Theor Appl Genet (2015) 128:2203–2212
DOI 10.1007/s00122-015-2578-x

ORIGINAL ARTICLE

Genome-wide association study of phosphorus-deficiency-tolerance traits in *Aegilops tauschii*

Yaxi Liu[1] · Lang Wang[1] · Mei Deng[1] · Zhanyi Li[1] · Yanli Lu[2] · Jirui Wang[1] · Yuming Wei[1] · Youliang Zheng[1]

Received: 6 February 2015 / Accepted: 1 July 2015 / Published online: 18 July 2015

Abstract

Key message **Using GWAS, 13 significant SNPs distributed on six of the seven *Aegilops tauschii* chromosomes (all but 5D) were identified, and several candidate P-deficiency-responsive genes were proposed from searches of public databases.**

Abstract *Aegilops tauschii*, the wheat (*Triticum aestivum*) D-genome progenitor, possesses numerous genes for stress resistance, including genes for tolerance of phosphorus (P) deficiency. Investigation of the genetic architecture of *A. tauschii* will help in developing P-deficiency-tolerant varieties of wheat. We evaluated nine traits in a population of 380 *A. tauschii* specimens under conditions with and without P application, and we performed genome-wide association studies for these traits using single nucleotide polymorphism (SNP) chips containing 7185 markers. Using a general linear model, we identified 119 SNPs that were significantly associated with all nine traits, and a mixed linear model revealed 18 SNPs associated with all traits. Both models detected 13 significant markers distributed on six of the seven *A. tauschii* chromosomes (all but 5D). Searches of public databases revealed several candidate/flanking genes related to P-deficiency tolerance. These genes were grouped in five categories by the types of proteins they encoded: defense response proteins, enzymes, promoters and transcription factors, storage proteins, or proteins triggered by P deficiency. The identified SNPs and genes contain essential information for cloning genes related to P-deficiency tolerance in *A. tauschii* and wheat, and they provide a foundation for breeding P-deficiency tolerant wheat cultivars.

Abbreviations

AM	Association mapping
AP	Applied phosphorus
BLAST	Basic local alignment search tool
CV	Coefficient of variation
DH	Double haploid
DNA	Deoxyribose nucleic acid
GLM	General linear model

Communicated by M. J. Sillanpaa.

Y. Liu and L. Wang contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-015-2578-x) contains supplementary material, which is available to authorized users.

✉ Youliang Zheng
ylzheng@sicau.edu.cn

Yaxi Liu
yaxi.liu@hotmail.com

Lang Wang
wlsuccess@foxmail.com

Mei Deng
183433899@qq.com

Zhanyi Li
lizhanyi2008@163.com

Yanli Lu
luyanli@sicau.edu.cn

Jirui Wang
wangjirui@gmail.com

Yuming Wei
yuming71@hotmail.com

[1] Triticeae Research Institute, Sichuan Agricultural University, Wenjiang, Chengdu 611130, China

[2] Maize Research Institute, Sichuan Agricultural University, Wenjiang, Chengdu 611130, China

GWAS	Genome wide association study
IWGSC	International Wheat Genome Sequencing Consortium
LD	Linkage disequilibrium
MAF	Minor allele frequency
MCMC	Markov Chain Monte Carlo
MLM	Mixed linear model
MTAs	Marker trait associations
NAP	No-applied phosphorus
NCBI	National Center for Biotechnology Information
P	Phosphorus
PDTI	Phosphorus deficiency tolerance index
PTs	Phosphate transporters
QTL	Quantitative trait loci R/S, root to shoot ratio
RDM	Root diameter
RDW	Root dry weight
RF	Root forks
RIL	Recombinant inbred line
RL	Root length
RSA	Root surface area
RT	Root tips
SDW	Shoot dry weight
SNP	Single nucleotide polymorphism
TDW	Total dry weight

Introduction

Phosphorus (P) is an important macronutrient for crop production; P is necessary for plant growth and metabolic processes including nucleic acid synthesis, glycolysis, photosynthesis, respiration, carbohydrate metabolism, and redox reactions (Marschner 1995). The total amount of P in soils is generally adequate to support plant growth, but P is largely unavailable for uptake by plants because it readily complexes with metal cations (e.g., aluminum and iron in acid soils and calcium and magnesium in alkaline soils), which leads to a deficiency of available P in soils (Holford 1997). P deficiency is a major limitation for plant growth worldwide (Lynch 2007; Vance et al. 2003). Fertilizers are commonly applied to alleviate this deficiency; however, fertilization is not a comprehensive solution because of limited bioavailability, potential environmental problems such as eutrophication, and the exhaustion of resources of phosphate rock (Vance et al. 2003; Yang et al. 2010). An alternative solution that is effective and more sustainable is the production of P-deficiency-tolerant cultivars (Fageria et al. 2008; Su et al. 2006).

Wheat (*Triticum aestivum*) is a staple food for much of the global population. Wheat yield has been severely restricted by P deficiency in arable areas in China and worldwide (Dai et al. 2010; Zhang et al. 2005), and development of low-P-tolerant wheat is critical. Wild genotypes represent a potential source of novel genes and promoters, and thus are an important resource for wheat improvement. There is significant genetic variation for low-P tolerance among wild relatives of wheat (Farooq 2009). *Aegilops tauschii,* the diploid progenitor of the D genome of hexaploid wheat, can be used to improve wheat resistance to P deficiency. The usefulness of *A. tauschii* arises from the close relationship between its genome and the D genome because of the recent origin of hexaploid wheat (Padulosi et al. 1996). The *A. tauschii* genome exhibits great genetic diversity, with numerous genes for stress resistance that can provide a valuable resource in crop breeding (Schneider et al. 2008). *A. tauschii* genes can be incorporated into the wheat genome through wide hybridization, specifically via bridge crosses (tetraploid wheat × *A. tauschii*) that lead to generation of synthetic hexaploids upon colchicine treatment or by spontaneous induction (Mujeeb-Kazi et al. 1996; Schneider et al. 2008). Therefore, understanding the genetic basis of tolerance to P deficiency in *A. tauschii* can contribute to the breeding of tolerant wheat.

