

## Genome-wide Association Study for Anther Length in Some Elite Bread Wheat Germplasm

XIAOPENG SONG, JIE FENG, ZIXIA CUI, CHUANLIANG ZHANG and DAOJIE SUN\*

College of Agronomy, Northwest A&F University, Yangling, Shaanxi, P.R. China

\*Corresponding author: chinawheat@163.com

### Abstract

Song X., Feng J., Cui Z., Zhang C., Sun D. (2018): Genome-wide association study for anther length in some elite bread wheat germplasm. Czech J. Genet. Plant Breed., 54: 109–114.

The anther is a crucial organ for the development of the spike in bread wheat (*Triticum aestivum* L.). Long anthers contain large amounts of pollen grains; thus, they are favourable for cross-pollination and increase resilience against adverse environmental conditions. We conducted a genome-wide association study (GWAS) of anther length in 305 elite wheat lines evaluated during 2013–2015 in two locations and two growing seasons. The mapping panel was genotyped using a high-density Illumina iSelect 90K single nucleotide polymorphism (SNP) array. The GWAS used 18763 SNPs and identified 17 markers associated with anther length in wheat. The loci were mainly distributed across the chromosomes 3A, 3B and 7B. Further studies are required to determine if these are candidate genomic regions of anther length. In addition, anther length had high heritability, and positive correlations between anther length and grain weight per spike were observed.

**Keywords:** bread wheat lines; GWAS; marker-trait association; SNP; spike traits

Bread wheat is a major cereal crop for the increasing global population. However, the rate of increase in wheat yields has declined over the past 30 years in association with the changing temperature and precipitation patterns associated with climate change (OLESEN *et al.* 2011). Although significant progress has been made in enhancing various components of wheat yield (WU *et al.* 2014), fewer advances in yield gain have been achieved. Therefore, it is useful to study new strategies, such as the breeding of hybrid wheat varieties, to promote global food security (WHITFORD *et al.* 2013). It is estimated that the increase in grain yield that can be gained from hybrid wheat varieties varies from about 5 to 20% (LONGIN *et al.* 2012). LONGIN *et al.* (2013) demonstrated the superior traits of hybrid wheat compared with those of the parent plants in terms of grain yield. Wheat lines that display high levels of pollination are crucial for hybrid seed production (LANGER *et al.* 2014). The dissection of genetic resources can be useful to enhance such traits to produce improved hybrid wheat varieties.

Wheat lines with long anthers are better pollinators (NGUYEN *et al.* 2015), which, as mentioned above, is crucial in the breeding of hybrid wheat varieties. In recent decades, several studies have been conducted to improve the pollination traits of wheat, with particular emphasis on the favourable effects of anther length (AL). MILOHNIC and JOST (1970) reported that the anther size was positively correlated with the number of pollen grains ( $r = 0.87$ ). LANGER *et al.* (2014) also observed a positive correlation between AL and pollen mass in wheat ( $r = 0.60$ ). Similarly, NGUYEN *et al.* (2015) reported that AL was associated with pollen grain number in wheat-rye chromosome addition lines ( $r = 0.93$ ). Based on these studies, AL appears to be a key spike trait for improving outcrossing in wheat. Thus, it is useful to study the genetic basis of AL to guide the breeding of superior wheat varieties.

Association mapping, a complementary approach to quantitative trait loci (QTL) mapping, offers a useful method to effectively dissect the genomic regions of complex traits. Compared to traditional

QTL mapping, association studies are cost-effective and time-saving (MYLES *et al.* 2009). As a result of advances in next-generation DNA sequencing, single nucleotide polymorphism (SNP) genotyping can be performed using a high-throughput method, and SNP markers across the genomes have been widely used in association mapping (TALAS *et al.* 2012). Recently, high-density SNP gene-chips, such as the 9K and 90K SNP chips, have provided a superior approach to genome-wide association studies (GWASs) in wheat (CORMIER *et al.* 2014; LOPES *et al.* 2015). Moreover, several studies also demonstrated that GWAS was useful for dissecting various complex traits, such as nitrogen use efficiency and yield trait (CORMIER *et al.* 2014; AIN *et al.* 2015; LOPES *et al.* 2015). Thus, the GWAS approach is appropriate to use in studies of AL in wheat. To the best of our knowledge, no GWASs of AL have been conducted.

