

# Analysis of agronomic and domestication traits in a durum × cultivated emmer wheat population using a high-density single nucleotide polymorphism-based linkage map

Justin D. Faris · Qijun Zhang · Shiaoman Chao ·  
Zengcui Zhang · Steven S. Xu

Received: 28 May 2014 / Accepted: 15 August 2014 / Published online: 4 September 2014  
© Springer-Verlag Berlin Heidelberg (outside the USA) 2014

## Abstract

**Key message** Development of a high-density SNP map and evaluation of QTL shed light on domestication events in tetraploid wheat and the potential utility of cultivated emmer wheat for durum wheat improvement.

**Abstract** Cultivated emmer wheat (*Triticum turgidum* ssp. *dicoccum*) is tetraploid and considered as one of the eight founder crops that spawned the Agricultural Revolution about 10,000 years ago. Cultivated emmer has non-free-threshing seed and a somewhat fragile rachis, but mutations in genes governing these and other agronomic traits occurred that led to the formation of today's fully domesticated durum wheat (*T. turgidum* ssp. *durum*).

Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the US Department of Agriculture. USDA is an equal opportunity provider and employer.

Communicated by Takao Komatsuda.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00122-014-2380-1) contains supplementary material, which is available to authorized users.

J. D. Faris · S. Chao · Z. Zhang · S. S. Xu  
USDA-Agricultural Research Service, Cereal Crops Research  
Unit, Red River Valley Agricultural Research Unit, Fargo,  
ND 58102, USA

J. D. Faris (✉)  
USDA-ARS, NCSL, 1605 Albrecht BLVD, Fargo,  
ND 58102-2765, USA  
e-mail: justin.faris@ars.usda.gov

Q. Zhang  
Department of Plant Sciences, North Dakota State University,  
Fargo, ND 58108, USA

Here, we evaluated a population of recombinant inbred lines (RILs) derived from a cross between a cultivated emmer accession and a durum wheat variety. A high-density single nucleotide polymorphism (SNP)-based genetic linkage map consisting of 2,593 markers was developed for the identification of quantitative trait loci. The major domestication gene *Q* had profound effects on spike length and compactness, rachis fragility, and threshability as expected. The cultivated emmer parent contributed increased spikelets per spike, and the durum parent contributed higher kernel weight, which led to the identification of some RILs that had significantly higher grain weight per spike than either parent. Threshability was governed not only by the *Q* locus, but other loci as well including *Tg-B1* on chromosome 2B and a putative *Tg-A1* locus on chromosome 2A indicating that mutations in the *Tg* loci occurred during the transition of cultivated emmer to the fully domesticated tetraploid. These results not only shed light on the events that shaped wheat domestication, but also demonstrate that cultivated emmer is a useful source of genetic variation for the enhancement of durum varieties.

## Introduction

Wheat was domesticated in the Fertile Crescent of the Middle East and today is a major source of sustenance for humans providing about 20 % of the caloric intake. The two primary cultivated forms of modern wheat are common (bread) wheat (*Triticum aestivum* L.,  $2n = 6x = 42$ , AABBDD genomes), and durum (macaroni) wheat (*T. turgidum* ssp. *durum* L.,  $2n = 4x = 28$ , AABB genomes). Common wheat, which is hexaploid, accounts for about 95 % of the global wheat crop and is used primarily to

make breads, cookies, cakes, crackers, noodles, etc., and durum wheat, which accounts for about 5 % of the global wheat crop, is used to make pasta and other semolina-based products.

Durum and common wheat arose from the same evolutionary lineage. About a half million years ago (Huang et al. 2002; Chalupska et al. 2008), a hybridization event between the wild diploid wheat *T. urartu* Tumanian ex Gandylia ( $2n = 2x = 14$ ; AA genome) as the donor of the A genome (Dvorak et al. 1993) and a close relative of the diploid goatgrass *Aegilops speltoides* Tausch ( $2n = 2x = 14$ , SS genome) as the donor of the B genome (Dvorak and Zhang 1990; Blake et al. 1999; Huang et al. 2002; Chalupska et al. 2008; Salse et al. 2008) resulted in the formation of tetraploid wild emmer wheat *T. turgidum* ssp. *dicoccoides* (Körn.) Thell ( $2n = 4x = 28$ , AABB genomes).

Wild emmer had a very brittle rachis, which provided a natural mechanism of seed dispersal because the spikelets would disarticulate and fall to the ground upon maturity. However, about 10,000 years ago, forms of emmer wheat with non-brittle rachises emerged in the Fertile Crescent (Nesbitt and Samuel 1996) due to genetic mutations in the brittle rachis (*Br*) genes on chromosomes 2A (Peng et al. 2003; Peleg et al. 2011; Thanh et al. 2013), 3A and 3B (Watanabe and Ikebata 2000; Nalam et al. 2006; Li and Gill 2006; Thanh et al. 2013). This non-shattering form is referred to as cultivated emmer wheat (*T. turgidum* ssp. *dicoccum* L.,  $2n = 4x = 28$ , AABB genomes), and was instrumental in spawning the Agricultural Revolution of the Neolithic Times due to the fact that the non-brittle rachis trait allowed early farmers to more efficiently harvest the grain (Tzarfati et al. 2013).

Although cultivated emmer had a non-brittle rachis, the rachis was still somewhat fragile and the seed hulled (non-free-threshing). Mutations in genes governing seed threshability (see below) led to the formation of free-threshing *T. turgidum* subspecies, such as durum wheat, with fully tough rachises. About 8,000 years ago (Huang et al. 2002), *T. turgidum* (most likely a free-threshing form) hybridized with the diploid *A. tauschii* Coss. ( $2n = 2x = 14$ , DD genome) leading to the formation of hexaploid common wheat (Kihara 1944; McFadden and Sears 1946).