Most of the traits related to low P tolerance are quantitative, which makes it difficult to identify the genetic differences that underlie the phenotype of interest. Linkage analysis is currently the most common approach for detecting quantitative trait loci (QTL) that correspond to complex traits in plants. Classical linkage mapping using bi-parental crosses is limited to revealing information about two alleles at a given locus or a few loci segregating in a population (Navakode et al. 2014). In addition, the genetic resolution of detected QTLs is poor (10–30 cM) where limited numbers of recombination events occur during the development of the mapping population (Flint-Garcia et al. 2003; Holland 2007). Furthermore, linkage analysis can sample only a small fraction of possible alleles in the parental population, and it is costly and time-consuming to develop mapping populations.

Alternatively, association mapping (AM) or linkage disequilibrium (LD) mapping, relies on existing natural populations or designed populations of plants to overcome the constraints inherent to linkage mapping (Pasam et al. 2012). AM has emerged as a powerful tool for resolving complex variation in traits and for identifying different loci or novel and superior alleles in natural populations (Zhu et al. 2008). LD mapping was first introduced in genetic mapping studies in humans (Hästbacka et al. 1992; Lander and Schork 1994) but has subsequently been considered for plant research. LD mapping identifies ancestral recombination events and takes into account all major alleles present in the population to identify significant marker–phenotype associations. It is possible to identify significantly associated genomic regions using a suite of mapped markers by

exploiting non-random associations of alleles at nearby loci (Mackay and Powell 2007). Successful mapping depends on the quality of phenotypic data, population size, and the degree of LD in a population (Flint-Garcia et al. 2005). Generally, the power of association studies depends on the degree of LD between genotyped markers and functional polymorphisms. The decay of LD varies across species, among intraspecific populations, and among loci within a given genome (Caldwell et al. 2006; Gupta et al. 2005; Tenaillon et al. 2001).

Genome-wide association, which exploits marker polymorphisms across all chromosomes, is one of the strategies on which LD mapping is based (Hirschhorn and Daly 2005). Genome-wide association studies (GWAS) have become increasingly popular and powerful in human and animal genetics, and the emergence of high-throughput, cost-effective genotyping platforms has made AM an attractive approach for QTL mapping in plants (Atwell et al. 2010). GWAS based on analyses of candidate genes have examined flowering-time genes in barley (Stracke et al. 2009), *Dwarf8* and the phytoene synthase locus in maize (Palaisa et al. 2003; Thornsberry et al. 2001), the *rhg-1* gene in soybean (Li et al. 2009), the Psy1-A1 locus in wheat (Singh et al. 2009), and many candidate genes in *Arabidopsis* (Ehrenreich et al. 2009; Zhao et al. 2007), In wheat, different association panels have been used in AM studies to identify loci that control agronomic traits (Bordes et al. 2013; Breseghello and Sorrells 2006; Crossa et al. 2007; Neumann et al. 2011) and quality traits (Bordes et al. 2011; Ravel et al. 2009).

Hence, few GWAS of P-deficiency-tolerant traits have used *Aegilops tauschii*. We aimed to investigate marker–trait associations for P-deficiency tolerance based on a whole-genome AM approach using single nucleotide polymorphism (SNP) markers in a core collection of 380 *A. tauschii* accessions of diverse origin. Additional objectives were to identify SNP markers that are highly associated with P-deficiency-tolerant traits, and to scan candidate gene responses to phosphorus deficiency. These identified genes and SNPs will provide important information for cloning P-deficiency-tolerance-related genes in *A. tauschii*.

Materials and methods

Plant materials and phenotypic evaluation

The natural population used for the association analyses included 380 *Aegilops tauschii* accessions collected by the Triticeae Research Institute of Sichuan Agricultural University (SAU). The *A. tauschii* were grown in a phytotron in Wenjiang, Sichuan Province, China, from September 2012 to March 2013. Hydroponic experiments with two phosphorus conditions, applied phosphorus (AP) and no-applied phosphorus (NAP), were performed using a completely randomized design with four replications per treatment. The AP and NAP treatments contained Hoagland's nutrient solution (Hoagland and Arnon 1950) modified to optimize growth of *A. tauschii*, with and without $NH_4H_2PO_4$ (1 mmol/L), respectively. The modified Hoagland's nutrient solution consisted of $Ca(NO_3)_2{\cdot}4H_2O$ (4 mmol/L), KNO_3 (6 mmol/L), $MgSO_4{\cdot}7H_2O$ (2 mmol/L), H_3BO_3 (46 μmol/L), Na·Fe·EDTA (100 μmol/L), $MnCl_2$ (9.146 μmol/L), $ZnSO_4$ (0.76 μmol/L), $CuSO_4$ (0.32 μmol/L), and $(NH_4)6Mo_7O_{24}$ (0.0161 μmol/L).

