Genome-wide association study of pre-harvest sprouting resistance in Chinese wheat founder parents

Yu Lin1,*, Shihang Liu1,*, Yaxi Liu1, Yujiao Liu1,*, Guoyue Chen1, Jie Xu2, Mei Deng1, Qiantao Jiang1, Yuming Wei1, Yanli Lu2 and Youliang Zheng1

1Triticeae Research Institute, Sichuan Agricultural University, Wenjiang, Chengdu, P.R. China.
2Maize Research Institute, Sichuan Agricultural University, Wenjiang, Chengdu, P.R. China.

Abstract

Pre-harvest sprouting (PHS) is a major abiotic factor affecting grain weight and quality, and is caused by an early break in seed dormancy. Association mapping (AM) is used to detect correlations between phenotypes and genotypes based on linkage disequilibrium (LD) in wheat breeding programs. We evaluated seed dormancy in 80 Chinese wheat founder parents in five environments and performed a genome-wide association study using 6,057 markers, including 93 simple sequence repeat (SSR), 1,472 diversity array technology (DArT), and 4,492 single nucleotide polymorphism (SNP) markers. The general linear model (GLM) and the mixed linear model (MLM) were used in this study, and two significant markers (tPt-7980 and wPt-6457) were identified. Both markers were located on Chromosome (Chr.) 3AS and 4AL, respectively.

Keywords: General linear model, linkage disequilibrium, marker-trait associations, mixed linear model, seed dormancy.
Received: August 17, 2016; Accepted: February 28, 2017.

Introduction

In Chinese wheat breeding programs, the crossing parents, known as “founder parents,” are wheat varieties or germplasm lines with excellent properties, such as resistance to numerous biotic and abiotic stresses, high combining ability, and strong heritability of superior agronomic traits, that are used for the development of improved cultivars with wide application value (Zhuang, 2003; Chen et al., 2013). Over the past seven decades, more than 2,000 wheat (Triticum aestivum L.) cultivars have been released through pedigree selection programs in China. Most of these, however, can be traced back to only 16 ancestral founder parents (Zhuang, 2003; Li et al., 2012). Founder parents have been used in multiple studies on genetic diversity (Wang et al., 2007; Li et al., 2008; 2012), quantitative trait loci (QTL) mapping (Lin et al., 2006; Su et al., 2006; Ma et al., 2007) and association studies (Brescaghello and Sorrells, 2006a; Crossa et al., 2007; Xiao et al., 2011).

PHS is the germination of grains within a physiologically mature wheat spike before harvest (Groos et al., 2002; Jaiswal et al., 2012). PHS in wheat (Triticum aestivum L.) is the result of the early break in seed dormancy under humid and wet conditions prior to harvest (Rikiishi and Maekawa, 2010; Kulwal et al., 2012), causing significant losses in grain weight and end-product quality. Percent germination (PG) is a variable commonly used to characterize the resistance to PHS based on seed dormancy (Knox et al., 2005, 2012; Rasul et al., 2009; Singh et al., 2010; Kumar et al., 2015). Previous studies reported that almost all 21 chromosomes of hexaploid wheat contain QTL for PHS resistance (Roy et al., 1999; Zanetti et al., 2000; Mares and Mrva, 2001; Flintham et al., 2002; Kulwal et al., 2005; Chen et al., 2008; Fofana et al., 2009; Kulwal et al., 2010; Kim et al., 2014; Somyong et al., 2014). One major QTL TaPHSI mapped on Chromosome (Chr.) 3AS has been cloned (Nakamura et al., 2011; Liu et al., 2011), whereas another major QTL on Chr. 4AL has been fine mapped with SNP markers (Cabral et al., 2014; Barrero et al., 2015; Lin et al., 2015).

AM, also known as LD mapping, relies on existing natural populations or specially designed populations to overcome the constraints of linkage mapping (Pasam et al., 2012). AM is a powerful tool for resolving complex trait variations and identifying different loci and/or novel and superior alleles in natural populations (Zhu et al., 2008). In recent years, association studies have been extensively used to discover and validate QTL or genes for important traits and map candidate genes in many crop plants. In
heritability was defined as $H = \frac{VG}{VG + VE}$, where $VG$ was estimated in SAS 8.1 (SAS Institute, Cary, NC, USA). Broad-sense heritability was also used (Piepho et al., 2008). BLUP was calculated across the five environments using the MIXED procedure (SAS Institute, Cary, NC, USA). BLUP was also used (Piepho et al., 2008). BLUP was calculated across the five environments using the MIXED procedure in SAS 8.1 (SAS Institute, Cary, NC, USA). Broad-sense heritability was defined as $H = \frac{VG}{VG + VE}$, where $VG$ and $VE$ are estimates of genetic and environmental variance, respectively (Smith et al., 1998).

