Bulked QTL-Seq identified a major QTL for the awnless trait in spring wheat cultivars in Qinghai, China

Dongxia Wang¹, Dong Cao¹,†, Yuan Zong¹,†, Yun Li¹,†, Jinmin Wang¹, Zongren Liu¹ and Baolong Liu¹

¹State Key Laboratory of Plateau Ecology and Agriculture, Qinghai University, Xining, Qinghai, PR China; †Department of Agriculture and Forestry, College of Agriculture and Animal Husbandry, Qinghai University, Xining, Qinghai, PR China; Qinghai Province Key Laboratory of Crop Molecular Breeding, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining, Qinghai, PR China; Laboratory of Wheat Quality Improvement, the Innovative Academy of Seed Design, Chinese Academy of Sciences, Xining, Qinghai, PR China

ABSTRACT
The awnless trait is the favorite trait for wheat breeding in Qinghai, China, but the major gene underlying the trait is unknown. This study aimed to analyze a recombinant inbred line (RIL) population containing 112 lines by crossing common wheat varieties GY448 (awnless) and GY115 (awned) using genotyping-by-sequencing analysis. A total of 56.62 Gb of clean sequence data were generated. A major gene was identified for the awnless trait based on 101,275 single-nucleotide polymorphisms by quantitative trait locus (QTL)-seq analysis. The region from 685.50 to 709.77 Mb on chromosome 5AL contained 334 genes and 394 transcripts, including B1, the major gene controlling the awnless trait. A 25-bp indel located in the B1 promoter region has a diagnostic value for the awnless/awned genotypes in the RIL population. A total of 42 spring wheat cultivars in Qinghai, China, were analyzed. Further, 23 awnless cultivars had the alleles of the GY448 parent, while 19 awned cultivars had the alleles of the GY115 parent. B1 is crucial in the awnless trait in wheat cultivars in Qinghai, China. The findings of this study might guide the breeding of new wheat cultivars in this region.

ARTICLE HISTORY
Received 11 September 2020
Accepted 25 November 2020

KEYWORDS
Awnless; B1; Qinghai; QTL-seq; wheat

Introduction
The awn is a fibrous extension of the lemma in the florets of some grass species, including wheat, barley and rice. It is important in the evolution and production of grasses belonging to the family Poaceae. The study of awn length has been a research focus for many years. The awn has photosynthetic activity and is a key photosynthetic organ, especially under adverse conditions [1–3]. During the grain-filling stage, the awn has a high photosynthetic rate, and its photosynthetic products are an essential source of wheat grain assimilates [4,5]. The awn contributes to up to 42% of the grain yield in wheat [1]. Genes associated with photosynthesis and chlorophyll and carotenoid biosynthesis are preferentially expressed [6]. In addition, the awn is a prominent feature of Triticum species and an important morphological marker for distinguishing different wheat varieties. It also protects against animal and seed dispersal [7,8]. Before the mechanization of agriculture, farmers preferred the awnless wheat cultivars because a long awn could puncture the skin. Moreover, the awn can also stab the oral cavity of sheep and cattle when feeding. In Qinghai, China, the wheat is still harvested by hand, and the straw is fed to the Tibetan sheep and cows. Thus, the major gene identification for the awnless trait is beneficial for wheat breeding in Qinghai, China.