Seeds were transplanted in the phytotron 8 days after germination; after 26 days, plants were evaluated using a WinRhizo Pro 2008a image analysis system (Régent Instr. Inc., Quebec, Canada) to obtain the following characteristics: root length (RL), root diameter (RDM), root surface area (RSA), root tips (RT), and root forks (RF). The plants were then separated into shoots and roots, stored in paper bags, heated at 105 °C for 30 min to kill living cells, and dried at 75 °C until constant mass to obtain root dry weight (RDW), shoot dry weight (SDW) and total dry weight (TDW). Root to shoot ratio (R/S), phosphorus deficiency tolerance index (PDTI) and *S*-value are also used as indicators of tolerance to P deficiency (Table 1). RSA, RT, and RF values covered a broad range, so we log-transformed these values for analysis. Descriptive analyses, analysis of variance, and correlation analyses were performed for each trait using SPSS for Windows, Version 20.0 (IBM Corp., 2011). The estimate of heritability (*H*) is defined by the formula $H = VG/(VG + VE)$, where *VG* and *VE* represent estimates of genetic and environmental variance, respectively (Smith et al. 1998).

10K Infinium iSelect SNP array and SNP genotyping

Construction of the *Aegilops tauschii* 10 K SNP array has been reported in detail (Luo et al. 2013). A total of 7185 polymorphic SNP markers in the array were uniquely mapped on the *A. tauschii* genetic map and on the physical map of the *A. tauschii* genome built from bacterial artificial chromosome clones (Luo et al. 2013). SNPs were assayed according to the manufacturer's protocol (Illumina, San Diego, CA) at the Genome Center, University of California, Davis. Normalized Cy3 and Cy5 fluorescence for each DNA sample was graphed with the GenomeStudio program (Illumina), resulting in genotype clustering for each SNP marker. Details for SNP genotyping of 380 *A. tauschii* accessions are provided in our previous study (Wang et al. 2013).

Population structure

The Bayesian inference program Structure 2.3.3 (Falush et al. 2007; Pritchard et al. 2000) was used to assess

Table 1 The tested traits used in the study

Abbreviated name	Full name	Description	Unit
RL	Root length	The average root length of four plants	cm
RDW	Root dry weight	The average root dry weight of four plants	g
SDW	Shoot dry weight	The average shoot dry weight of four plants	g
RDM	Root diameter	The average root diameter of four plants	mm
RSA	Root surface area	The average root surface of four plants	cm^2
RT	Root tips	The average number of root tips of four plants	count
RF	Root forks	The average number of root forks of four plants	count
TDW	Total dry weight	The total weight of RDW and SDW	g
R/S	Root to shoot ratio	The ratio of root dry weight to shoot dry weight	\
PDTI	Phosphorus deficiency tolerance index	The ratio of traits under no-applied phosphorus condition to traits under applied phosphorus	\
S value	*S* value	An indicator to measure tolerant ability of phosphorus deficiency	\

population structure with the 7185 polymorphic SNP markers (Luo et al. 2013) mapped on the *A. tauschii* genetic map. The linkage ancestry and correlated allele frequency models were used. A total of 100 burn-in iterations followed by 100 Markov Chain Monte Carlo (MCMC) iterations for $K = 1$–10 clusters were used to identify the optimal range of K. For each K, five independent runs were produced. The optimal value of K was determined using the delta-K method (Evanno et al. 2005). Here, $K = 4$ was used, and the whole panel was divided into Subp1, Subp2, Subp3, and Subp4 based on our previous research (Wang et al. 2013).

Association analysis

Marker-trait associations (MTAs) were calculated in Tassel 2.1 using a general linear model (GLM) and mixed linear model (MLM). The models included a GLM with Q-matrix and an MLM with Q-matrix and kinship (K)-matrix as correction for population structure, and 6905 SNP markers with MAF >0.05. Bonferroni-corrected thresholds at $\alpha = 1$ were used as the cutoffs (Yang et al. 2014). When the number of markers was 6905 SNPs, at $\alpha = 1$, the Bonferroni-corrected thresholds for the p values were 145×10^{-6}, with corresponding $-\log p$ values of 3.839. Significant markers are shown with a Manhattan plot drawn in the R Language and Environment for Statistical Computing (R version 3.0.3; http://www.r-project.org/). Important p value distributions (observed p values against cumulative p values, negative $\log_{10}$ scale) are shown with a quantile–quantile plot drawn in R. To find candidate/flanking genes or trait-related proteins, we blasted the National Center for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/) using the IWGSC BLAST results, and then we did a direct BLASTx search of the NCBI database (Table S1).

Results

Phenotypic evaluation

Significant phenotypic variation was observed for all nine traits (see Fig. S1, Additional file 2). The means of all traits were significantly different between the AP and NAP treatments. The mean values of RT, RF and R/S were lower under AP than under NAP conditions. RL, RDW, SDW, RDM, RSA, and TDW were lower under NAP conditions (Table 2). The coefficient of variation (CV) ranged from 5 to 36 % under AP condition and from 13 to 41 % under NAP condition (Table 2). Differences between AP and NAP were significant for all traits except for RDM and R/S, indicating that the tested traits were significantly affected by P supply. Heritability ranged from 55.0 to 81.8 % for AP conditions and from 68.5 to 90.5 % for NAP conditions. Heritability of all traits was generally higher under NAP (Table 2).

Marker-trait association analysis

The Bonferroni-corrected threshold ($-\log p > 3.839$, $\alpha = 1$) was used as a cutoff to identify MTAs. A total of 119 and 18 significant markers were detected by GLM and MLM, respectively. Of these, 13 were detected by both models (Table 3; Fig. S2). The significant SNPs detected by both models were distributed on six of seven chromosomes (all but 5D) (Table S1; see Additional file 1).