Analysis of variance (ANOVA) was performed using SAS 8.1 (SAS Institute), whereas descriptive statistics and Pearson’s correlation using SPSS 20.0 (IBM Corp., Armonk, NY, USA).

Genotyping

Three types of molecular markers, SSRs, DArTs, and SNPs, were used in this study. DNA was extracted from young seedlings using the CTAB method (Saghai-Maroof et al., 1994) and then, sent to Triticate Pty. Ltd. (Canberra, Australia) for whole-genome profiling using DArT markers. The panel was also genotyped using the Illumina 9K iSelect SNP chip assay (Cavanagh et al., 2013). Both DArT and SNP markers were filtered to contain < 10% missing values and a minor allele frequency (MAF) threshold > 5%. An additional set of 93 SSR markers distributed across the wheat genome was also screened against the genome-wide association study (GWAS) population. For SSR analysis, PCR was performed as described by Sood et al. (2009). PCR products were separated in 8% polyacrylamide gels and visualized by silver staining.

Population structure

Population structure was estimated with a set of 4,492 SNP markers with an MAF threshold > 0.05 using STRUCTURE 2.3.3, which implements a model-based Bayesian cluster analysis (Pritchard et al., 2000; Wang et al., 2013). Five runs of STRUCTURE were performed with a $K$ set between 1 and 10 using the admixture model with 20,000 replicates for burn-in and 20,000 replicates during analysis. The optimal $K$-value was determined using the delta $K$ method (Evanno et al., 2005).

LD analysis

LD estimates and significant differences between 3,285 SNPs with genetic distances based on the Illumina 9K SNP consensus map (Cavanagh et al., 2013) were calculated in TASSEL 3.0. The LD squared allele-frequency correlation ($r^2$), which contains both mutational and recombination history, was evaluated for linked/syntenic and unlinked loci (p < 0.05). LD decay scatter plots were generated using $r^2$ and the genetic map distance between markers. Non-linear regression equations were performed for LD decay distances.

Association analysis

Population structure ($Q$ matrix) from STRUCTURE was used as covariate for the GLM and MLM, and a marker-based kinship matrix ($K$) obtained using TASSEL was used in the MLM. MTAs were calculated in TASSEL using: i) the GLM with a $Q$ matrix and ii) the MLM with a $Q$ matrix and kinship. PG from the five environments, mean values, and BLUP values were used in the analysis. All models used 6,057 informative markers (93 SSRs, 1,472 DArTs, and 4,492 SNPs) with a MAF threshold > 0.05. A Bonferroni-corrected threshold at $\alpha = 1$ was used as the cutoff (Yang et al., 2014). For 6,057 markers, the Bonferroni-corrected $p$-value threshold at $\alpha = 1$ was 1.65 x 10^-4 with a corresponding -log $p$-value of 3.782. Significant
markers were visualized with a Manhattan plot drawn in R 3.0.3 (R core team, Vienna, Austria). Important p-value distributions (observed vs. cumulative p-values on a -log10 scale) were displayed on a quantile-quantile plot drawn in R. To find candidate genes, flanking genes, and trait-related proteins, we performed a Basic Local Alignment Search Tool (BLAST) search against the International Wheat Genome Sequencing Consortium database (IWGSC) using the marker sequences. A BLASTN search against the National Center for Biotechnology Information (NCBI) database was performed using 5 kb of the best survey sequence around the significant marker from IWGSC BLAST results, and genes from the best hits were listed.

Results

Phenotypic evaluation

PG was significantly different (p < 0.001) among genotypes, environments, and genotype-environment interactions with a heritability of 0.76 (Table 1). Descriptive statistics for the 80 founder parents are shown in Table 2. Pearson’s correlation was used to investigate PG across years, as well as BLUP and mean values (Table 3). Significant correlations were observed between all PG values, except for those of 12WJ and 13YA. BLUP and mean values were also highly correlated with the data from the five environments (Table 3). Frequency histograms of BLUP and mean values of the 80 founder parents are shown in Figure 1.

Table 1 - Variance analysis for percent germination (PG) in five environments.

| Variable | DF | Sum of squares | Mean square | F-value |
|----------|----|----------------|-------------|---------|
| Environments | 4 | 31.19 | 7.80 | 892.13*** |
| Replications | 5 | 1.65 | 0.33 | 37.72*** |
| Genotypes | 79 | 22.69 | 0.29 | 32.86*** |
| G × E | 316 | 18.09 | 0.06 | 6.72*** |

Abbreviation: DF, degrees of freedom. *** Significant at p < 0.001.