In the present study, a 90K Illumina iSelect array was used to identify polymorphic SNPs in the genomes of 305 bread wheat genotypes. The objectives of the study were to (1) analyse the correlations between AL and the major spike traits and (2) to identify the SNP markers associated with AL using the GWAS.

## MATERIAL AND METHODS

**Plant materials.** A total of 305 genetically diverse varieties were used in the GWAS (Table S1 in Electronic Supplementary Material (ESM)). These materials included the core parents and elite cultivars, and introduced varieties used by wheat breeding programs that have had significantly increased wheat yields in China. The field trials were performed at Yangling and Anyang during two growing seasons (2013–2015). All materials were planted in randomized complete blocks with three replications. Each plot contained three 2-m long rows, and the rows were separated by 30 cm. In addition, appropriate field management, including weed, disease and pest control, was performed to avoid adverse influences on the yield. The two pairs of florets were sampled from the central portion of six random spikes per plot and the ALs were measured before flowering using Vernier calipers. Spike length (SL), number of sterile spikelets per spike (St), number of spikelets per spike (SPS), grain number per spikelet (GNS), and grain number per spike (GNPS) were measured, and the means of five randomly selected spikes were calculated per plot at the late grain-filling stage. Spikelet density per spike (Sds) was calculated from the SPS to SL ratio.

At harvest, the grain weight per spike (GWS) was calculated from the average value of the grain weight in five randomly selected spikes per plot, and the thousand kernel weight (TKW) was measured by weighing duplicate samples of 500 kernels per plot.

**Genotyping for SNP markers.** The 305 bread wheat varieties were genotyped using the 90K SNP assay according to the manufacturer's protocol. Genotypes were called using the GenomeStudio software program (Illumina). Heterozygous loci were treated as missing data, and loci with more than 15% missing values were discarded from the dataset. Markers with a minor allele frequency less than 0.05 across all samples were excluded in subsequent analysis. A consensus genetic map of the SNP markers was used to assign the chromosomal position. In total, 18763 informative SNPs were retained and used in the GWAS.

**Statistical analysis.** Best linear unbiased estimates (BLUEs) for all measured traits were estimated, assuming fixed genotypic effects. The GenStat software (Ver. 16, 2013, VSN International) was used to calculate the value of BLUEs applying the Mixed models REML module and the Linear mixed models function. Analysis of variance (ANOVA) was performed using the PROC MIXED procedure with SAS software according to the methods of ZHAI *et al.* (2016). The broad-sense heritabilities ( $h^2$ ) of each trait were calculated across all environments as

$$h^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_{ge}^2/e + \sigma_e^2/re)$$

where:

$\sigma_g^2$  – genotypic variance

$\sigma_{ge}^2$  – genotype by environment interaction variance

$\sigma_e^2$  – estimates of residual error variances

$e, r$  – number of environments and replicates per environment, respectively

Pearson's correlation coefficients were calculated with the PROC CORR procedure in the SAS 9.2 software.

**Genome-wide association analysis (GWAS).** A GWAS was conducted in the GAPIT implemented in R (LIPKA *et al.* 2012). The population structure represented by the Q matrix was assessed in the structure software with 590 unlinked markers (PRITCHARD *et al.* 2000). A  $P$ -value threshold of  $10^{-4}$  was used to define a marker as statistically significant.

## RESULTS

**Phenotypic evaluation and analysis of correlations among traits.** Based on the BLUE values across

<https://doi.org/10.17221/70/2017-CJGPB>

Table 1. Statistical analysis of phenotypic data and broad-sense heritabilities ( $h^2$ ) for anther length (AL) and other spike traits