Two major genes prohibit the free-threshing trait in wheat: the *q* locus on chromosome arm 5AL and the homoeologous tenacious glume (*Tg*) gene set on the short arms of homoeologous group 2 chromosomes. A mutation in the *q* allele on 5AL about 10,000 years ago resulted in the formation of the hyperfunctionalized *Q* allele (Simons et al. 2006; Zhang et al. 2011), which confers the free-threshing character and also influences a repertoire of other domestication-related traits including rachis fragility,

glume tenacity, plant height, and heading time (Watkins 1940; Sears 1956; Muramatsu 1963, 1979, 1985, 1986; Kato et al. 1999, 2003; Faris and Gill 2002; Faris et al. 2003, 2005; Simons et al. 2006; Zhang et al. 2011). The molecular cloning of the *Q* gene revealed that it is a member of the AP2 class of transcription factors (Faris et al. 2003; Simons et al. 2006). For reasons argued in other articles (Matsuoka 2011; Dvorak et al. 2012; Faris 2014), the mutation from *q* to *Q* most likely happened in a tetraploid *T. turgidum* subspecies, which was then involved in the amphiploidization event with *A. tauschii* that resulted in the formation of hexaploid wheat thereby transferring the *Q* allele from tetraploid to hexaploid wheat (Simons et al. 2006).

Although the first hexaploid wheat probably had *Q*, it would have been non-free-threshing due to acquisition of the *Tg-D1* gene from *A. tauschii*, which confers very tenacious glumes and is epistatic to *Q* (Kerber and Rowland 1974). This trait was first observed by Kerber and Dyck (1969) and later mapped by Kerber and Rowland (1974) to the short arm of chromosome 2D using cytogenetic stocks. More recent mapping experiments have confirmed the location of *Tg-D1* on 2DS relative to molecular markers (Jantaturiyarat et al. 2004; Nalam et al. 2007; Sood et al. 2009; Dvorak et al. 2012).

Simonetti et al. (1999) evaluated a wild emmer (*T. turgidum* ssp. *dicoccoides*) × durum (*T. turgidum* ssp. *durum*) population for QTL associated with threshability and identified a major QTL on 5AL corresponding to the *Q* locus as well as a QTL on 2BS, which was suggested to be a homoeoallele of *Tg-D1* on 2DS. Faris et al. (2014a) recently demonstrated that the gene conferring tenacious glumes on 2BS was indeed homoeologous to *Tg-D1*, and therefore designated *Tg-B1*. Faris et al. (2014a) also showed that chromosome 2A of at least some wild emmer accessions harbored factors causing reduced threshability, suggesting the possible existence of a *Tg-A1* allele on 2AS as well. Dvorak et al. (2012) also provided evidence for the existence of *Tg-A1* alleles. Therefore, in regards to seed threshability, wild emmer would have the genotype *Tg-A1Tg-A1/Tg-B1Tg-B1/qq*.

On the contrary, fully domesticated durum wheat would necessarily have mutations at all three loci and have the genotype *tg-A1tg-A1/tg-B1tg-B1/QQ* because it is completely free-threshing. Cultivated emmer wheat (*T. turgidum* ssp. *dicoccum*) is intermediate to wild emmer and domesticated durum wheat, both in terms of position on the evolutionary lineage and in degree of threshability, i.e., cultivated emmer is non-free-threshing, but not as difficult to thresh as wild emmer. Cultivated emmer is known to have the *q* allele, which itself confers the non-free-threshing trait and a somewhat fragile rachis, but it is not known if cultivated emmer might possess *Tg-A1*, *Tg-B1*, or both alleles.

In other words, while it is known that the mutation of  $q$  to  $Q$  occurred during the transition of cultivated emmer to durum wheat, it is not known if the mutations from  $Tg-A1$  to  $tg-A1$  and  $Tg-B1$  to  $tg-B1$  occurred during the transition of wild to cultivated emmer, or from cultivated emmer to fully domesticated durum.

Other reports of the evaluation of domestication traits in tetraploid wheat mostly involved populations derived from crosses between wild emmer and durum wheat (Simonetti et al. 1999; Peng et al. 2003) or wild emmer and cultivated emmer (Thanh et al. 2013), but to our knowledge, studies investigating transitions in domestication traits between cultivated emmer and durum wheat have not been conducted. Here, we report the identification of QTL associated with domestication, spike morphology, and yield-related traits in a cultivated emmer  $\times$  durum wheat recombinant inbred line (RIL) population. The results provide further insights into the genetic factors and events that shaped the domestication of both durum and bread wheat.

## Materials and methods

### Plant materials

A population consisting of 200 recombinant inbred lines (RILs) was developed by crossing the durum wheat variety Ben (PI 596557) with the cultivated emmer wheat accession PI 41025 and advancing by the single seed-decent method to the  $F_{7,8}$  generation. Ben is a North Dakota hard amber durum variety (Elias and Miller 1998) and PI 41025 was collected near Samara, Russia. Seed of PI 41025 was obtained from the USDA-ARS National Small Grains Collection in Aberdeen, ID. The population of 200 RILs will hereafter be referred to as the BP025 population.

### Phenotypic and statistical analysis

The parents and BP025 population were evaluated for the number of days to heading (DTH), spike length (SL), number of spikelets per spike (SPS), spike compactness (CMP), kernels per spike (KPS), grain weight per spike (GWS), thousand kernel weight (TKW), rachis fragility (FRA), and threshability (THR). All plants were grown in the greenhouse in a completely randomized design with one plant per pot, and each pot was considered as an experimental unit. Plants were grown in six-inch clay pots with Metro Mix 902 soil (Hummert International, Earth City, MO) and supplemental fertilizer at a temperature of 21 °C and a 16/8 h light/dark photoperiod. Traits DTH, SL, SPS, CMP, FRA, and THR were evaluated in two greenhouse seasons with three replications, whereas data for KPS, GWS, and TKW were evaluated with two replications.