Population structure

A set of 4,492 SNP markers was used to estimate the underlying population structure. Delta K declined after K = 3 and again increased. Using K = 3 as inferred by delta K, the population was divided into sub-population (Subp)1, Subp 2, and Subp 3. Based on their pedigrees, Subp 1 included the founder parent Fan 6 and its derivatives; Subp 2 included 23 founder parents and their derivatives; whereas Subp 3 only included 11 founder parents.

LD analysis

The extent of LD was estimated using SNP data (1,547 loci from the A genome, 1,510 loci from the B genome, and 228 from the D genome). Linked and unlinked locus pairs were detected among the 80 founder parents (p < 0.05). The map distance for which LD fell below the $r^2$ threshold of 0.3 was determined. This is a frequently used LD threshold in GWAS (Ardlie et al., 2002; Shifman et al., 2003; Khatkar et al., 2008; Maccaferri et al., 2015). The LD decay distances ($r^2 = 0.3$) were approximately 1.94, 5.73, 1.42, and 17.48 cM for the whole genome and the A, B, and D genomes, respectively (Figure 2). The D genome had the most significant LD, whereas the B genome had the least significant LD.

Table 3 - Correlation coefficients for percent germination (PG).

|          | PG-12WJ | PG-12YA | PG-13WJ | PG-13YA | PG-14CZ | Mean   |
|----------|---------|---------|---------|---------|---------|--------|
| PG-12YA  | 0.604** | 0.351** |         |         |         |        |
| PG-13WJ  | 0.430** |         | 0.305** |         |         |        |
| PG-13YA  | 0.167   | 0.344** |         |         |         |        |
| PG-14CZ  | 0.654** | 0.680** | 0.556** | 0.436** |         |        |
| Mean     | 0.766** | 0.839** | 0.630** | 0.581** | 0.903** | 0.995**|
| BLUP     | 0.722** | 0.833** | 0.626** | 0.580** | 0.912** |        |

Abbreviation: 12WJ, Wenjiang 2012; 12YA, Ya’an 2012; 13WJ, Wenjiang 2013; 13YA, Ya’an 2013; 14CZ, Chongzhou 2014; BLUP, the best linear unbiased prediction.

** significant at p < 0.01.
Significant MTAs for PHS resistance

The Bonferroni-corrected threshold (-log p > 3.782, \( \alpha = 1 \)) was used as a cutoff (Yang et al., 2014) to identify MTAs. A total of 68 significant markers was detected in the five environments by the GLM, with a phenotypic variation of 14.43–30.51% (Table S1; Figures S1–S5). Based on the Illumina 9K SNP consensus map (Cavanagh et al., 2013) and the Triticarte consensus map 3.0 (Chrom_Wheat_ConserverMap_version_3; Triticarte Pty. Ltd.), most of the significant markers were mapped on Chrs. 1A, 1B, 2A, 2B, 3A, 4A, 4B, 5A, 5B, 6A, and 7A, whereas 10 markers remained unmapped. The significant markers were distributed on all the seven chromosomes of the A genome, but none on the D genome. Besides, the DArT marker \( wPt-4301 \) was significant in both 12WJ and 14CZ. In the MLM, three DArT makers were detected with a phenotypic variation of 21.19–36.82% (Table 4), and two of these markers mapped on Chr. 1B (104.00 cM and

Figure 1 - Frequency histograms (a) of BLUPs in 80 Chinese wheat founder parents, (b) of mean values in 80 Chinese wheat founder parents.

Figure 2 - Scatter plots of significant \( r^2 \) values and genetic distance (cM) (p < 0.05) of locus pairs on A, B, D, and whole genomes in 80 Chinese wheat founder parents.
Table 4 - Markers showing significant association with pre-harvest sprouting resistance in the mixed linear model.

| Environment | Marker   | Locus | Locus position (cM) | -Log10(0) | Marker R² (%) |
|-------------|----------|-------|---------------------|------------|---------------|
| 13YA        | wPt-668205 | 1A    | 473.36              | 3.92       | 21.19         |
| 13WJ        | wPt-6457  | 1B    | 104.00              | 4.47       | 36.82         |
| 13WJ        | tPt-7980  | 1B    | 365.57              | 5.34       | 29.94         |

* 13WJ; Wenjiang 2013; 13YA, Ya’an 2013.
* a significant at the threshold of -log10(0) = 3.782.
* b the chromosome and position of DArT markers were from the Triticarte consensus map 3.0 (http://www.triticarte.com.au/).

365.57 cM) were also significant in the GLM (Table 4; Table S1; Figures S1–S5).

BLUP and mean values were also calculated for the five environments and analyzed to interpret the results (Table S1; Figures S6, S7). In the GLM, 11 and 10 significant markers were detected using BLUP and mean values, respectively. Eight of these markers were detected using both BLUP and mean values and mapped on Chr. 2B, 3A, 3D, and 5A. However, no significant markers were found by the MLM using BLUP or mean values.