Traits	Descriptive statistics			Variance parameters			
	min	max	mean	$\sigma_g^2$	$\sigma_{ge}^2$	$\sigma_e^2$	$h^2$
AL (mm)	2.7	6.0	3.8	2.36	1.58	1.84	0.81
DH (days)	187.1	207.6	195.8	6.24	2.16	2.73	0.89
SL (cm)	6.1	13.4	9.2	4.25	1.33	1.58	0.90
TKW (g)	18.2	48.3	38.0	20.83	4.06	10.76	0.92
St	0.9	4.3	2.5	5.21	12.36	13.21	0.55
SPS	13.5	24.2	18.1	4.17	1.24	1.13	0.91
GNS	2.2	4.2	3.3	1.32	0.86	1.42	0.80
GNPS	20.2	57.1	34.4	1.83	2.67	6.25	0.61
Sds (spikelets/cm)	1.5	2.6	2.0	0.42	0.26	0.67	0.78
GWS (g)	0.7	2.6	1.8	2.53	1.42	1.61	0.84

DH – days to heading; SL – spike length; TKW – thousand kernel weight; St – number of sterile spikelets per spike; SPS – number of spikelets per spike; GNS – grain number per spikelet; GNPS – grain number per spike; Sds – spikelet density per spike; GWS – grain weight per spike

four environments, the minimum, maximum, mean variance parameters, and broad-sense heritability of each trait were estimated (Table 1). The TKW and SPS ranged from 18.2 to 48.3 g, and 13.5 to 24.2, with averages of 38.0 g and 18.1 g, respectively, and had the highest broad-sense heritability. The AL ranged from 2.7 to 6.0 mm, with an average of 3.8 mm (Figure S1 in ESM), and showed high heritability (0.81), indicating that AL is a stable trait and mainly controlled by genetic effects. St showed low heritability (0.55), and a low-to-medium heritability was observed in GNPS.

**Analysis of correlations among traits.** Pearson's correlation coefficients were calculated based on the averages across all trials (Table 2). The maximum positive correlation value ( $r = 0.64$ ) was observed

between AL and GWS, followed by that between AL and GNPS ( $r = 0.54$ ). Similarly, GNS and TKW showed a moderate correlation with AL ( $r = 0.51$  and  $r = 0.46$ , respectively). The correlation coefficient between AL and SL was low (0.32). The Sds showed a negative correlation ( $r = -0.42$ ) with AL. However, St showed no significant correlation with AL. GWS was also positively correlated with GNPS and SPS ( $r = 0.80$  and  $r = 0.76$ , respectively).

**Population structure and marker-trait association (MTA) for the anther length.** The analysis of population structure indicated that the association mapping panel had five major subpopulations ( $K = 5$ ) (Figure S2 in ESM) based on the approach of EVANNO *et al.* (2005). Overall, 17 markers associated with AL were detected at  $P < 0.0001$  using GAPIT; MTAs were

Table 2. Pearson's correlation coefficient for anther length (AL) and various spike traits in bread wheat germplasm

Variables	AL	SL	TKW	St	SPS	GNS	GNPS	Sds
SL	0.32**							
TKW	0.46**	-0.08						
St	-0.12	-0.15	0.14					
SPS	0.18	-0.34**	-0.35**	-0.15				
GNS	0.51**	0.14	0.06	-0.13	0.70*			
GNPS	0.54**	0.18	0.14	-0.28	0.85*	0.66**		
Sds	-0.42*	-0.70**	-0.24	0.18	0.12**	0.13	0.16*	
GWS	0.64**	0.14	0.55**	0.12	0.76**	0.55**	0.80*	0.07

\*, \*\*statistically significant at  $P < 0.05$ , 0.01, respectively; AL – anther length; SL – spike length; TKW – thousand kernel weight; St – number of sterile spikelets per spike; SPS – number of spikelets per spike; GNS – grain number per spikelet; GNPS – grain number per spike; Sds – spikelet density per spike; GWS – grain weight per spike