Under AP conditions, five significant markers were detected by GLM for traits R/S and TDW, and the MLM detected two significant markers for each of two traits. The r^2 values provide an estimate of phenotypic variation explained by the markers, which ranged from 4.32 to 5.34 % (Table 3). The GLM detected 10, 1, 32, 4, 4, and 7

Table 2 Phenotypic variation in *Aegilops tauschii* traits under applied phosphorus (AP) and no-applied phosphorus (NAP) conditions

Trait	Unit	Condition	Mean ± SD	CV (%)	*F* value[a]	h_B^b (%)
RL	cm	AP	16.8 ± 4.8	29	413.59**	61.8
		NAP	10.4 ± 3.8	37		74.0
RDW	g	AP	0.029 ± 0.010	34	145.00**	80.1
		NAP	0.018 ± 0.006	35		90.0
SDW	g	AP	0.07 ± 0.02	32	171.61**	79.6
		NAP	0.04 ± 0.01	32		86.8
RDM	mm	AP	0.9 ± 0.3	30	1.28[ns]	55.0
		NAP	0.8 ± 0.3	35		68.5
RSA[c]	cm^2	AP	1.7 ± 0.5	27	18.46**	75.3
		NAP	1.6 ± 0.4	24		83.8
RT[c]	count	AP	2.5 ± 0.4	17	13.93**	65.9
		NAP	2.5 ± 0.6	23		75.6
RF[c]	count	AP	3.1 ± 0.4	13	5.64*	73.3
		NAP	3.1 ± 0.5	16		79.7
TDW	g	AP	0.10 ± 0.03	27	176.98**	81.8
		NAP	0.05 ± 0.02	31		90.5
R/S	/	AP	0.4 ± 0.1	36	0.06[ns]	72.5
		NAP	0.5 ± 0.2	41		84.1

R/S root-to-shoot ratio, *RDM* root diameter, *RDW* root dry weight, *RF* root forks, *RL* root length, *RSA* root surface area, *RT* root tips, *RV* root volume, *SDW* shoot dry weight, *TDW* total dry weight

[a] * and **: significant at $p < 0.05$ and 0.01, respectively; ns: not significant

[b] Broad-sense heritability of the tested traits

[c] Traits RSA, RT and RF were translated into logarithmic transformation

significant markers for RDW, RF, RL, RSA, RT, and SDW, respectively; no significant markers were detected for these traits by MLM (Table 3).

Under NAP conditions, one significant marker for SDW and two for TDW were detected by GLM only. Significant markers for R/S, RDM, RL, and RSA were detected by both models (Table 3). No significant markers for RT, RF, or RDW were detected by either GLM or MLM under NAP conditions.

The PDTI and S-value are used as indicators of tolerance to P deficiency. Twelve SNPs significantly associated with the two indicators were revealed by GLM, and 10 were revealed by MLM; a relatively large amount of phenotypic variation (from 5.01 to 5.82 %) in PDTI was explained by the markers (Table 3).

We performed a BLAST search of the International Wheat Genome Sequencing Consortium (IWGSC, http://www.wheat-genome.org/) using the SNP sequences. The chromosomal locations were different from the best hits in the IWGSC. For example, the SNP marker *GDRF1KQ02FTOCX_181* was located on 1D in a genetic map constructed by Luo et al. (2013), but we found it located on 4AS according to the IWGSC BLAST results. Similarly, *GBUVHFX01B-GL2Y_284*, *GBF1XID01ECUNB_109*, *BG607326ATwsnp2*, *GBUVHFX02HETOV_295*, *GA8KES401CAC7D_237*, and *GDRF1KQ02IR7MF_260* were located on 1D, 3D, 3D, 3D, 6D, and 6D, respectively, in the genetic map; however, IWGSC BLAST identified their locations as 4AL, 3B, 3B, 3B, 6BS, and 6AS, respectively.

Putative candidate genes associated with significant loci

The putative/flanking genes associated with significant loci are listed in Table S1. Noteworthily, we identified a few candidate phosphorus-deficiency-response genes, each of which was associated with different traits. Fox example, candidate gene *TaAP2-B* was associated with traits RDM-PDIT and R/S-PDTI. We also found a candidate vernalization-requirement gene, *VRN2*, suggesting that vernalization may be related to P-deficiency tolerance.

The identified candidate genes could be roughly divided into five groups according to the types of proteins they encoded. The first group consisted of genes including *putative alliin lyase* and *Pm3*. The second group of candidate genes included *PMM-D1*, *Acc-1*, *LR34*, and *cytochrome P450*. Genes in the third group included *WM4, WM5A*, *WM5B*, *TaAP2-B*, *TaAP2-D*, and *m7*. The fourth group only included *Glo-2* and *Glu-Dty*. The last group consisted of *ASY1*, *Tamyb10-D1*, *pdil7-2*, *Rht-A*, *Wknox1b*, and *H2A*.