DTH was measured as the number of days from planting to emergence of the first spike beyond the flag leaf. Measurements for SL, SPS, and CMP were conducted as described in Faris et al. (2014). Briefly, four main spikes were harvested at maturity and placed in a dryer at 32 °C for at least 24 h. SL was determined by measuring the spikes from the base of the first spikelet to the tip of the most terminal spikelet excluding the awns. SPS was calculated by counting the number of spikelets, and CMP was determined by dividing the spike length by the number of spikelets. KPS was calculated by counting the number of seeds produced by the four main spikes and dividing the number by 4. Similarly, GWS was determined by weighing the total seed sample from the four main spikes and dividing that number by 4. TKW for each line was determined by weighing 100 seeds and multiplying the number by 10.

The same four spikes from each plant used to calculate SL, SPS, CMP, KPS, and GWS were used to measure FRA and THR. FRA was measured by manually breaking the spike near the middle of the spike and scoring the ease of breakage on a scale from 1 to 4, where 1 = most difficult to break (as difficult as the fully tough rachis parent Ben); 2 = somewhat difficult to break but not as difficult as Ben; 3 = somewhat easy to break but not as easy as PI 41025; and 4 = very easy to break (as easy as PI 41025). THR was scored precisely as described in Faris et al. (2014a) using the 1–4 scale for ease of threshing.

For traits SL, SPS, FRA, and THR, the scores for the four spikes from each plant were averaged to derive the mean score for each plant within the replicate. For all traits, the mean scores for each plant were subjected to Bartlett's Chi squared test for homogeneity of error variances among replications using PROC GLM in the SAS program version 9.3 (SAS institute 2011). Error variances were homogeneous among replications for all traits with the exception of DTH (data not shown). Therefore, the scores of each replicate for traits SL, SPS, CMP, KSP, GWS, TKW, FRA, and THR were used to calculate the overall mean for each RIL, and the mean values were used in subsequent analyses. Because the DTH data were not statistically homogeneous, means were not calculated and each replicate was analyzed separately for associated QTL. Fisher's Least Significant Difference (LSD) test was used to determine significant differences among means of the RILs at the 0.05 level of probability.

### Genotyping and linkage mapping

The BP025 population along with the parental lines was genotyped using the iSelect array containing 9000 wheat single nucleotide polymorphism (SNP) markers (Cavanagh et al. 2013). The SNP markers were designated with "IWA" followed by their index number. The genotyping assay was

carried out using Illumina's Infinium assay following the manufacturer's protocols. SNP clustering and genotype calling were performed using Illumina's GenomeStudio software v.2011.1. The genotype callings were manually inspected to correct cluster shifts due to copy number differences and to ensure call accuracy for every SNP.

In addition, several simple sequence repeat (SSR) markers with known chromosome locations were selected from each chromosome to help anchor genetic linkage groups to specific chromosomes. These SSR primers were selected from BARC (Song et al. 2005), CFA (Sourdille et al. 2003), CFD (Sourdille et al. 2004), WMS (marker designations = 'gwm') (Röder et al. 1998), GDM (Pestsova et al. 2000), and WMC (Somers et al. 2004) sets and tested for polymorphism between Ben and PI 41025 using the PCR conditions described by Röder et al. (1998).

The SSR primer set FCP650, which detects an SSR within the last intron of the *Q* gene (Simons et al. 2006) and has the marker designation *Xfcp650(Q)*, was used to map the *Q* gene on chromosome 5A. The primer sequences for the FCP650 set are 5' GCACTAGCTAATTCAGTGGTTAGATTTGCTCA 3' and 5' ATTCAGTGGTAGCAACAGTTTCAGTAAGCTGG 3' and an annealing temperature of 65 °C was used.

Four sequence-tagged site (STS) markers previously developed by Abeysekara et al. (2010) for mapping the *Tsc2* gene on chromosome 2B were also mapped in the BP025 population. One of the four markers was derived from the expressed sequence-tag (EST) BF484320, and the other three were developed from the tentative

consensus (TC) sequences TC454839, TC432379, and TC422669.

Linkage maps were assembled using the computer program MapDisto 1.7 (Lorieux 2012). Markers were first organized into groups using the 'find groups' command with a minimum LOD = 3.0 and a maximum theta of 0.30. The 'order' sequence command was used to establish the initial order of markers within a linkage group. Subsequent interrogation of the sequence using the 'check inversions,' 'ripple order,' and 'drop locus' commands was conducted to determine the best map. Map distances were calculated using the Kosambi mapping function (Kosambi 1944).

#### QTL analysis

The DTH data for individual replicates and the overall means for all other traits were used in QTL analysis. Composite interval mapping (CIM) was performed using the computer program QGene 4.3 (Joehanes and Nelson 2008). A 1,000-iteration permutation test revealed a critical LOD threshold of 3.2 for the 0.05 level of probability.