Table 3. Description of main SNP markers associated with the anther length

Trait	SNP	Chr	Position (cM)	Alleles	$-\log_{10}$ ( <i>P</i> -value)	Maf	Marker $R^2$	All eff
AL	wsnp_BE637864B_Ta_1_1	1B	62	280(A)/20(C)	4.18	0.07	0.19	0.17
	wsnp_Ex_c53364_56625806	3A	87	253(G)/48(A)	5.84	0.16	0.21	0.13
	Excalibur_c13242_1178	3A	88	264(G)/28(A)	4.64	0.09	0.20	0.16
	RAC875_c25375_236	3B	62	222(C)/70(A)	5.18	0.23	0.21	–0.13
	Kukri_s116671_78	3B	63	227(A)/76(G)	5.45	0.25	0.21	–0.12
	BobWhite_c19155_246	5A	67	205(G)/87(A)	5.22	0.29	0.29	0.12
	wsnp_JD_c13673_13606066	7B	136	224(G)/77(A)	4.39	0.25	0.21	–0.12

Anther length (AL), SNP markers, chromosomes (Chr), position, alleles,  $-\log_{10}$  (*P*-value), minor allele frequency (Maf), marker  $R^2$ , allelic effects (All eff) are shown in the table for each significant association

identified in seven major genomic regions within different chromosomes. Table 3 shows the 7 major markers that were associated with AL. Among the significant MTAs, chromosomes 3A and 3B had the highest number of MTAs, followed by 5A. Most of the MTAs were distributed on chromosomes 3A, 3B and 5A, and they accounted for 19–29% of the variation in AL. This indicates that the genetic regions of 3A, 3B and 5A chromosomes could provide the MTAs for AL.

## DISCUSSION

**The evaluation of anther length.** SAITO *et al.* (2001) reported that AL was associated with the QTLs of cold tolerance in rice. The anther of wheat could be sensitive to adverse fluctuations in climate, such as late spring coldness that can affect its development. However, most bread wheat germplasm resources possess relatively small anthers and fewer pollen grains compared to rye anthers, and this limits their potential as elite pollinators for hybrid wheat seed production. Several studies on AL in wheat have revealed a positive correlation with the pollen grain number (DE VRIES 1974; LANGER *et al.* 2014). In a recent study of wheat-rye chromosome addition lines, AL was positively correlated with the number of pollen grains per anther (NGUYEN *et al.* 2015). These studies indicate that increasing AL could produce more pollen grains, which would have favourable effects on pollination traits in wheat. Furthermore, the heritability of AL was high in this study. The AL of Chuanmai 41 was 6.0 mm, which renders this variety a useful resource for improving the pollination traits of bread wheat. However, few studies have been carried out on the correlations between

AL and spike traits in wheat; hence, further studies on this topic would be useful. We observed several correlations between AL and other spike traits. AL showed a significant positive relationship with GWS; likewise, GNPS, GNS and TKW showed moderately significant correlations with AL. In contrast, a negative correlation was observed between AL and Sds. Our results show that AL could be a suitable trait for improving some important agronomic trait such as GWS, GNPS, GNS and TKW in wheat.

The AL has been documented as influencing the number of extruded anthers (LONGIN *et al.* 2012) and has a strong effect on the quantity of shed pollen (LANGER *et al.* 2014). Several studies indicated that cultivars with high anther extrusion were quite resistant to the disease Fusarium head blight (FHB) (BUERSTMAYR & BUERSTMAYR 2015). Furthermore, some fully cleistogamous wheat varieties with non-opening self-pollinating flowers could also contribute to increased FHB resistance, so it is also very interesting to observe correlations between AL and FHB in a subsequent study. In a recent study, BOEVEN *et al.* (2016) reported that the dwarfing allele had a negative effect on anther extrusion, thus, this suggests that the correlation between AL and plant height is promising to investigate.

**Marker-trait associations and population structure.** Population structure should be taken into account in the GWAS since it could lead to non-functional false associations. Therefore, we treated the population structure as a covariate in the GAPIT to reduce the influence of population stratification.