Table 3 Genome-wide association study of *Aegilops tauschii* traits

Trait		GLM					MLM					No. shared[c]
		No. sig[a]	Average −log(*P*)	Range −log(*P*)	Average R^2 (%)[b]	Range R^2 (%)[b]	No. sig[a]	Average −log(*P*)	Range −log(*P*)	Average R^2 (%)[b]	Range R^2 (%)[b]	
AP	R/S	5	4.17	3.87–4.59	4.60	4.22–5.11	2	4.29	4.29–4.29	5.23	5.23–5.23	2
	RDW	10	4.97	3.87–5.75	5.34	4.04–6.33						
	RF	1	4.49		5.24							
	RL	32	4.61	3.88–7.31	4.57	3.76–7.49						
	RSA	4	4.23	4.01–4.53	4.92	4.53–5.35						
	RT	4	4.08	3.91–4.37	4.71	4.51–4.97						
	SDW	7	4.14	3.97–4.51	4.32	4.11–4.77						
	TDW	5	4.18	4.06–4.21	4.45	4.31–4.49	2	3.90	3.85–3.95	4.63	4.56–4.70	
NAP	R/S	8	4.22	3.84–5.18	4.86	4.39–6.11	2	4.12	4.09–4.15	5.02	4.97–5.06	1
	RDM	4	4.22	4.00–4.35	4.80	4.52–4.96	1	3.96		4.77		1
	RL	24	4.55	3.88–6.05	4.96	4.14–6.75	1	3.89		4.62		1
	SDW	1	4.22		4.47							
	TDW	2	3.97	3.91–4.02	4.18	4.12–4.25						
PDTI	R/S	2	5.10	4.52–5.69	5.82	5.09–6.54	2	4.20	4.09–4.31	5.01	4.86–5.16	2
	RDM	4	4.83	3.91–5.48	5.46	4.31–6.26	6	4.44	3.85–5.17	5.38	4.54–6.41	4
	RF	2	4.20	4.17–4.22	5.28	5.16–5.39	1	3.86		5.09		1
	RSA	3	4.58	4.58–4.58	5.33	5.33–5.33						
S value		1	4.81		5.66		1	4.65		6.07		1
Total		119					18					13

AP applied phosphorus conditions, *NAP* no-applied phosphorus conditions, *GLM* general linear model, *MLM* mixed linear model

[a] The total number of significantly associated SNPs detected by GLM and MLM at the threshold of –log 10(p) = 3.84

[b] R^2 value showing the percentage of phenotypic variation explained

[c] The number of significant SNPs detected by both models

Discussion

Loci controlling P-deficiency-tolerant traits

Here, we reported the outcome of a whole-genome AM approach for identifying genomic regions that condition responses to AP and NAP in *Aegilops tauschii* by using 7185 SNP markers genotyped in a core collection of 380 natural *A. tauschii* accessions. Linkage mapping can also detect QTL using different segregating populations tested in different environments. Although there are few reports about QTL related to P-tolerant traits in *A. tauschii*, numerous QTL related to such traits were detected in wheat by linkage mapping. Comparing these QTL to our results will help us to identify loci that control P-deficiency tolerance.

Su et al. (2006) identified 39 QTL related to shoot dry weight, tiller number, shoot P uptake, and shoot P-utilization-efficiency traits under P-deficient and P-sufficient conditions; those loci were distributed on 21 chromosomal regions. In addition, *Xgwm271.2* and *Xgwm121* were closely linked with vernalization requirement genes (*VRN-A1* on chromosome 5A and *VRN-D1* on chromosome 5D). We identified a candidate gene, VRN2, at significant locus *contig02393_383*, which was located on 4D (76 cM) and was associated with trait R/S-PDTI. This indicates that vernalization-requirement genes probably affect P-deficiency tolerance in wheat. These findings suggest the importance of exploring the relationship between P-deficiency tolerance and vernalization requirement genes. Cao et al. (Cao et al. 2001) identified QTLs for P-use efficiency using an F7 recombinant inbred line (RIL) population; they found that a region between *Xfba354* and *Xfba69* on chromosome 7A was important and that 5A and 2D were vital for controlling P-use efficiency. We identified three significant loci related to trait R/S-NAP on chromosome 2D, indicating that 2D may be an important chromosome for P-deficiency tolerance in wheat. Li et al. (2007) identified 30 QTL distributed on 14 chromosomes and related to root responses to P deficiency; among these, three loci on chromosome 7D were related to traits RDW and RL. By comparison, in our study, significant loci were identified on chromosome 7D that were related to traits RDM-PDIT. Su et al. (2009) identified 36 QTL for P-use efficiency distributed on 14 chromosomes and related to shoot dry weight, biomass yield, grain characteristics, P uptake, and P-utilization efficiency.

Among these, chromosomes 2D, 6D, and 7D were detected as controlling biomass yield; we were unable to achieve the same results.

Significantly associated loci distributed on all seven chromosomes were detected by GLM and MLM, and some traits were associated with multiple chromosomes (Table S1). For example, RDM–PDTI was associated with SNPs on 1D and 7D. Massman et al. (2011) suggested that significant SNP markers separated by less than 5 cM could be considered part of one QTL. Markers *GDRF1KQ02FTOCX_181*, *GA8KES4 02GIF8J_42*, and *GBUVHFX01BGL2Y_284* (chromosome 1D) were separated by less than 1 cM and could be considered a QTL related to RDM–PDTI. Similarly, *GA8KES4 01CAC7D_237* and *GDRF1KQ02IR7MF_260* were related to R/S–AP and also can be considered as a QTL (Table S1).

Until the map of the wheat genome is completed, the loci identified here as being associated with P-tolerance traits cannot be directly compared with QTL reported by other researchers. In addition, because the genome of *A. tauschii* is very close but not identical to the D-genome of wheat, only approximate chromosomal locations that control P-deficiency-tolerant traits can be inferred.

In diploid *A. tauschii*, the marker *GDRF1KQ02FTOCX_181* was located on chromosome 1D, but it was on chromosome 4AS in hexaploid wheat. Similarly, marker *GBUVHFX01 BGL2Y_284* was on 1D in *A. tauschii* but on 4AL in wheat; *GA8KES401CAC7D_237* was on 6D in *A. tauschii* but on 6BS in wheat; and *GDRF1KQ02IR7MF_260* was on 6D in *A. tauschii* but on 6AS in wheat. In view of the present technical conditions, mistakes in allocation might have been made by one group or the other.