## Results

### Map construction

Genotyping of the BP025 population with the 9 K iSelect SNP chip resulted in 2,461 segregating markers that were

**Table 1** Chromosome assignment and distribution of markers, length of chromosome linkage groups, and marker density of maps generated in the Ben × PI 41025 (BP025) recombinant inbred population

Chromosome	EST-STS	SSR	SNP	Total markers	Length (cM)	cM/marker	Number of distorted markers (%)
1A	0	7	272	279	147.8	0.53	44 (16)
1B	0	10	125	135	167.6	1.24	19 (14)
2A	0	11	155	166	227.3	1.37	94 (57)
2B	4	28	303	335	237.2	0.71	214 (64)
3A	0	13	190	203	177.8	0.94	13 (6)
3B	0	8	232	240	177.8	0.74	64 (27)
4A	0	7	124	131	124.1	0.95	16 (12)
4B	0	4	85	89	118.5	1.33	0 (0)
5A	0	12	140	152	217.3	1.43	41 (27)
5B	0	7	163	170	202.8	1.19	97 (57)
6A	0	3	144	147	143.0	0.97	51 (35)
6B	0	4	172	176	145.8	0.83	7 (4)
7A	0	9	201	210	181.1	0.86	56 (27)
7B	0	5	155	160	176.3	1.10	5 (3)
A genome	0	62	1,226	1,288	1,218.4	0.95	315 (24)
B genome	4	66	1,235	1,305	1,226.0	0.94	406 (31)
Total	4	128	2,461	2,593	2,444.4	0.94	721 (28)

used for genetic linkage mapping (Table 1, Online Resource 1). In addition to the SNP markers, 128 SSR markers and four EST-STS markers were mapped in the BP025 population for a total of 2,593 markers, which detected 967 unique loci and accounted for a total genetic map length of 2,444.4 cM. This resulted in an average whole-genome marker density of one marker every 0.94 cM.

The average number of markers per chromosome was 185.2 and the number ranged from 89 for chromosome 4B to 335 for chromosome 2B (Table 1). Similarly, the average genetic length over all chromosome maps was 174.6 cM, and the map lengths for individual chromosomes ranged from 118.5 cM for 4B to 237.2 cM for 2B. Chromosome 1A was the most densely covered with one marker every 0.53 cM, and chromosome 5A was the least dense with one marker per 1.43 cM.

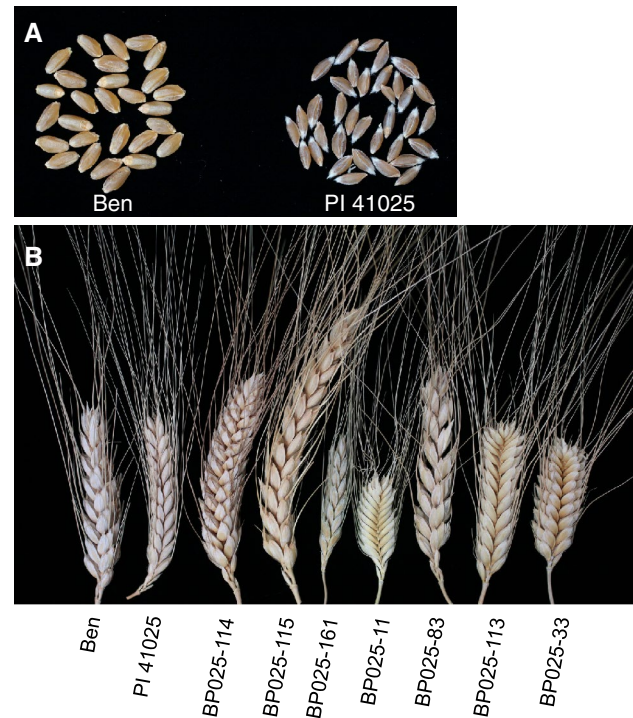
The number of markers, genetic map lengths, and marker densities for the A and B genomes were nearly equal. The A genome had 1,288 markers spanning 1,218.4 cM for a density of one marker per 0.95 cM, and the B genome had 1,305 markers spanning 1,226.0 cM with a marker density of one per 0.94 cM (Table 1).

Of the 2,593 markers, a total of 721 (28 %) had segregation ratios that significantly ( $P < 0.05$ ) deviated from the expected 1:1 ratio (Table 1, Online Resource 1). Nearly all the distorted markers occurred in large clusters (Online Resource 1). With the exception of chromosome 4B, all chromosomes harbored some distorted markers with chromosome 2B having the most (214, 64 %). Chromosome 5B possessed a large cluster of the most severely distorted markers ( $P < 0.00001$ ), which coincides with a genomic region of segregation distortion in tetraploid wheat studied in detail by Kumar et al. (2007).

#### Trait evaluations

The DTH data were not homogeneous and therefore each replicate was evaluated separately in the QTL analysis. However, we used overall means from the three replicates to determine the mean and range of the BP025 population, and to evaluate correlations with other traits. The average DTH for the parents of the RIL population differed by only a single day with Ben and PI 41025 having values of 59.67 and 58.67, respectively, and were therefore not significantly different (Table 2). However, the population ranged from 42.67 to 87.0 DTH indicating that the two parents possess different genes governing heading time.

Seed of Ben and PI 41025 along with mature spikes of Ben, PI 41025, and representative spikes from several BP025 RILs are presented in Fig. 1. Ben had an average SL that was 0.66 cm shorter than the average SL of PI 41025, but they did not differ significantly (Figs. 1, 2; Table 2). The average SL among all the RILs in the BP025 population was 6.0 cm, which was the midpoint value between the