AL is a key trait for increasing the number of pollen grains. Although significant progress in enhancing the rate of outcrossing in wheat has been made in recent decades (SIMONS *et al.* 2006; NGUYEN *et al.* 2015),



<https://doi.org/10.17221/70/2017-CJGPB>

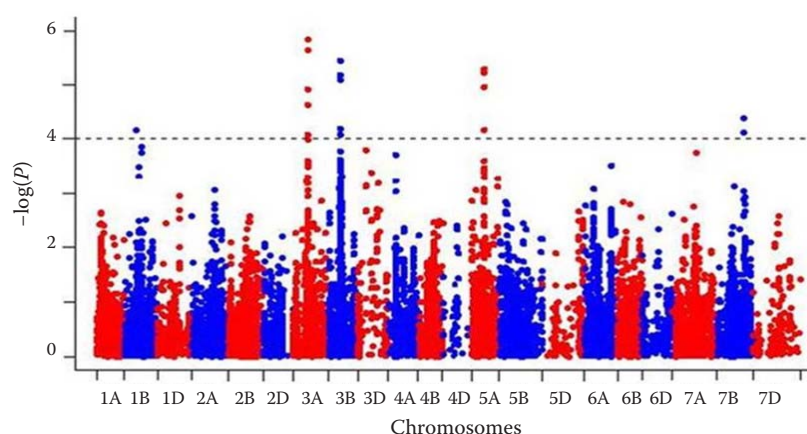


Figure 1. The genome-wide association study (GWAS) for anther length of wheat based on normalized averages across four environments

Manhattan plot of  $P$  values depicting genomic regions associated with anther length using 18 763 SNP markers; the horizontal line indicates a significant threshold for the anther length

the role of AL appears to have been ignored in most studies. To date, few genes have been identified as effective in increasing wheat AL. Dissection of the genomic regions affecting the AL in a GWAS offers a promising approach. In this study, 17 MTAs were observed on chromosomes 1B, 3A, 3B, 5A, and 7B, and no MTAs were present on genome D (Figure 1). More specifically, these genomic regions were distributed mainly on chromosomes 3A, 3B and 5A. These genomic regions merit further research, but due to a lack of data from previous studies, it is currently difficult to align and compare the QTLs identified by our study. However, we conducted a comparative analysis with the rice genome, and these SNPs associated with AL were collinear with rice chromosome 1 (11.95–28.77 Mb), 1 (3.92–4.07 Mb), 6 (7.92–26.19 Mb), and 9 (4.07–20.86 Mb), respectively (Table S2 in ESM). Based on these potential genomic regions, it would be useful to construct near-isogenic lines and mutagenesis to further dissect these loci. In addition, it was observed that the wheat-rye chromosome addition lines significantly increased AL and pollen grain number (NGUYEN *et al.* 2015). Further studies are required to characterize the novel genes related to AL in wheat, which would increase our understanding of the genetic mechanisms and provide elite pollinators for the production of hybrid wheat seeds.

**Acknowledgements.** This work was supported by the National 973 Program on Key Basic Research Project (2014CB138100) and Natural Science Foundation of Shanxi

Province of China (2015JM3094). The authors are grateful to Prof. Z. HE and X. XIA, for kindly providing SNP markers and plant materials.

## References

- Ain Q., Rasheed A., Anwar A., Mahmood T., Imtiaz M., Mahmood T., Xia X.C., He Z.H., Quraishi U.M. (2015): Genome-wide association for grain yield under rainfed conditions in historical wheat cultivars from Pakistan. *Frontiers in Plant Science*, 6: 743.
- Boeven P.H.G., Longin C.F.H., Leiser W.L., Kollers S., Ebmeyer E., Würschum T. (2016): Genetic architecture of male floral traits required for hybrid wheat breeding. *Theoretical and Applied Genetics*, 129: 2343–2357.
- Buerstmayr M., Buerstmayr H. (2015): Comparative mapping of quantitative trait loci for Fusarium head blight resistance and anther retention in the winter wheat population Capo × Arina. *Theoretical and Applied Genetics*, 128: 1519–1530.
- Cormier F., Gouis F.J., Dubreuil P., Lafarge S., Praud S. (2014): A genome-wide identification of chromosomal regions determining nitrogen use efficiency components in wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, 127: 2679–2693.
- De Vries A.P. (1974): Some aspects of cross-pollination in wheat (*Triticum aestivum* L.). 3. Anther length and number of pollen grains per anther. *Euphytica*, 23: 11–19.
- Evanno G., Regnaut S., Goudet J. (2005): Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, 14: 2611–2620.
- Langer S.M., Longin C.F.H., Würschum T. (2014): Phenotypic evaluation of floral and flowering traits with relevance