In common wheat (Triticum aestivum L.), awn development is mainly regulated by five loci, including Tipped1 (B1), Tipped2 (B2), Tipped3 (B3), Hooded (Hd) and A [9–11]. B1, B2, B3 and Hd are the main dominant inhibitors of awn development. The effect of B1 is the strongest, while A promotes awn formation. Most studies have focused on B1, B2 and Hd. B1 is located on chromosome arms 5AL [3,5,12,13]. TraesCS5A02G542800, distinguishing different wheat varieties. It also protects against animal and seed dispersal [7,8].
a candidate gene of the B1 locus, encodes C2H2 zinc finger protein and has been isolated in wheat [14–16]. The promoter variations of B1 are closely associated with awn development in awnless and awned individuals [14]. B2 and Hd are, respectively, located on chromosome arms 6BL [5,12,17] and 4AS [5,10]. B2 can cause the entire panicle to have a short awn, and Hd is associated with a hooked phenotype. However, markers tightly linked to B2 were not found [12], which might be related to the presence of B2 in the centromere region [18]. The locus inhibiting awn development is also located on chromosome 6B in durum wheat [19]. A combination of the three recessive hd, b1 and b2 alleles is fully awned, while a combination of two dominant loci, such as Hd and B1, or Hd and B2, or B1 and B2, is sufficient to induce complete awn inhibition [10–12]. Moreover, T locus, controlling the awnless phenotype, is located on the chromosome 5AL in Triticum carthlicum Nevski and is different from B1 [20,21]. In the Aegilops tauschii variety anthera, the major awnless locus Antr was located on the distal region of chromosome 5DS and inhibited awn elongation in wheat and its relatives [22].

Bulked segregant analysis (BSA) is an efficient method used to rapidly identify markers closely linked to a gene of interest [23]. Some efficient mapping approaches, including quantitative train locus (QTL)-seq, MutMap and specific locus amplified fragment sequencing (SLAF-seq), were developed by combining BSA and next-generation sequencing of a DNA pool, and the candidate genomic region was rapidly identified in rice [24,25], peanut [26], maize [27], and capsicum [28]. In this study, a recombinant inbred line (RIL) was built by a cross between common wheat varieties GY 448 and GY 115 (awned). QTL-seq was used to identify the candidate genomic region tightly linked to the awning inhibitor loci. Furthermore, the diagnostic markers were designed to understand the role of QTL-seq in the awnless trait formation in the cultivars in Qinghai, China.

Materials and methods

Plant materials
An awnless/awned wheat cross was made between a pair of varieties with contrasting awn length: GY448 and GY115 (Figure 1). The parental GY448 is an awnless variety, while GY115 shows an awned phenotype. GY448 was the favorite cultivar in Qinghai, China. It had been planted for 20 years and occupied 50% of the plant area. The F2 progenies derived from the GY448/GY115 cross were further developed into RILs using the single-seed descent approach. A total of 112 F9-generation plants were planted in a randomized complete block design with three replicates in 2018 and 2019 in Xining (101°74'E, 36°56'N), Qinghai, China. Each replicate row was 2 m long, with a 20-cm row-to-row distance. The parental lines and RIL lines were scored qualitatively for an awn phenotype as either awnless or awned by visual inspection [29]. Ultimately, 21 extreme individual plants with a fully awnless phenotype and 21 extreme plants with a fully awned phenotype belonging to RILs were selected for the mapping and validation of the target QTL controlling awn length.

Genotyping-by-sequencing analysis
A total of 21 individual plants with the awnless trait belonging to RILs were sampled as an awnless bulk, and 21 individuals with the awned trait belonging to RILs were sampled as an awned bulk. Genomic DNAs of the female parent GY448, male parent GY115 and RIL individuals were extracted from leaves using the methods proposed by Yan et al. [30]. Bulked samples were obtained by mixing equimolar DNA. Genomic DNA from the parents and two bulks were used to construct the libraries for genotyping-by-sequencing (GBS) analysis. GBS libraries were prepared using protocols described in Poland et al. [31]. Libraries were sequenced using Illumina paired-end sequencing.
technology with read lengths of 150 bp on an Illumina HiSeq X instrument by Novogene Bioinformatics Institute (Beijing, China).

**Identification of single-nucleotide polymorphisms**

Raw reads obtained from the four samples were subjected to a set of quality control procedures using FastQC v0.11.5 and Trimmomatic-0.38. After excluding the reads with low-quality bases (quality score < 20) and length less than 50 bp, clean reads were mapped against the wheat reference genome (https://urgi.versailles.inra.fr/download/iwgsc/IWGSC_RefSeq_Assemblies/v1.0/) using the Burrows-Wheeler Aligner (BWA) software with the BWA-MEM algorithm [32], and applied for single-nucleotide polymorphisms (SNP) calling using the GATK software. SNPs with read depth less than five, RMS mapping quality less than 20, and QUAL less than 50 were filtered out. Only the SNPs homozygous in either parent and polymorphic between the parents were prepared for the further QTL-seq analysis.