Putative candidate genes from best hits of significant loci

Based on BLAST search against IWGSC and NCBI using the marker sequences, numerous putative/flanking genes were identified from the best hits of significant loci (Table S1). The identified candidate genes were roughly divided into groups based on the type of encoded proteins. The first group included genes encoding enzymes such as Acc-1 (Chalupska et al., 2008), Acc-2 (Chalupska et al., 2008), CAC3 (Ke et al., 2000), starch synthase I gene (Li et al., 1999), UBA1 (Hatfield et al., 1990), and waxy gene (Takeuchi T, Sato M, Suzuki T, Yoshimura Y, Nakamichi K, Kobayashi S, Nishimura T, Ikenaga M and Sato N. unpublished). The second group included genes encoding regulatory proteins such as vp1D (McKibbin et al., 2002), Rht-A (Wu et al., 2013), Rht-B (Wu et al., 2013), VRN3 (Yan et al., 2006), VRN-A1 (Ivanicova et al., 2016), VRN-B1 (Guedira M, Xiong M, Johnson J, Marshall D and Brown-Guedira G. unpublished), and PRR73 (Cockram et al., 2012). The third group included genes encoding transport proteins such as AACT1 (Silva-Navas et al., 2012) and HAK11 (Banuelos et al., 2002). The fourth group included genes encoding proteins related to stem rust resistance, such as Adj2 (Brueggeman et al., 2008), and the fifth group included genes encoding other proteins such as AIP2-1 (Gao DY, Ma YZ, Xia LQ and Xu ZS. unpublished), GRMZM2G043657-like gene (Jiao et al., 2012), Hox-1 (Wicker et al., 2009), NP30_C3 (Zang et al., 2011), NP35_C3 (Zang et al., 2011), and Ty3 (Hudakova et al., 2001).

Discussion

LD decay distance is important in genome-wide association studies. In the present study, LD decay distance indicated the high marker density and precision of association mapping. Many different LD decay distances have been previously reported in wheat. Chao et al. (2007) reported genome-wide LD estimates of less than 1 cM for genetically linked locus pairs with $r^2 < 0.2$ ($p < 0.01$) and less than 10 cM between loci among 43 elite U.S. wheat cultivars using 242 genomic SSRs; Cormier et al. (2014) reported LD decay distances of 1.12 cM for the whole genome and 0.52, 0.70, and 2.14 cM for the A, B, and D genomes, respectively among 214 European elite varieties using 23,603 genome-wide distributed SNPs; and Maccaferri et al. (2015) reported a LD decay distance of 1.6 cM ($r^2 = 0.3$) for the whole genome among 1,000 spring wheat genotypes using 4,585 SNPs. In the present study, the LD decay distance ($r^2 = 0.3$) was approximately 1.94 cM for the whole genome and 5.73, 1.42, and 17.48 cM for the A, B, and D genomes, respectively (Figure 2). The LD values differed significantly among the three wheat genomes, whereas the D genome had the greatest LD, similar to that reported in previous studies (Cormier et al., 2014; Edae et al., 2014).

In the present study, the Bonferroni-corrected threshold ($-\log p > 3.782; \alpha = 1$) was used as a cutoff to identify MTAs. Significantly associated loci distributed on 12 chromosomes (Figure 3; Table S1) were detected by both the GLM and MLM in the five environments. The GLM identified more markers than the MLM. One marker was detected in both 12WJ and 14CZ. Two significant markers (tPt-7980 and wPt-6457) were detected by both models (Tables 4 and S1; Figure 3), and the two QTL associations with tPt-7980 and wPt-6457, located on Chr. 1B, explained 37% and 30% of the variation, respectively. Based on the mean and BLUP values, eight significant markers were detected. However, no significant markers were detected in all environments. PHS is a complex genetic trait significantly affected by the environment, and maturation at different times under different climatic conditions can affect seed development and confound the phenotype (Kulwal et al., 2012). Thus, it is difficult to detect significant markers in all tested environmental conditions. In the present study, both tPt-7980 and wPt-6457 were significant in the GLM and MLM. Thus, further studies are needed to investigate the application of the two markers in breeding programs through marker-assisted selection.
reference genome, we assumed that identified in the present study. Due to the incomplete wheat map 3.0, but it was distant from and located on Chr. 1BS based on the Triticarte consensus detected one QTL associated with the marker wPt-6457 (Table 4). The QTL sensitivity map 3.0, a chromosomal position similar to that of Triticarte consensus map 3.0 (Table 4). Kulwal et al. (2009) reported the PHS resistance and divided them into five groups based on the types of encoding proteins. Among these genes, vPlD and AIP2-1 that have been reported to be related to PHS resistance (McKibbin et al., 2002; Gao et al., 2014) were identified using the sequences of significant markers by the GLM, but not by the MLM. The gene vPlD involved in germination hence related to PHS resistance (Bailey et al., 1999; Kulwal et al., 2005; Imtiaz et al., 2008), whereas Chr. 4A contains Phs 1 (Torada et al., 2008). In present study, we found some candidate genes putatively linked to PHS resistance and divided them into five groups based on the distance between the two QTL was less than 1 cM based on the GLM results; b: MLM results; c: Q-Q plots of GLM and MLM.