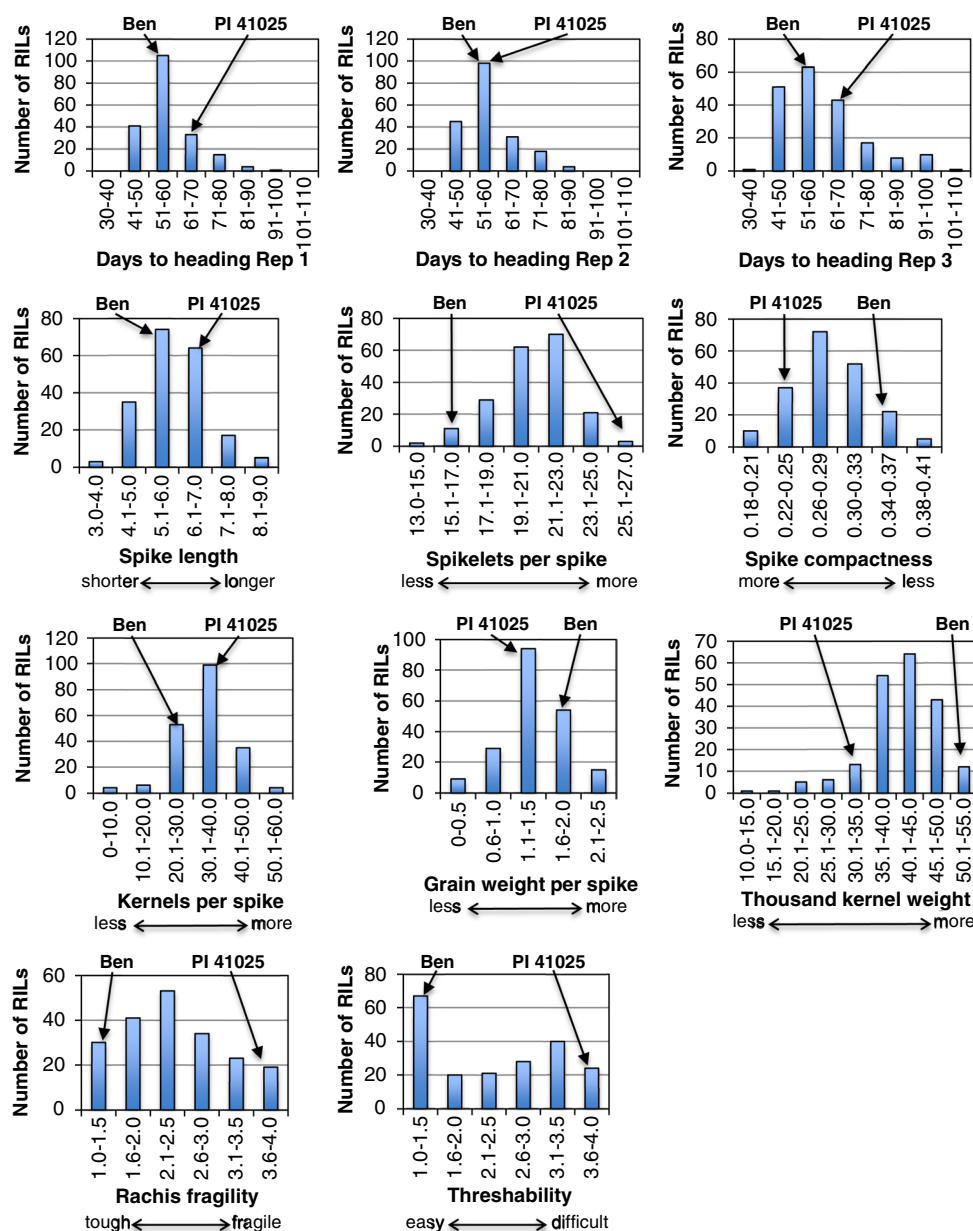
**Fig. 1** Seed and spike morphology of the parents and several recombinant inbred lines of the Ben  $\times$  PI 41025 (BP025) population. (a) Seed of the durum variety Ben (left) and the cultivated emmer wheat accession PI 41025 (right). (b) Mature spikes of Ben, PI 41025, and several recombinant inbred lines of the BP025 population

**Table 2** Parental and population means, ranges, and least significant differences (LSD) at the 0.05 level of probability ( $P < 0.05$ ) for the traits days to heading (DTH), spike length (SL), spikelets per spike (SPS), spike compactness (CMP), kernels per spike (KPS), grain weight per spike (GWS), thousand kernel weight (TKW), rachis fragility (FRA), and threshability (THR)

Trait	Mean			BP population range	LSD (0.05)
	Ben	PI 41025	BP population		
DTH	59.67	58.67	58.79	42.67–87.00	8.89
SL	5.94	6.60	6.00	3.36–8.81	0.68
SPS	16.53	26.64	20.68	13.03–26.03	2.18
CMP	0.36	0.25	0.29	0.18–0.42	0.03
KPS	29.13	36.0	33.64	8.13–54.63	10.96
GWS	1.60	1.04	1.40	0.16–2.35	0.60
TKW	55.12	28.84	40.84	10.28–54.33	8.90
FRA	1.33	3.92	2.36	1.00–4.00	1.12
THR	1.00	3.75	2.26	1.00–4.00	1.00

average parental SL values. However, RIL population had SL values that ranged from 3.36 to 8.81 cm, again suggesting that the parents harbor different genes governing SL.

Ben and PI 41025 had average SPS values of 16.53 and 26.64, respectively (Fig. 2; Table 2). Therefore, on average,



**Fig. 2** Histograms of the Ben  $\times$  PI 41025 (BP025) recombinant inbred population for traits days to heading (DTH), spike length (SL), spikelets per spike (SPS), spike compactness (CMP), kernels per spike (KPS), grain weight per spike (GWS), thousand kernel weight

(TKW), rachis fragility (FRA), and threshability (THR). For DTH, histograms for each individual replication were generated because the replications were not statistically homogeneous and therefore the data was not combined

PI 41025 spikes had 10.11 more spikelets per spike than did Ben. The RILs of the BP025 population had an average of 20.68 SPS, and the range across the population was 13.03–26.03 suggesting that PI 41025 was the sole contributor of genetic factors conferring an increased number of SPS.

Spike compaction (CMP) was calculated by dividing the SL by the SPS. Therefore, CMP values are a function of SL and SPS. Spikes of PI 41025 were significantly more compact than those of Ben with average CMP values of 0.25 and 0.36, respectively (Fig. 2; Table 2). Therefore, although

Ben had slightly (albeit insignificantly) shorter spikes than PI 41025, it had less compact spikes than PI 41025 due to the fact that PI 41025 had significantly more SPS than did Ben. The BP025 population had an average overall CMP value of 0.29 and a range of 0.18–0.42 suggesting that the increased SPS in PI 41025 was not the only factor contributing to CMP and that both parents likely harbor gene(s) that influence this trait.