**QTL-seq analysis**

A statistical approach through calculating the SNP index was used to identify the closely linked SNPs for the target trait [24]. The SNP index and Δ (SNP index) for each SNP position were calculated, and corresponding SNP-index graphs were plotted for the awnless bulk and the awned bulk to identify the candidate regions responsible for the differences in awn length. The SNP index and Δ (SNP index) were calculated for each position as previously described [24]. The absolute Δ (SNP-index) value for the closely linked SNPs should be significantly higher than the threshold at the 99% level of significance. BLAST software was employed to annotate the genes in the candidate genomic region.

**Promoter isolation of B1**

The two primers (F: 5’- CATCCGTATGTAGTTTGTATTG-3’; R: 5’- TTCTTGAAGCTG CGTGAAG-3’) were used to isolate the promoter regions of B1 in GY 448 and GY 115. Based on the difference in the promoter region, two primers (ALI-1pF: 5’- GGAAATCACGAATTCCAC-3’; ALI-1pR: 5’- GCAGCATTATTACTGTCAAC-3’) were designed and used to identify variations in the RIL population and 42 wheat varieties in Qinghai, China. Polymerase chain reaction amplifications were conducted, and the products were confirmed by electrophoresis in 1.0% or 2.5% agarose gel. The bands of the expected size were purified from the gel using an EasyPure Quick Gel Extraction Kit (TransGen, Beijing, China). The targeted DNA fragments were ligated into a PGEMT Easy Vector (Promega, WI, USA), which were transformed into Escherichia coli DH 5a competent cells. The positive clones were picked out and sent to a commercial company (Sangon Biotech Co., Ltd., Shanghai, China) for sequencing.

**Results and discussion**

**Identification of a major QTL for awn length by QTL-seq analysis**

Obvious phenotypic differences between the parents GY448 (awnless) and GY115 (awned) were observed (Figure 1). The RIL population (GY448/GY115) used in this study had high phenotypic variability for awn length. Of the 112 RIL lines evaluated in the initial screening of the population, 21 showed a fully awnless phenotype and were similar to GY448, while 21 showed a fully awned phenotype and were similar to GY115. Awnless and awned lines were observed in two additional summer field trials, and the trait was steady. Therefore, awnless bulks were constituted by mixing equimolar DNA from 21 RILs with an awnless phenotype. Similarly, awned bulks were constituted by mixing equimolar DNA from 21 RILs with an awned phenotype for QTL-seq analysis. The GBS data were generated for four samples: awnless parent GY448, awned parent GY115, awnless bulk and awned bulk. After filtering, the clean data were generated for GY448 (1.16 Gb), GY115 (1.22 Gb), awnless bulk (31.53 Gb) and awned bulk (22.71 Gb) (Supplemental Table S1). A total of 8.41, 8.83, 227.52 and 163.71 million clean reads were obtained for GY448, GY115, awnless bulk and awned bulk, with the Q30 percentage of 89.52%, 89.59%, 91.04% and 91.45%, respectively (Supplemental Table S1).

When aligning the clean reads to a common wheat reference genome, the SNP was identified for GY448, GY115, awnless bulk and awned bulk. A total of 101,275 high-quality SNPs homozygous for each parent and showing polymorphism between the parents were developed for further QTL-seq analysis. A distribution diagram of the SNPs used in association analysis was drawn by plotting the number of SNPs using a sliding window of 5 Mb against the genome position (Supplemental Figure S1). To identify the QTL for awn length, the Δ (SNP index) was calculated for the selected 101,275 SNPs (Figure 2, Supplemental Figure S1). Δ (SNP-index) analysis identified a significant genomic region for Qal.nwipb-5AL (the B1 locus),
which was located between 685.50 and 709.77 Mb on chromosome 5AL (Figure 2). The candidate region with a 24.27-Mb interval contained 221 SNPs (Supplemental Table S2). The $D$ (SNP index) for the SNPs located in the candidate genomic region was greater than the threshold and close to 0.70 with a confidence level of 99% (Figure 2).