Based on linkage mapping, Munkvold et al. (2009) detected two QTL (one DArT and one SSR marker) on Chr. 1B in three environments, accounting for 7% and 4% of the variation, respectively. Singh et al. (2014) reported the QTL Qphs.spa-1B associated with the markers tPt-8831, wPt-4603, and wPt-3582 in different environments and located on Chr. 1BS (104.00 cM) based on the Triticarte consensus map 3.0, a chromosomal position similar to that of wPt-6457 (Table 4). The QTL Qphs.spa-1A was also reported by Singh et al. (2014) and might be the same with wPt-668205 that detected in our study, since the distance between the two QTL was less than 1 cM based on the Triticarte consensus map 3.0 (Table 4). Kulwal et al. (2012) detected one QTL associated with the marker wPt-666564 and located on Chr. 1BS based on the Triticarte consensus map 3.0, but it was distant from wPt-6457 and tPt-7980 identified in the present study. Due to the incomplete wheat reference genome, we assumed that wPt-6457 was a previously reported QTL, whereas tPt-7980 might be a novel QTL.

In the present study, the GLM identified more markers than the MLM. The MLM can reduce the probability of false positives and the Type 1 error in association mapping (Yu et al., 2006). The quantile-quantile (Q-Q) plot also showed that the MLM was more conservative than GLM (Figure 3). However, some significant associations may not be detected using only the MLM (false negatives), because they do not pass the false discovery rate criteria. Thus, we run a BLAST search using all the significant markers sequence from both models and listed the best hit genes.

Previous studies revealed that Chr. 3 contains the R genes and the vivipary gene series TaVp1 (orthologs of maize genes) and Vp1, which are involved in germination and are possibly related to PHS resistance (McKibbin et al., 1999; Kulwal et al., 2005; Imtiaz et al., 2008), whereas Chr. 4A contains Phs 1 (Torada et al., 2008). In present study, we found some candidate genes putatively linked to PHS resistance and divided them into five groups based on the types of encoding proteins. Among these genes, vPlD and AIP2-1 that have been reported to be related to PHS resistance (McKibbin et al., 2002; Gao et al., 2014) were identified using the sequences of significant markers by the GLM, but not by the MLM. The gene vPlD involved in germination hence related to PHS resistance (Bailey et al., 1999; Kulwal et al., 2005; Imtiaz et al., 2008). The wheat AIP2s could negatively regulate the ABA signaling pathway and play important roles in seed germination, and thus wheat PHS resistance. In addition, other candidate genes, encoding enzymes, regulatory proteins, protection proteins, transport proteins, and other proteins were identified and might regulate or control PHS resistance. Hoshino et al. (1989) found that a large vernalization requirement delays germination in winter wheat areas where the late wheat is subjected to ear sprouting by monsoon rain. Cao et al. (2016) also detected a QTL associated with PHS resistance on the short arm of Chr. 7B where Vrn-B3 is located. Therefore, vernalization genes may be involved in the regulation of PHS resistance. In the GLM, we also found the significant marker wsnp_CAP11_rep_c6622_304459 located on Chr. 7BS linked to the gene Vrn-3 (Vrn-B3) (Table S1), which may be the same QTL with that reported previously. Overall, our data provided a basis for elucidating the genetic mechanisms of PHS resistance in Chinese wheat founder parents.

**Conclusion**

We performed a genome-wide association study of pre-harvest sprouting resistance among 80 Chinese wheat founder parents using 6,057 markers. The LD decay distances ($r^2 = 0.3$) were approximately 1.94 cM for the whole genome and 5.73, 1.42, 17.48 cM for the A, B, and D genomes, respectively. Two significant marker-trait associations were detected using the GLM and MLM. Twenty-
two candidate genes that might control PHS resistance were identified at or near significant loci. The significantly associated loci identified in this study are potential candidates for cloning genes related to PHS resistance and may be used in wheat breeding programs.

Acknowledgments

This study was financially supported by the National Natural Science Foundation of China (31560388 and 31301317), Outstanding Youth Foundation of the Department of Science and Technology of Sichuan Province (2016JQ0040), Key Technology Research and Development Program of the Department of Science and Technology of Sichuan Province (2016NZ0057), and the International Science & Technology Cooperation Program of the Bureau of Science and Technology of Chengdu (2015-GH03-00008-HZ).