Ben and PI 41025 averaged 29.13 and 36.0 kernels per spike (KPS), respectively (Fig. 2; Table 2). Therefore,

**Table 3** Correlation coefficients between the mean values of the traits days to heading (DTH), spike length (SL), spikelets per spike (SPS), spike compactness (CMP), kernels per spike (KPS), grain

weight per spike (GWS), thousand kernel weight (TKW), rachis fragility (FRA), and threshability (THR)

Trait	DTH	SL	SPS	CMP	KPS	GWS	TKW	FRA	THR
DTH	1.0								
SL	−0.36**	1.0							
SPS	0.07	0.30**	1.0						
CMP	−0.36**	0.73**	−0.37**	1.0					
KPS	−0.12	0.37**	0.43**	0.04	1.0				
GWS	−0.28**	0.40**	0.28**	0.16*	0.87**	1.0			
TKW	−0.37**	0.30**	−0.04	0.30**	0.30**	0.69**	1.0		
FRA	−0.08	0.14*	−0.11	0.23**	−0.15*	−0.15*	−0.12	1.0	
THR	−0.08	0.38**	0.14*	0.27**	0.04	−0.03	−0.07	0.41**	1.0

\*  $P < 0.05$ \*\*  $P < 0.01$ 

although PI 41025 averaged 10.11 more SPS than Ben, it averaged only seven more kernels. The BP025 population averaged 33.64 KPS but had a range of 8.13–54.63 KPS suggesting that both PI 41025 and Ben harbored genes that confer increased KPS, but the genes were different.

Although PI 41025 produced more seed than Ben, the GWS for Ben (1.60 g) was more than that for PI 41025 (1.04 g) (Fig. 2; Table 2). The RIL population averaged 1.40 g and ranged from 0.16 to 2.35 g, again indicating that Ben and PI 41025 both contribute to increased GWS, but through different genes.

For TKW, Ben had much heavier seed at 55.12 g compared to only 28.84 g for PI 41025 (Fig. 2; Table 2). The RIL population averaged 40.84 g and ranged from 10.28 to 54.33 g. It is likely that PI 41025 does not harbor alleles that contribute to increased seed weight because no RIL had a heavier TKW than Ben.

Ben had a fully tough rachis and was completely free-threshing resulting in average FRA and THR values of 1.33 and 1.00, respectively (Fig. 2; Table 2). Conversely, PI 41025 had a relatively fragile rachis that broke easily and seed that was not free-threshing resulting in average FRA and THR values of 3.92 and 3.75, respectively. The BP025 population had average FRA and THR values of 2.36 and 2.26, respectively, and a range of 1.00–4.00 for both traits suggesting that Ben was responsible for contributing the genetic factors controlling a tough rachis and free-threshing seed.

The correlation coefficients among the six traits evaluated are presented in Table 3. DTH was negatively correlated with SL, CMP, GWS, and TKW. SL had a positive correlation with all traits. In addition to a positive correlation with SL, KPS, and GWS; SPS had a weak positive correlation with THR and a strong negative correlation with CMP. CMP had strong negative correlations with DTH and SPS, and strong positive correlations with SL, TKW,

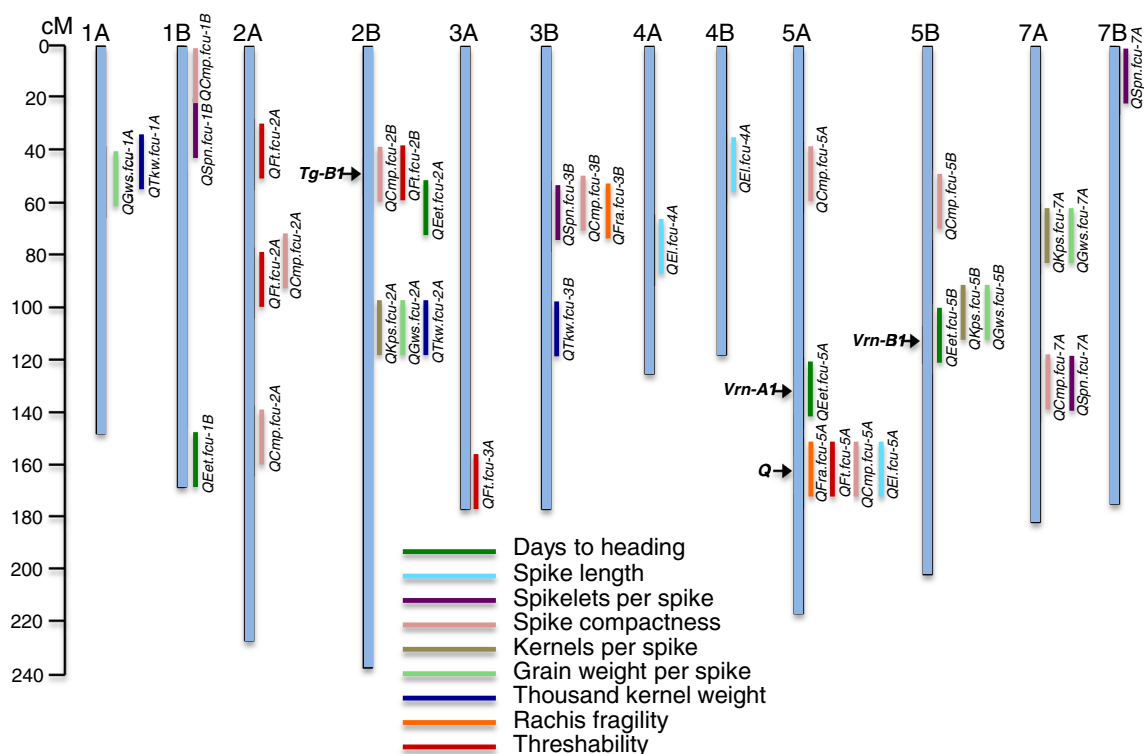
FRA, and THR as well as a weaker positive correlation with GWS. KPS was strongly correlated with SL, SPS, GWS, and TKW, and it had a weak negative correlation with FRA. GWS had weak negative correlations with DTH and FRA, and stronger positive correlations with SL, SPS, CMP, KPS, and TKW. FRA had positive correlations with SL, CMP, and THR, and weak negative correlations with KPS and GWS. THR was positively correlated with SL, SPS, CMP, and FRA.