Awn length, as a quantitative trait, is controlled by multiple loci [9–11]. The mapping and isolation of QTLs for awn length is important for a better understanding of molecular mechanisms underlying the awn formation and elongation. Previous studies preliminarily identified the loci of awn suppression based on many mapping populations [3,5,9,12,17], and two loci, $B1$ and $Hd$, were detected by the aforementioned studies. In the present study, QTL-seq, the rapid QTL identification method, was used to identify QTLs for awn length. One major QTL for awn length, $Qal.nwipb-5AL$, was detected in a candidate region (about 24.27 Mb). The QTL was located in the terminal region of 5AL and was in the same region as the $B1$ locus. Previous studies revealed that $B1$ was 8.0 cM [5], 5.4 cM [3] distal to $Xgwm291$. The $B1$ locus was located in the telomeric region of chromosome 5AL (306 cM) reported by Yoshioka et al. [12]. The interval of the $B1$ locus was further shortened in this study, thus laying the foundation for the map-based cloning of $B1$.

**B1 was responsible for the awnless trait of GY448**

A total of 334 genes and 394 transcripts, including TraesCS5A02G542800, were annotated by the BLAST software (Supplemental Table S3). TraesCS5A02G542800 was confirmed to be a candidate transcription repressor encoding C2H2 zinc finger protein underlying awn suppression at the $B1$ locus in wheat [14–16]. ALI-1, the underlying gene of the $B1$ locus, caused cell proliferation stagnation and cell number reduction during awn development [14]. The promoter region differences of ALI-1 were associated with awn development [14]. The genomic sequence of TraesCS5A02G542800 was isolated from GY448 and GY115, and no genetic variations existed in the coding region. The promoter regions of GY448 and GY115 were compared with the 89 wheat accessions in a previous study [14]. Four SNPs and one indel existed in the promoter regions of GY448 and GY115 (Supplemental Table S4). The promoter
variations of GY448 were similar to those in wheat accession Cadenza of group 1 [14], while the variations of NY115 were similar to those in XiShanBianSui of group 3 [14] (Supplemental Table S4). A molecular marker, ALI-1p, was developed based on a 25-bp deletion of the promoter region. All individuals with awnless and awned phenotypes were diagnosed using the marker ALI-1p (Supplemental Figure S3). These results suggested that B1 was the functional gene for the awnless phenotype in Qinghai, China.

Allele diversity in the wheat cultivars in Qinghai, China

GY448 is planted widely in Qinghai, China, and is also one of the main parents for new spring wheat cultivar breeding in Qinghai, China. In this study, 42 spring wheat cultivars were collected, while 23 and 19 wheat cultivars showed an awnless phenotype and an awned phenotype, respectively (Supplemental Table S5). All 23 wheat cultivars with an awnless phenotype contained a 25-bp deletion in the promoter region, which was the same as that in GY448 (Figure 3).

In Western China, several special landraces exist because of the reclusive environment in high mountains; e.g. Yunnan hulled wheat (Triticum aestivum ssp. Yunnanense, AABBDD), Tibetan wheat (T. aestivum ssp. tibetanum Shao) and Xinjiang wheat (T. petropavlovskyi Udats et Migusch) [33]. The wheat landrace contains some particular traits. Yunnan hulled wheat contains the hulled trait, which could be obtained in other regions. Moreover, Yunnan hulled wheat also contains some rare variations in the grain hardness gene. Qinghai province is linked to Tibet and Xinjiang geographically. The geographic environment is similar to that in Tibetan and Xinjiang provinces. The awnless trait is the favorite trait in this region because the awns can puncture the skin and stab the oral cavity of sheep and cattle. The artificial selection of the awnless trait might induce the mutation of major genes relative to the awnless trait. Unfortunately, the major gene controlling the awnless trait of GY448 resided in the same region as B1. B1 was identified to encode C2H2 zinc finger protein, causing cell proliferation stagnation and cell number reduction and suppressing awn development [14–16]. B1 was responsible to the awnless trait in GY448. In addition, 23 wheat cultivars with the awnless trait in Qinghai, China, also contained a 25-bp deletion, suggesting that B1 was vital in the awnless trait in Qinghai, China. The findings of this study can guide the breeding of new cultivars with the awnless trait in Qinghai, China.