References

Ardlie KG, Kruglyak L and Seielstad M (2002) Patterns of linkage disequilibrium in the human genome. Nat Rev Genet 3:566-566.

Bailey PC, McKibbin RS, Lenton JR, Holdsworth MJ, Flintham JE and Gale MD (1999) Genetic map locations for orthologous Vp1 genes in wheat and rice. Theor Appl Genet 98:281-284.

Banuelos MA, Garcia-deblas B, Cubero B and Rodriguez-Navarro (2015-GH03-00008-HZ).

Bailey PC, McKibbin RS, Lenton JR, Holdsworth MJ, Flintham JE, Adlam R, Bassoi M, Holdsworth M and Gale M (2014) Genome dynamics explain the evolution of flowering time CCT domain gene families in the Poaceae. PLoS One 7:e45307.

Borrell J, Ravel C, Le Gouis J, Lapierre A, Charmet G and Balfourier F (2011) Use of global wheat core collection for association analysis of flour and dough quality traits. J Cereal Sci 54:137-147.

Bordes J, Ravel C, Jaubertte JP, Duperrier B, Gardet O, Heumez E, Pissavy AL, Charmet G, Gouis JE, Balfourier F, et al. (2013) Genomic regions associated with the nitrogen limitation response revealed in a global wheat core collection. Theor Appl Genet 126:805-822.

Breshegello F and Sorrells ME (2006a) Association analysis as a strategy for improvement of quantitative traits in plants. Crop Sci 46:1323-1330.

Breshegello F and Sorrells ME (2006b) Association mapping of kernel size and milling quality in wheat (Triticum aestivum L.) cultivars. Genetics 172:1165-1177.

Brueggeman R, Druka A, Nirmala J, Cavileer T, Drader T, Ros-toks N, Mirlohi A, Gill BU, Kudrna D, Whelaw C, et al. (2008) The stem rust resistance gene Bpg5 encodes a protein with nucleotide-binding-site, leucine-rich, and protein kinase domains. Proc Natl Acad Sci U S A 105:14970-14975.

Cabral AL, Jordan MC, McCartney CA, You FM, Humphreys DG, MacLachlan R and Pozniak CJ (2014) Identification of candidate genes, regions and markers for pre-harvest sprouting resistance in wheat (Triticum aestivum L.). BMC Plant Biol 14:340.

Cao LZ, Hayashi K, Tokui M, Mori M, Miura H and Onishi K (2016) Detection of QTLs for traits associated with pre-harvest sprouting resistance in bread wheat (Triticum aestivum L.). Breeding Sci 66:260-270.

Cavanagh CR, Chao SM, Wang SC, Huang BE, Stephen S, Kiani S, Forrest K, Saintenac C, Brown-Guedira GL, Akhunova A, et al. (2013) Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. Proc Natl Acad Sci U S A 110:8057-8062.

Chalupka D, Lee HY, Faris JD, Evrard A, Chalhoub B, Hasel-korn R and Gornicki P (2008) Acc homeologs and the evo-lution of wheat genomes. Proc Natl Acad Sci U S A 105:9691-9696.

Chao S, Zhang W, Dubcovsky J and Sorrells M (2007) Evaluation of genetic diversity and genome-wide linkage disequi-librium among US wheat (Triticum aestivum L.) germplasm representing different market classes. Crop Sci 47:1018-1030.

Chen CX, Cai B and Bai GH (2008) A major QTL controlling seed dormancy and pre-harvest sprouting resistance on chromosome 4A in a Chinese wheat landrace. Mol Breed 21:351-358.

Chen GY, Liu W, He JY, Gou LL, Yu M, Chen SS, Wei YM and Zheng YL. (2013) Specific loci for adult-plant resistance to stripe rust in wheat founder parent Fan 6 and their genetic dissection in its derivatives. Zhongguo Nong Ye Ke Xue 39:827-836.

Cockram J, Thiel T, Steurnagel B, Stein N, Taudien S, Bailey PC and O’Sullivan DM (2012) Genome dynamics explain the evolution of flowering time CCT domain gene families in the Poaceae. PLoS One 7:e45307.

Cornier F, Le Gouis J, Dubreuil P, Lafarge S and Praud S (2014) A genome-wide identification of chromosome regions determining nitrogen use efficiency components in wheat (Triticum aestivum L.). Theor Appl Genet 127:2679-2693.

Crosa J, Burgueno J, Dreiseckacker S, Vargas M, Herrera-Foessel SA, Lillemo M, Singh RP, Trethowan R, Warburton M, Franco J, et al. (2007) Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. Genetics 177:1889-1913.

Edae EA, Byrne PF, Halem SD, Lopes MS and Reynolds MP (2014) Genome-wide association mapping of yield and yield components of spring wheat under contrasting moisture regimes. Theor Appl Genet 127:791-807.