<https://doi.org/10.17221/70/2017-CJGPB>

- for hybrid breeding in wheat (*Triticum aestivum* L.). *Plant Breeding*, 133: 433–441.
- Lipka A.E., Tian F., Wang Q., Peiffer J., Li M., Bradbury P.J., Gore M.A., Buckler E.S., Zhang A. (2012): GAPIT: genome association and prediction integrated tool. *Bioinformatics*, 28: 2397–2399.
- Longin C.F.H., Mühleisen J., Maurer H.P., Zhang H., Gowda M., Reif J.C. (2012): Hybrid breeding in autogamous cereals. *Theoretical and Applied Genetics*, 125: 1087–1096.
- Longin C.F.H., Gowda W., Mühleisen J., Ebmeyer E., Kazman E., Schachschneider R., Schacht J., Kirchhoff M., Zhao Y., Reif J.C. (2013): Hybrid wheat: quantitative genetic parameters and consequences for the design of breeding programs. *Theoretical and Applied Genetics*, 126: 2791–2801.
- Lopes M.S., Dreisigacker S., Peña R.J., Sukumaran S., Reynolds M.P. (2015): Genetic characterization of the wheat association mapping initiative (WAMI) panel for dissection of complex traits in spring wheat. *Theoretical and Applied Genetics*, 128: 453–464.
- Milohnic J., Jost M. (1970): Pollen production and anther extrusion of wheat (*Triticum aestivum* L. Em Thell.). *Acta Agronomica Hungarica*, 19: 17–23.
- Myles S., Peiffer J., Brown P.J., Ersoz E.S., Zhang Z., Costich D.E., Buckler E.S. (2009): Association mapping: critical considerations shift from genotyping to experimental design. *The Plant Cell*, 21: 2194–2202.
- Nguyen V., Fleury D., Timmins A., Laga H., Hayden M., Mather D., Okada T. (2015): Addition of rye chromosome 4R to wheat increases anther length and pollen grain number. *Theoretical and Applied Genetics*, 128: 953–964.
- Olesen J.E., Trnka M., Kersebaum K.C., Skjelvåg A.O., Seguin B., Peltonen-Sainio P., Rossi F., Kozyra J., Micale F. (2011): Impacts and adaptation of European crop production systems to climate change. *European Journal of Agronomy*, 34: 96–112.
- Pritchard J.K., Stephens M., Donnelly P. (2000): Inference of population structure using multilocus genotype data. *Genetics*, 155: 945–959.
- Saito K., Miura K., Nagano K., Hayano-Saito Y., Araki H., Kato A. (2001): Identification of two closely linked quantitative trait loci for cold tolerance on chromosome 4 of rice and their association with anther length. *Theoretical and Applied Genetics*, 103: 862–868.
- Simons K.J., Fellers J.P., Trick H.N., Zhang Z., Tai Y.S., Gill B.S., Faris J.D. (2006): Molecular characterization of the major wheat domestication gene *Q*. *Genetics*, 172: 547–555.
- Talas F., Wurschum T., Reif J.C., Parzies H.K., Miedaner T. (2012): Association of single nucleotide polymorphic sites in candidate genes with aggressiveness and deoxynivalenol production in *Fusarium graminearum* causing wheat head blight. *BMC Genetics*, 13: 14.
- Whitford R., Fleury D., Reif J.C., Garcia M., Okada T., Korzun V., Langridge P. (2013): Hybrid breeding in wheat: technologies to improve hybrid wheat seed production. *Journal of Experimental Botany*, 64: 5411–5428.
- Wu W., Li C., Ma B., Shah F., Liu Y., Liao Y. (2014): Genetic progress in wheat yield and associated traits in China since 1945 and future prospects. *Euphytica*, 196: 155–168.
- Zhai S.N., He Z.H., Wen W.E., Jin H., Liu J.D., Zhang Y., Liu Z.Y., Xia X.C. (2016): Genome-wide linkage mapping of flour color-related traits and polyphenol oxidase activity in common wheat. *Theoretical and Applied Genetics*, 129: 377–394.

Received for publication May 7, 2017

Accepted after corrections January 10, 2018

Published online March 9, 2018