Putative candidate/flanking gene analysis

In response to deficiency of available P in the rhizosphere, plants have developed specialized morphological, physiological, and biochemical adaptations to modify the rhizosphere and increase the ability of their root systems to utilize inorganic P (Pi) (Hinsinger et al. 2009; Vance et al. 2003). The genes controlling P uptake or utilization are usually responsive to P deficiency. At the molecular level, P is acquired by plant roots through phosphate transporters (PTs) in the cell membranes of epidermal cells and root hairs. Pi is loaded into the apoplastic space of the xylem and transported to shoots, and the process is also mediated by PTs (Marschner 1995). A number of PT genes have been well characterized, and putative high-affinity PTs have been identified, including nine genes (*AtPht1;1–AtPht1;9*) in *Arabidopsis* and 13 genes (*OsPT1–OsPT13*) in rice (Misson et al. 2004; Paszkowski et al. 2002). These Pht1 family members are generally induced by low-P stress and are involved in mediating P acquisition under P deficiency (Ai et al. 2009; Vance et al. 2003). In rice, Wissuwa et al. (Wissuwa and Ae 2001; Wissuwa et al. 1998) identified a major QTL (*Pup1*) for P-deficiency tolerance in the traditional *aus*-type rice variety Kasalath. Gamuyao et al. (2012) found a *Pup1*-specific protein kinase gene and named it phosphorus-starvation tolerance 1 (*PSTOL1*). Subsequent research showed that *PSTOL1* enhances early root growth, thereby enabling plants to acquire more P and other nutrients. Only a few PT members have been characterized in wheat; these include *TaPht2;1*, *TaPht1;4*, and *TaPht2* (Guo et al. 2013, 2014; Liu et al. 2013). Because of the very large (40-fold larger than rice), highly repetitive hexaploid genome of wheat, and the complex and quantitative nature of P-related traits, there are few reports on putative genes controlling P-deficiency tolerance in wheat or *A. tauschii*.

We identified a few putative candidate genes associated with phenotypic traits under P-deficient conditions and divided these genes into five groups based on the types of proteins they encoded. Group 1 encoded defense-response proteins under P-deficient conditions. These proteins are very important under P stress, as demonstrated by previous studies (Cantu et al. 2013; Li et al. 2008). Group 2 encoded various enzymes, suggesting that numerous biochemical pathways were activated under P deficiency. These enzymes might regulate P content and accelerate reuse of P in plants. Group 3 encoded transcription factors and promoters and might facilitate synthesis of response proteins under P stress. Group 4 encoded storage proteins and thus are likely to be necessary in *A. tauschii* for P uptake under P-deficient conditions. This would enable absorption of P from nutrient stores, in accordance with previous reports on soybean, in which storage proteins were regulated in response to P stress (Tang et al. 2001). Group 5 candidate genes controlled diverse metabolic functions in shoot apical meristems and tissue proteins, including asynapsis, proanthocyanidin synthesis in testa, and disulfide bond formation and isomerization during protein folding. These genes are probably triggered by P deficiency and play a crucial role in P-deficiency tolerance in *A. tauschii*. These new findings reflect the complexity of P-deficiency-tolerance mechanisms and the large numbers of genes involved in these mechanisms. These data provide a basis for dissecting the genetic mechanism of P-deficiency tolerance in *A. tauschii* and further in wheat.

Conclusion

We performed genome-wide association studies for phosphorus-deficiency-tolerance traits in a population containing 380 *Aegilops tauschii* accessions using 7185 polymorphic SNP markers. Twenty-seven significant markers were detected by both GLM and MLM. At significant loci and flanking regions, we identified candidate

genes for response to phosphorus-deficiency, including defense response proteins, enzymes, promoters and transcription factors, storage proteins, and proteins that might be triggered by phosphorus deficiency. The identified SNPs and genes offer information for cloning genes related to phosphorus-deficiency tolerance in *A. tauschii* and wheat and provide a foundation for breeding phosphorus-deficiency-tolerance wheat cultivars.

Author contribution statement YL partially designed this study and helped to conduct data analysis and revise the manuscript. LW carried out the phenotypic evaluation and drafted the manuscript. MD and ZL participated in the phenotypic evaluation. YL performed the association mapping and sequence alignment. JW participated in the design of the study and performed the statistical analysis. YW participated in the design of the study. YZ coordinated the study and helped to draft the manuscript. All authors have read and approved the final manuscript.

Acknowledgments We thank Assad Siham (ICARDA, Syria), Jon W. Raupp (Kansas State University, USA), Shuhei Nasuda (Komugi, Japan) and Harold Bockelman (USDA, USA) for plant materials. This work was supported by the International Science and Technology Cooperation Program of China (No. 2015DFA30600), the National Basic Research Program of China (2014CB147200), and the National Natural Science Foundation of China (31301317, 31171555).

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

Ai P, Sun S, Zhao J, Fan X, Xin W, Guo Q, Yu L, Shen Q, Wu P, Miller AJ (2009) Two rice phosphate transporters, OsPht1; 2 and OsPht1; 6, have different functions and kinetic properties in uptake and translocation. Plant J 57:798–809

Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT (2010) Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. Nature 465:627–631

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

Bordes J, Ravel C, Jaubertie J, Duperrier B, Gardet O, Heumez E, Pissavy A, Charmet G, Le Gouis J, Balfourier F (2013) Genomic regions associated with the nitrogen limitation response revealed in a global wheat core collection. Theor Appl Genet 126:805–822

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

Caldwell KS, Russell J, Langridge P, Powell W (2006) Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, Hordeum vulgare. Genetics 172:557–567