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.

Flintham JE, Adlam R, Bassoi M, Holdsworth M and Gale M (2002) Mapping genes for resistance to spraying damage in wheat. Euphytica 126:39-45.

Fofana B, Humphreys DG, Rasul G, Cloutier S, Brule-Babel A, Woods S, Lukow Om and Somers DJ (2009) Mapping quantititative trait loci controlling pre-harvest sprouting resistance in a red 9 white seeded spring wheat cross. Euphytica 165:509-521.

Gao DY, Xu ZS, HeY, Sun YW, Ma YZ and Xia LQ (2014) Functional analyses of an E3 ligase gene AIP2 from wheat in
**Arabidopsis** revealed its roles in seed germination and pre-harvest sprouting. J Integr Plant Biol 56:480-491.

Groos C, Gay G, Perretat MR, Gervais L, Bernard M, Dedryver F and Charmet G (2002) Study of the relationship between pre-harvest sprouting and grain color by quantitative trait loci analysis in a white × red grain bread-wheat cross. Theor Appl Genet 104:39-47.

Hatfield PM, Callis J and Vierstra RD (1990) Cloning of ubiquitin-activating enzyme from wheat and expression of a functional protein in Escherichia coli. J Biol Chem 265:15813-15817.

Hoshino T, Tomooka TN, Fukunaga K and Seko H (1989) Testing methods of pre-harvest sprouting and genealogical pedigree of pre-harvest sprouting resistant cultivars in wheat. Jpn J Breed 39:365-372.

Hudakova S, Michalek W, Presting GG, ten Hoopen R, dos Santos K, Jasencazkova Z and Schubert I (2001) Sequence organization of barley centromeres. Nucleic Acids Res 29:5029-5035.

Imtiaz M, Ogbonnaya FC, Osman J and van Ginkel M (2008) Characterization of quantitative trait loci controlling genetic variation for pre-harvest sprouting in synthetic backcross-derived wheat lines. Genetics 178:1725-1736.

Ivanicova Z, Jakobson I and Reis D (2016) Characterization of new allele influencing flowering time in bread wheat introgressed from Triticum turgidum ssp. viviparum. N Biotechnol 33:718-727.

Jaiswal V, Mir RR, Mohan A, Balyan HS and Gupta PK (2012) Association mapping for pre-harvest sprouting tolerance in common wheat (Triticum aestivum L.). Euphytica 188:89-102.

Jiao Y, Zhao H, Ren L, Song W, Zeng B, Guo J, Wang B, Liu Z, Chen J, Li Wei, et al. (2012) Genome-wide genetic changes during modern breeding of maize. Nat Genet 44:812-815.

Ke J, Wen TN, Nikolau BJ and Wurttele ES (2000) Coordinate regulation of the nuclear and plastidic genes coding for the subunits of the heteromeric acetyl-coenzyme A carboxylase. Plant Physiol 122:1057-1072.

Khartak MS, Nicholas FW, Collins AR, Zenger KR, Cavanagh JL, Brar IS, Schnabel RD, Taylor JF and Raudsm AHW (2008) Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. BMC Genomics 9:187.

Kim KH, Jung JU, Kang CS, Shin SH, Park JC, Son JH and Park CS (2014) Mapping of QTL for pre-harvest sprouting resistance and seed dormancy in a doubled haploid population. Plant Breed Biotechnol 2:247-256.

Knox RE, Clarke FR, Clarke JM and Fox SL (2005) Genetic analysis of pre-harvest sprouting in a durum wheat cross. Euphytica 143:261-264.

Knox RE, Clarke FR, Clarke JM, Fox SL, DDePaum RM and Singh AK (2012) Enhancing the identification of genetic loci and transgressive segregants for preharvest sprouting resistance in a durum wheat population. Euphytica 186:193-206.

Kulwal P, Ishikawa G, Benscher D, Feng Z, Yu LX, Jadhav A, Mehetre and Sorrells ME (2012) Association mapping for pre-harvest sprouting resistance in white winter wheat. Theor Appl Genet 125:793-805.

Kulwal PL, Kumar N, Gaur A, Khurana P, Khurana JP, Tyagi AK, Balyan HS and Gupta PK (2005) Mapping of a major QTL for pre-harvest sprouting tolerance on chromosome 3A in bread wheat. Theor Appl Genet 111:1052-1059.

Kulwal PL, Mir RR, Kumar S and Gupta PK (2010) QTL analysis and molecular breeding for seed dormancy and pre-harvest sprouting tolerance in bread wheat. J Plant Biol 37:59-74.

Kumar S, Knox RE, Clarke FR, Pozniak CJ, DePaum RM, Cuthbert RD and Fox S (2015) Maximizing the identification of QTL for pre-harvest sprouting resistance using seed dormancy measures in a white-grained hexaploid wheat population. Euphytica 205:1-23.