Conclusions

Bulked QTL-Seq analysis was carried out to identify the candidate gene of the awnless trait in an RIL population (GY448/GY115 cross). The results demonstrated the candidate region was located on chromosome 5AL, and contained 334 genes and 394 transcripts. B1, the major gene controlling the awnless trait, was responsible to the awnless trait in GY448. A 25-bp indel of the B1 promoter region has a diagnostic value for the awnless/awned genotypes in the RIL population and 42 spring wheat cultivars. B1 is crucial in the awnless trait in wheat cultivars in Qinghai, China. The findings of this study might guide the breeding of new wheat cultivars in this region.

Funding

This study was financially supported by the Project of Qinghai Science and Technology Department (2019-ZJ-982Q, 2020-SF-132), West Light Talent Program of the Chinese Academy of Sciences, the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA24030102-4).
Disclosure statement
The authors have declared that no competing interests exist.

References
[1] Maydup ML, Antonietta M, Guiamet JJ, et al. The contribution of ear photosynthesis to grain filling in bread wheat (Triticum aestivum L.). Field Crop Res. 2010;119(1):48–58.
[2] Rebetzke GJ, Bonnett DG, Reynolds MP. Awns reduce grain number to increase grain size and harvestable yield in irrigated and rainfed spring wheat. J Exp Bot. 2016;67(9):2573–2586.
[3] Wu Q, Chen Y, Fu L, et al. QTL mapping of flag leaf traits in common wheat using an integrated high-density SSR and SNP genetic linkage map. Euphytica. 2016;208(2):337–351.
[4] Li XJ, Wang HG, Li HB, et al. Awns play a dominant role in carbohydrate production during the grain-filling stages in wheat (Triticum aestivum L.). Physiol Plant. 2006;127(4):701–709.
[5] Kosuge K, Watanabe N, Kuboyama T, et al. Cytological and microsatellite mapping of mutant genes for spherical grain and compact spikes in durum wheat. Euphytica. 2008;159(3):289–296.
[6] Abebe T, Wise RP, Skadsen RW. Comparative transcriptional profiling established the awn as the major photosynthetic organ of the Barley spike while the lemma and the palea primarily protect the seed. Plant Genome. 2009;2(3):247–259.
[7] Elbaum R, Zaltzman L, Burgert I, et al. The role of wheat awns in the seed dispersal unit. Science. 2007;316(5826):884–886.
[8] Li XF, Du B, Wang HG. Awn anatomy of common wheat (Triticum aestivum L.) and its relatives. Caryologia. 2010;63(4):391–397.
[9] McIntosh RA. Catalogue of gene symbols for wheat. 12th International wheat genetics symposium, Yokohama 2013. pp. 6–7.
[10] McIntosh RA, Hart GE, Devos KM, et al. Catalogue of gene symbols for wheat. Proc 9th Int Wheat Genet Symp. Saskatoon. 1998;5:235.
[11] Sourdille P, Cadalen T, Gay G, et al. Molecular and physical mapping of genes affecting awning in wheat. Plant Breed. 2002;121(4):320–324.
[12] Yoshioka M, lehisa JCM, Ohno R, et al. Three dominant awnless genes in common wheat: fine mapping, interaction and contribution to diversity in awn shape and length. Plos One. 2017;12(4):e0176148.
[13] Zhang JP, Zhang ZY, Sun XM, et al. Natural alleles of GLA for grain length and awn development were differentiably domesticated in rice subspecies japonica and indica. Plant Biotechnol. J. 2019;17(8):1547–1559.
[14] Wang DZ, Yu K, Jin D, et al. Natural variations in the promoter of awn length inhibitor 1 (ALI-1) are associated with awn elongation and grain length in common wheat. Plant J. 2020;101(5):1075–1090.