#### QTL analysis

A total of four QTL associated with DTH were identified (Fig. 3; Table 4). The DTH QTL was identified on chromosomes 1B, 2B, 5A, and 5B, and designated *QEet.fcu-1B*, *QEet.fcu-2B*, *QEet.fcu-5A*, and *QEet.fcu-5B*, respectively. PI 41025 contributed the earliness effects of the 1B and 2B QTL, and they explained as much as 16.5 and 9.7 % of the variation in DTH, respectively. Ben contributed the earliness effects of the 5A and 5B QTL, which explained up to 19.2 and 17.6 % of the DTH variation for the different replicates, respectively.

Three QTL associated with SL were identified on chromosomes 4A, 4B, and 5A (Fig. 3), and they were designated *QEl.fcu-4A*, *QEl.fcu-4B*, and *QEl.fcu-5A* (Table 4). Ben contributed effects for decreased spike length at all three loci. The 4A, 4B, and 5A QTL had LOD values of 6.5, 7.0, and 26.5 and explained 9.1, 7.7, and 27.3 % of the variation in SL, respectively. *QEl.fcu-5A* peaked at marker *Xfcp650(Q)* on 5A indicating that the effects for decreased spike length for this QTL were conferred by the *Q* allele of Ben.

Four significant QTL were associated with SPS, and PI 41025 contributed the effects for increased SPS at all four loci (Fig. 3; Table 4). The four QTL, designated *QSpn.fcu-1B*, *QSpn.fcu-3B*, *QSpn.fcu-7A*, and *QSpn.fcu-7B*, were



**Fig. 3** Illustration of the chromosomal locations of the 37 quantitative trait loci (QTL) associated with the nine traits evaluated. A centiMorgan (cM) scale is shown on the left. The known positions of the

*Q*, *Tg-B1*, *Vrn-A1*, and *Vrn-B1* loci are indicated in black. Chromosomes 6A and 6B are not shown because no QTL was detected on them in this research

located on chromosomes 1B, 3B, 7A, and 7B, respectively. The LOD values for these QTL ranged from 4.7 to 7.8, and they explained from 6.2 to 11.8 % of the variation in SPS.

Nine QTL were found to be significantly associated with CMP (Fig. 3; Table 4). These consisted of single QTL on chromosomes 1B (*QCmp.fcu-1B*), 2B (*QCmp.fcu-2B*), 3B (*QCmp.fcu-3B*), 5B (*QCmp.fcu-5B*), and 7A (*QCmp.fcu-7A*), and two QTL each on chromosomes 2A (*QCmp.fcu-2A.1* and *QCmp.fcu-2A.2*) and 5A (*QCmp.fcu-5A.1* and *QCmp.fcu-5A.2*). With the exception of *QCmp.fcu-5A.2*, the effects for increased CMP at all the other QTL were contributed by PI 41025. These eight QTL had LOD values ranging from 4.2 to 6.8 and explained from 2.0 to 7.0 % of the variation in CMP. Ben contributed the effects for increased CMP at the *QCmp.fcu-5A.2* locus. This QTL had an LOD value of 36.2 and explained 31.6 % of the variation in CMP. The marker *Xfcp650(Q)* defined the peak of this QTL indicating that the *Q* allele from Ben was the genetic factor underlying the effects of the QTL.

Three QTL associated with KPS were identified (Fig. 3; Table 4). The QTL *QKps.fcu-2B* was located on chromosome 2B with an LOD of 17.7 and explained 15.2 % of the variation. Effects for increased KPS at this locus were contributed by Ben. PI 41025 contributed increased KPS for the other

two QTL, which were located on chromosomes 5B and 7A and designated *QKps.fcu-5B* and *QKps.fcu-7A*, respectively. These two QTL each explained 5.6 % of the variation.

The four QTL associated with GWS were located on chromosomes 1A, 2B, 5B, and 7A (Fig. 3; Table 4). The 2B QTL, *QGws.fcu-2B*, had the largest effects with an LOD of 14.0 and explaining 16.6 % of the variation. *QGws.fcu-1A*, *QGws.fcu-5B*, and *QGws.fcu-7A* had LOD values of 9.0, 4.6, and 5.0, and explained 6.4, 7.0, and 2.9 % of the variation, respectively. Ben contributed the effects for increased GWS at the 2B and 1A loci, and PI 41025 contributed increased GWS at the 5B and 7A loci. A survey of the mean GWS for all RILs indicated that five lines (BP025-41, BP025-50, BP025-114, BP025-115, and BP025-118) had significantly more GWS than either parent. Investigation of the genotypes harbored by these five RILs at each of the four QTL revealed that BP025-114 and BP025-115, the two RILs with the highest GWS among all RILs, harbored alleles for increased GWS at all four loci (Fig. 4). BP025-41 had the beneficial alleles at all loci except for *QGws.fcu-7A* and BP025-118 had harbored alleles for increased GWS at all loci except for *QGws.fcu-1A*. BP025-50 harbored only one of the four QTL alleles for increased GWS, which was *QGws.fcu-7A*.