Li Q, Wang CY, Liu XL, Gao D and Ji W (2008) Genetic diversity of Xiaoyan 6 and its derivatives by SSR. Mai Lei Zuo Wu Xue Bao 28:950-955.

Li XJ, Xu X, Yang XM, Li XQ, Liu WH, Gao AN and Li LH (2012) Genetic diversity among a founder parent and widely grown wheat cultivars derived from the same origin based on morphological traits and microsatellite markers. Crop Pasture Sci 63:303-310.

Li Z, Rahman S, Kosar-Hashemi B, Mouille G, Appels R and Morell MK (1999) Cloning and characterization of a gene encoding wheat starch synthase I. Theor Appl Genet 98:1208-1216.

Lin F, Xue SL, Zhang ZZ, Zhang CQ, Kong ZX, Yao GQ, Tian DG, Zhu HL, Li CJ, Cao Y, et al. (2006) Mapping QTL associated with resistance to Fusarium head blight in the Nan-da 2419 × Wangshuibai population: II. Type I resistance. Theor Appl Genet 112:526-535.

Li M, Cai S, Wang S, Liu S, Zhang G and Bai G (2015) Genotyping-by-sequencing (GBS) identified SNP tightly linked to QTL for pre-harvest sprouting resistance. Theor Appl Genet 128:1385-1395.

Liu S, Bai G, Cai S and Chen C (2011) Dissection of genetic components of preharvest sprouting resistance in white wheat. Mol Breed 27:511-523.

Ma Z, Zhao D, Zhang C, Zhang Z, Xue S, Lin F, Kong Z, Tian D and Luo Q (2007) Molecular genetic analysis of five spike-related traits in wheat using RIL and immortalized F2 populations. Mol Gen Genomics 277:31-42.

Maccaferri M, Zhang J, Bulli P, Abate Z, Chao S, Cantu D, Bossolini E, Chen X, Pumphrey M, Dubcovsky J, et al. (2015) A genome-wide association study of resistance to stripe rust (Puccinia striiformis f. sp. triticci) in a worldwide collection of hexaploid spring wheat (Triticum aestivum L.). G3 (Bethesda) 5:449-465.

Mares DI and Mrva K (2001) Mapping quantitative trait loci associated with variation in grain dormancy in Australian wheat. Aust J Agric Res 52:1257-1265.

McKibbin RS, Wilkinson MD, Bailey PC, Flintham JE, Andrew LM, Lazzeri PA, Gade ML, Denton JR and Holdsworth (2002) Transcripts of Vp-1 homeologues are misspliced in modern wheat and ancestral species. Proc Natl Acad Sci U S A 99:10203-10208.

Munkvold JD, Tanaka J, Benscher D and Sorrells ME (2009) Mapping quantitative trait loci for preharvest sprouting resistance in white wheat. Theor Appl Genet 119:1223-1235.

Nakamura S, Abe F, Kawahigashi H, Nakazono K, Tagiri A, Matsumoto T, Utsugi S, Ogawa T, Handa H, Ishida H, et al. (2011) A wheat homolog of mother of FT and TFL1 acts in the regulation of germination. Plant Cell 23:3215-3229.

Lin et al. 627
Neumann K, Kobiljski B, Dencie S, Varshney RK and Borner A (2007) Genome-wide association mapping: A case study in bread wheat (*Triticum aestivum* L.). Mol Breeding 27:37-58.

Pasam RK, Sharma R, Malosetti M, van Eeuwijk FA, Haseneyer G, Kilian B and Graner A (2012) Genome-wide association studies for agronomical traits in a worldwide spring barley collection. BMC Plant Biol 12:16-37.

Piepho HP, Mühring J, Melching AE and Buchse A (2008) BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161:209-228.

Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-995.

Rasul G, Humphreys DG, Brule-Babel A, McCartney CA, Knox Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-995.

Piepho HP, Mühring J, Melching AE and Buchse A (2008) BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161:209-228.

Rasul G, Humphreys DG, Brule-Babel A, McCartney CA, Knox Pritchard JK, Stephens M and Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-995.
Supplementary material
The following online material is available for this article:
Table S1 – Significant markers in GLM and MLM
Figure S1 – Manhattan plots for 12WJ
Figure S2 – Manhattan plots for 12YA
Figure S3 – Manhattan plots for 13WJ
Figure S4 – Manhattan plots for 13YA
Figure S5 – Manhattan plots for 14CZ
Figure S6 – Manhattan plots for BLUPs
Figure S7 – Manhattan plots for mean values

Associate Editor: Everaldo Gonçalves de Barros

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License (type CC-BY), which permits unrestricted use, distribution and reproduction in any medium, provided the original article is properly cited.