Cantu D, Yang B, Ruan R, Li K, Menzo V, Fu D, Chern M, Ronald PC, Dubcovsky J (2013) Comparative analysis of protein-protein interactions in the defense response of rice and wheat. BMC Genom 14:166

Cao W, Jia J, Jin J (2001) Identification and interaction analysis of QTL for phosphorus use efficiency in wheat. Plant Nutr Fertil Sci 7:285–292

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

Dai XQ, Zhang HY, Spiertz J, Yu J, Xie GH, Bouman B (2010) Crop response of aerobic rice and winter wheat to nitrogen, phosphorus and potassium in a double cropping system. Nutr Cycl Agroecosys 86:301–315

Ehrenreich IM, Hanzawa Y, Chou L, Roe JL, Kover PX, Purugganan MD (2009) Candidate gene association mapping of Arabidopsis flowering time. Genetics 183:325–335

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

Fageria N, Baligar V, Li Y (2008) The role of nutrient efficient plants in improving crop yields in the twenty first century. J Plant Nutr 31:1121–1157

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol Ecol Notes 7:574–578

Farooq S (2009) Triticeae: the ultimate source of abiotic stress tolerance improvement in wheat. Salinity and Water Stress. Springer, Netherlands, pp 65–71

Flint-Garcia SA, Thornsberry JMB IV (2003) Structure of linkage disequilibrium in plants. Annu Rev Plant Biol 54:357–374

Flint-Garcia SA, Thuillet AC, Yu J, Pressoir G, Romero SM, Mitchell SE, Doebley J, Kresovich S, Goodman MM, Buckler ES (2005) Maize association population: a high-resolution platform for quantitative trait locus dissection. Plant J 44:1054–1064

Gamuyao R, Chin JH, Pariasca-Tanaka J, Pesaresi P, Catausan S, Dalid C, Slamet-Loedin I, Tecson-Mendoza EM, Wissuwa M, Heuer S (2012) The protein kinase Pstol1 from traditional rice confers tolerance of phosphorus deficiency. Nature 488:535–539

Guo C, Zhao X, Liu X, Zhang L, Gu J, Li X, Lu W, Xiao K (2013) Function of wheat phosphate transporter gene TaPHT2; 1 in Pi translocation and plant growth regulation under replete and limited Pi supply conditions. Planta 237:1163–1178

Guo C, Guo L, Li X, Gu J, Zhao M, Duan W, Ma C, Lu W, Xiao K (2014) TaPT2, a high-affinity phosphate transporter gene in wheat (Triticum aestivum L.), is crucial in plant Pi uptake under phosphorus deprivation. Acta Physiol Plant 36:1373–1384

Gupta PK, Rustgi S, Kulwal PL (2005) Linkage disequilibrium and association studies in higher plants: present status and future prospects. Plant Mol Biol 57:461–485

Hästbacka J, de la Chapelle A, Kaitila I, Sistonen P, Weaver A, Lander E (1992) Linkage disequilibrium mapping in isolated founder populations: diastrophic dysplasia in Finland. Nat Genet 2:204–211

Hinsinger P, Bengough AG, Vetterlein D, Young IM (2009) Rhizosphere: biophysics, biogeochemistry and ecological relevance. Plant Soil 321:117–152

Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6:95–108

Hoagland DR, Arnon DI (1950) The water-culture method for growing plants without soil. Circular California Agricultural Experiment Station 347

Holford I (1997) Soil phosphorus: its measurement, and its uptake by plants. Aust J Soil Res 35:227–240

Holland JB (2007) Genetic architecture of complex traits in plants. Curr Opin Plant Biol 10:156–161

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. Science 265:2037–2048

Li Z, Ni Z, Peng H, Liu Z, Nie X, Xu S, Liu G, Sun Q (2007) Molecular mapping of QTLs for root response to phosphorus deficiency at seedling stage in wheat (Triticum aestivum L.). Prog Nat Sci 10:1177–1184

Li A, Wang X, Leseberg CH, Jia J, Mao L (2008) Biotic and abiotic stress responses through calcium-dependent protein kinase (CDPK) signaling in wheat (Triticum aestivum L.). Plant Signal Behav 3:654–656

Li Y-H, Zhang C, Gao Z-S, Smulders MJM, Ma Z, Liu Z-X, Nan H-Y, Chang R-Z, Qiu L-J (2009) Development of SNP markers and haplotype analysis of the candidate gene for rhg1, which confers resistance to soybean cyst nematode in soybean. Mol Breeding 24:63–76

Liu X, Zhao X, Zhang L, Lu W, Li X, Xiao K (2013) TaPht1; 4, a high-affinity phosphate transporter gene in wheat (Triticum aestivum), plays an important role in plant phosphate acquisition under phosphorus deprivation. Funct Plant Biol 40:329–341

Luo MC, Gu YQ, You FM, Deal KR, Ma Y, Hu Y, Huo N, Wang Y, Wang J, Chen S (2013) A 4-gigabase physical map unlocks the structure and evolution of the complex genome of Aegilops tauschii, the wheat D-genome progenitor. Proc Natl Acad Sci USA 110:7940–7945

Lynch JP (2007) Turner review no. 14. Roots of the second green revolution. Aust J Bot 55:493–512

Mackay I, Powell W (2007) Methods for linkage disequilibrium mapping in crops. Trends Plant Sci 12:57–63

Marschner H (1995) Functions of mineral nutrients: macronutrients. Mineral Nutr Higher Plants 2:379–396

Massman J, Cooper B, Horsley R, Neate S, Dill-Macky R, Chao S, Dong Y, Schwarz P, Muehlbauer G, Smith K (2011) Genome-wide association mapping of Fusarium head blight resistance in contemporary barley breeding germplasm. Mol Breeding 27:439–454