[15] Huang DQ, Zheng Q, Melchkart T, et al. Dominant inhibition of awn development by a putative zinc-finger transcriptional repressor expressed at the B1 locus in wheat. New Phytol. 2020;225(1):340–355.
[16] Dewitt N, Guedira M, Lauer E, et al. Sequence-based mapping identifies a candidate transcription repressor underlying awn suppression at the B1 locus in wheat. New Phytol. 2020;225(1):326–339.
[17] Li H, Han YC, Guo XX, Xue F, et al. Genetic effect of locus B2 inhibiting awning in double-ditelsosomic 6B of Triticum durum DR147. Genet Resour Crop Evol. 2015;62(3):407–418.
[18] Luo JH, Liu H, Zhou TY, et al. An-1 encodes a basic helix-loop-helix protein that regulates awn development, grain size, and grain number in rice. Plant Cell. 2013;25(9):3360–3376.
[19] Prokopik DO, Ternovs'ka TK. SSR marking of genes used for control of beardedness in Durum wheat (Triticum durum Desf). Visn. Ukr. Tovar. Genet. Selekt. 2010;8:31–40.
[20] Korzun V, Röder MS, Worland AJ, et al. Intrachromosomal mapping of genes for dwarfing (Rht12) and vernalization response (Vrn1) in wheat using RFLP and microsatellite markers. Plant Breed. 1997;116(3):227–232.
[21] Haque MA, Takayama A, Watanabe N, et al. Cytological and genetic mapping of the gene for four-awned phenotype in Triticumcarthlicum Nevski. Genet Resour Crop Evol. 2011;58(7):1087–1093.
[22] Nishijima R, Ikeda TM, Takumi S. Genetic mapping reveals a dominant awn-inhibiting gene related to differentiation of the variety anathera in the wild diploid wheat Aegilops tauschii. Genetica. 2018;146(1):75–84.
[23] Salunkhe AS, Poornima R, Prince KSJ, et al. Fine mapping QTL for drought resistance traits in Rice (Oryza sativa L.) using bulk segregant analysis. Mol Biotechnol. 2011;49(1):90–95.
[24] Takagi H, Abe A, Yoshida K, et al. QTL-seq: Rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. Plant J. 2013;74(1):174–183.
[25] Kadambi G, Vemireddy LR, Srividhya A, et al. QTL-seq-based genetic analysis identifies a major genomic region governing dwarfness in rice (Oryza sativa L.). Plant Cell Rep. 2018;37(4):677–687.
[26] Cleveenger J, Chu Y, Chavarro C, et al. Mapping late leaf spot resistance in Peanut (Arachishypogaea) using QTL-seq reveals markers for marker-assisted selection. Front Plant Sci. 2018;9:83.
[27] Chen Q, Song J, Du WP, et al. Identification and genetic mapping for Rht-DM, a dominant dwarfing gene in mutant semi-dwarf maize using QTL-seq approach. Genes Genomics. 2018;40(10):1091–1099.
[28] Park M, Lee JH, Han K, et al. A major QTL and candidate genes for capsicainoid biosynthesis in the pericarp of Capsicum chinense revealed using QTL-seq and RNA-seq. Theor Appl Genet. 2019;132(2):515–529.
[29] Yuo T, Yamashita Y, Kanamori H, et al. A short internode (SHI) family transcription factor gene regulates awn elongation and pistil morphology in barley. J Exp Bot. 2012;63(14):5223–5232.
[30] Yan ZH, Wan YF, Liu KF, et al. Identification of a novel hmwglutenin subunit and comparison of its amino acid sequences. Euphytica. 2008;159(3):289–296.
acid sequence with those of homologous subunits. Sci Bull. 2002; 47:222–226.

[31] Poland JA, Brown PJ, Sorrells ME, et al. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PloS One. 2012;7(2):e32253

[32] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754–1760.

[33] Wang H, Wang XE, Chen PD, et al. Assessment of genetic diversity of Yunnan, Tibetan, and Xinjiang wheat using SSR markers. J Genet Genom. 2007;34(7):623–633.