrest K., Allen A., Chao S. and Huang E. 2014 Characterization of polyploid wheat genomic diversity using a high-density 90,000 SNP array. *Plant Biotechnol. J.* **12**, 787–796.

Waugh R., Jannink J. L., Muehlbauer G. J. and Ramsay L. 2009 The emergence of whole genome association scans in barley. *Curr. Opin. Plant Biol.* **12**, 218–222.

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

Würschum T., Maurer H. P., Kraft T., Janssen G., Nilsson C. and Reif J. C. 2011 Genome-wide association mapping of agronomic traits in sugar beet. *Theor. Appl. Genet.* **123**, 1121–1131.

Xiao Y. G., Liu J. J., Xia X. C., Chen X. M., Reynolds M. P. and He Z. H. 2013 Genetic analysis of early vigour in winter wheat using digital imaging. *Acta Agron. Sin.* **39**, 1935–1943.

Yu J. M. and Buckler E. S. 2006 Genetic association mapping and genome organization of maize. *Curr. Opin. Biotech.* **17**, 155–160.

Yang G. H., Xing Y. Z., Li S., Ding J. Z., Yue B., Deng K. *et al.* 2006 Molecular dissection of developmental behavior of tiller number and plant height and their relationship in rice (*Oryza sativa* L.). *Hereditas* **143**, 236–245.

Yang L., Shao H., Wu Q. X., Yu J., Ran C. F., Li L. Q. *et al.* 2013 QTLs mapping and epistasis analysis for the number of tillers and spike number per plant. *J. Triticeae Crops* **33**, 875–882.

Zanke C., Ling J., Plieske J., Kollers S., Ebmeye E., Korzun V. *et al.* 2014 Genetic architecture of main effect QTL for heading date in European winter wheat. *Plant Genet. Genomics*, doi: 10.3389/fpls.2014.00217.

Zhang J., Hao C., Ren Q., Chang X., Liu G. and Jing R. 2011 Association mapping of dynamic developmental plant height in common wheat. *Planta* **234**, 891–902.

Zhang K. P., Wang J. J., Zhang L. Y., Rong C. W., Zhao F. W., Peng T. *et al.* 2013 Association analysis of genomic loci important for grain weight control in elite common wheat varieties cultivated with variable water and fertiliser supply. *PLoS One* **8**, e57853.

Zhu Y. L., Song Q. J., Hyten D. L., Van C. P., Matukumalli L. K., Grimm D. R. *et al.* 2003 Single-nucleotide polymorphisms in soybean. *Genetics* **163**, 1123–1134.

Zhu C., Gore M., Buckler E. S. and Yu J. 2008 Status and prospects of association mapping in plants. *Plant Genome* **1**, 5–20.

Received 13 October 2015, in revised form 28 April 2016; accepted 6 May 2016
Unedited version published online: 11 May 2016
Final version published online: 25 January 2017

Corresponding editor: UMESH C. LAVANIA