Misson J, Thibaud M-C, Bechtold N, Raghothama K, Nussaume L (2004) Transcriptional regulation and functional properties of Arabidopsis Pht1; 4, a high affinity transporter contributing greatly to phosphate uptake in phosphate deprived plants. Plant Mol Biol 55:727–741

Mujeeb-Kazi A, Rosas V, Roldan S (1996) Conservation of the genetic variation of Triticum tauschii (Coss.) Schmalh. (Aegilops squarrosa auct. non L.) in synthetic hexaploid wheats (T. turgidum L. s. lat. x T. tauschii; 2n = 6x = 42, AABBDD) and its potential utilization for wheat improvement. Genet Resour Crop Ev 43:129–134

Navakode S, Neumann K, Kobiljski B, Lohwasser U, Börner A (2014) Genome wide association mapping to identify aluminium tolerance loci in bread wheat. Euphytica 198:401–411

Neumann K, Kobiljski B, Denčić S, Varshney R, Börner A (2011) Genome-wide association mapping: a case study in bread wheat (Triticum aestivum L.). Mol Breeding 27:37–58

Padulosi S, Hammer K, Heller J (1996) Hulled Wheats: Proceedings of the First International Workshop on Hulled Wheats, Castelvecchio Pascoli, Tuscany, Italy. Bioversity International

Palaisa KA, Morgante M, Williams M, Rafalski A (2003) Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. Plant Cell 15:1795–1806

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

Paszkowski U, Kroken S, Roux C, Briggs SP (2002) Rice phosphate transporters include an evolutionarily divergent gene specifically activated in arbuscular mycorrhizal symbiosis. Proc Natl Acad Sci USA 99:13324–13329

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

Ravel C, Martre P, Romeuf I, Dardevet M, El-Malki R, Bordes J, Duchateau N, Brunel D, Balfourier F, Charmet G (2009) Nucleotide polymorphism in the wheat transcriptional activator Spa influences its pattern of expression and has pleiotropic effects on grain protein composition, dough viscoelasticity, and grain hardness. Plant Physiol 151:2133–2144

Schneider A, Molnár I, Molnár-Láng M (2008) Utilisation of Aegilops (goatgrass) species to widen the genetic diversity of cultivated wheat. Euphytica 163:1–19

Singh A, Reimer S, Pozniak C, Clarke F, Clarke J, Knox R, Singh A (2009) Allelic variation at Psy1-A1 and association with yellow pigment in durum wheat grain. Theor Appl Genet 118:1539–1548

Smith S, Kuehl R, Ray I, Hui R, Soleri D (1998) Evaluation of simple methods for estimating broad-sense heritability in stands of randomly planted genotypes. Crop Sci 38:1125–1129

Stracke S, Haseneyer G, Veyrieras J-B, Geiger HH, Sauer S, Graner A, Piepho H-P (2009) Association mapping reveals gene action and interactions in the determination of flowering time in barley. Theor Appl Genet 118:259–273

Su J, Xiao Y, Li M, Liu Q, Li B, Tong Y, Jia J, Li Z (2006) Mapping QTLs for phosphorus-deficiency tolerance at wheat seedling stage. Plant Soil 281:25–36

Su J-Y, Zheng Q, Li H-W, Li B, Jing R-L, Tong Y-P, Li Z-S (2009) Detection of QTLs for phosphorus use efficiency in relation to agronomic performance of wheat grown under phosphorus sufficient and limited conditions. Plant Sci 176:824–836

Tang Z, Sadka A, Morishige DT, Mullet JE (2001) Homeodomain leucine zipper proteins bind to the phosphate response domain of the soybean VspB tripartite promoter. Plant Physiol 125:797–809

Tenaillon MI, Sawkins MC, Long AD, Gaut RL, Doebley JF, Gaut BS (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (Zea mays ssp. mays L.). Proc Natl Acad Sci USA 98:9161–9166

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES (2001) Dwarf8 polymorphisms associate with variation in flowering time. Nat Genet 28:286–289

Vance CP, Uhde-Stone C, Allan DL (2003) Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource. New Phytol 157:423–447

Wang J, Luo MC, Chen Z, You FM, Wei Y, Zheng Y, Dvorak J (2013) Aegilops tauschii single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. New Phytol 198:925–937

Wissuwa M, Ae N (2001) Genotypic variation for tolerance to phosphorus deficiency in rice and the potential for its exploitation in rice improvement. Plant Breeding 120:43–48

Wissuwa M, Yano M, Ae N (1998) Mapping of QTLs for phosphorus-deficiency tolerance in rice (Oryza sativa L.). Theor Appl Genet 97:777–783

Yang M, Ding G, Shi L, Feng J, Xu F, Meng J (2010) Quantitative trait loci for root morphology in response to low phosphorus stress in Brassica napus. Theor Appl Genet 121:181–193

Yang N, Lu Y, Yang X, Huang J, Zhou Y, Ali F, Wen W, Liu J, Li J, Yan J (2014) Genome wide association studies using a new nonparametric model reveal the genetic architecture of 17

agronomic traits in an enlarged maize association panel. PLoS Genet 10:e1004573

Zhang H, Cao F, Fang S, Wang G, Zhang H, Cao Z (2005) Effects of agricultural production on phosphorus losses from paddy soils: a case study in the Taihu Lake Region of China. Wetl Ecol Manag 13:25–33

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C, Tang C, Toomajian C, Zheng H, Dean C, Marjoram P (2007) An Arabidopsis example of association mapping in structured samples. PLoS Genet 3:e4

Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. Plant Genome 1:5–20