**Table 4** Quantitative trait loci associated with the traits evaluated in the Ben × PI 41025 (BP025) recombinant inbred population

Trait	QTL	Chromosome	Position (cM)	Marker or marker interval	LOD	$R^2 \times 100$	Additive effects <sup>a</sup>
DTH <sup>b</sup>	<i>QEet.fcu-1B</i>	1B	166	<i>IWA742–IWA4525</i>	12.3, 5.6, 4.7	16.5, 9.8, 9.3	−5.8, −3.0, −2.8
	<i>QEet.fcu-2B</i>	2B	64	<i>IWA546–IWA8083</i>	3.7, 7.0, 6.3	3.2, 9.7, 7.3	−3.0, −3.5, −2.9
	<i>QEet.fcu-5A</i>	5A	126	<i>IWA4276–IWA4805</i>	13.8, 3.6, 7.7	19.2, 9.5, 14.5	5.9, 2.8, 3.3
	<i>QEet.fcu-5B</i>	5B	116	<i>IWA3226–IWA6718</i>	17.4, 9.3, 13.5	17.2, 13.8, 17.6	5.7, 3.4, 3.7
SL	<i>QEl.fcu-4A</i>	4A	78	<i>IWA2816–IWA811</i>	6.5	9.1	−0.3
	<i>QEl.fcu-4B</i>	4B	54	<i>IWA75–Xcfa2091</i>	7.0	7.7	−0.3
	<i>QEl.fcu-5A</i>	5A	162	<i>Xfcp650(Q)</i>	26.5	27.3	−0.5
SPS	<i>QSpn.fcu-1B</i>	1B	28	<i>Xbarc119</i>	7.1	6.2	−0.6
	<i>QSpn.fcu-3B</i>	3B	62	<i>IWA3716–IWA6201</i>	4.7	6.2	−0.6
	<i>QSpn.fcu-7A</i>	7A	132	<i>IWA4910–IWA5912</i>	7.8	11.8	−0.8
	<i>QSpn.fcu-7B</i>	7B	10	<i>IWA1089</i>	5.7	7.4	−0.6
CMP	<i>QCmp.fcu-1B</i>	1B	6	<i>IWA7399–IWA4754</i>	5.0	2.8	0.01
	<i>QCmp.fcu-2A.1</i>	2A	84	<i>IWA581–IWA2758</i>	6.8	3.5	0.01
	<i>QCmp.fcu-2A.2</i>	2A	148	<i>IWA173–IWA8627</i>	4.8	2.0	0.01
	<i>QCmp.fcu-2B</i>	2B	52	<i>IWA2117–IWA4284</i>	4.2	2.4	0.01
	<i>QCmp.fcu-3B</i>	3B	58	<i>IWA3896–IWA1458</i>	5.5	3.5	0.01
	<i>QCmp.fcu-5A.1</i>	5A	46	<i>IWA8582–IWA5184</i>	4.4	7.0	0.01
	<i>QCmp.fcu-5A.2</i>	5A	162	<i>Xfcp650(Q)</i>	36.2	31.6	−0.03
	<i>QCmp.fcu-5B</i>	5B	64	<i>IWA1956–IWA1520</i>	6.5	6.4	0.01
	<i>QCmp.fcu-7A</i>	7A	132	<i>IWA4910–IWA5912</i>	12.6	4.9	0.01
KPS	<i>QKps.fcu-2B</i>	2B	114	<i>IWA1217–IWA652</i>	17.7	15.2	4.2
	<i>QKps.fcu-5B</i>	5B	106	<i>IWA4094–IWA6915</i>	5.4	5.6	−2.2
	<i>QKps.fcu-7A</i>	7A	72	<i>IWA1845–IWA7192</i>	4.3	5.6	−1.8
GWS	<i>QGws.fcu-1A</i>	1A	48	<i>IWA7328–IWA267</i>	9.0	6.4	0.2
	<i>QGws.fcu-2B</i>	2B	114	<i>IWA1217–IWA652</i>	14.0	16.6	0.2
	<i>QGws.fcu-5B</i>	5B	106	<i>IWA4094–IWA6915</i>	4.6	7.0	−0.1
	<i>QGws.fcu-7A</i>	7A	72	<i>IWA1845–IWA7192</i>	5.0	2.9	−0.1
TKW	<i>QTKw.fcu-1A</i>	1A	44	<i>IWA5435–IWA3338</i>	4.6	4.7	1.6
	<i>QTKw.fcu-2B</i>	2B	114	<i>IWA1217–IWA652</i>	11.9	8.3	6.5
	<i>QTKw.fcu-3B</i>	3B	116	<i>IWA2462–IWA4324</i>	4.3	5.2	1.6
FRA	<i>QFra.fcu-3B</i>	3B	60	<i>IWA3716–IWA6201</i>	7.1	6.8	−0.2
	<i>QFra.fcu-5A</i>	5A	162	<i>Xfcp650(Q)</i>	18.6	27.7	−0.4
THR	<i>QFt.fcu-2A.1</i>	2A	40	<i>IWA4441–IWA5087</i>	4.6	5.7	−0.2
	<i>QFt.fcu-2A.2</i>	2A	88	<i>IWA26–Xgwm372</i>	9.7	5.4	−0.2
	<i>QFt.fcu-2B</i>	2B	48	<i>Xwmc154–IWA5762</i>	16.9	17.2	−0.5
	<i>QFt.fcu-3A</i>	3A	172	<i>IWA4905–IWA7812</i>	7.0	2.2	−0.2
	<i>QFt.fcu-5A</i>	5A	162	<i>Xfcp650(Q)</i>	28.9	30.9	−0.6

*DTH* days to heading, *SL* spike length, *SPS* spikelets per spike, *CMP* spike compactness, *KPS* kernels per spike, *GWS* grain weight per spike, *TKW* thousand kernel weight, *FRA* rachis fragility, *THR* threshability

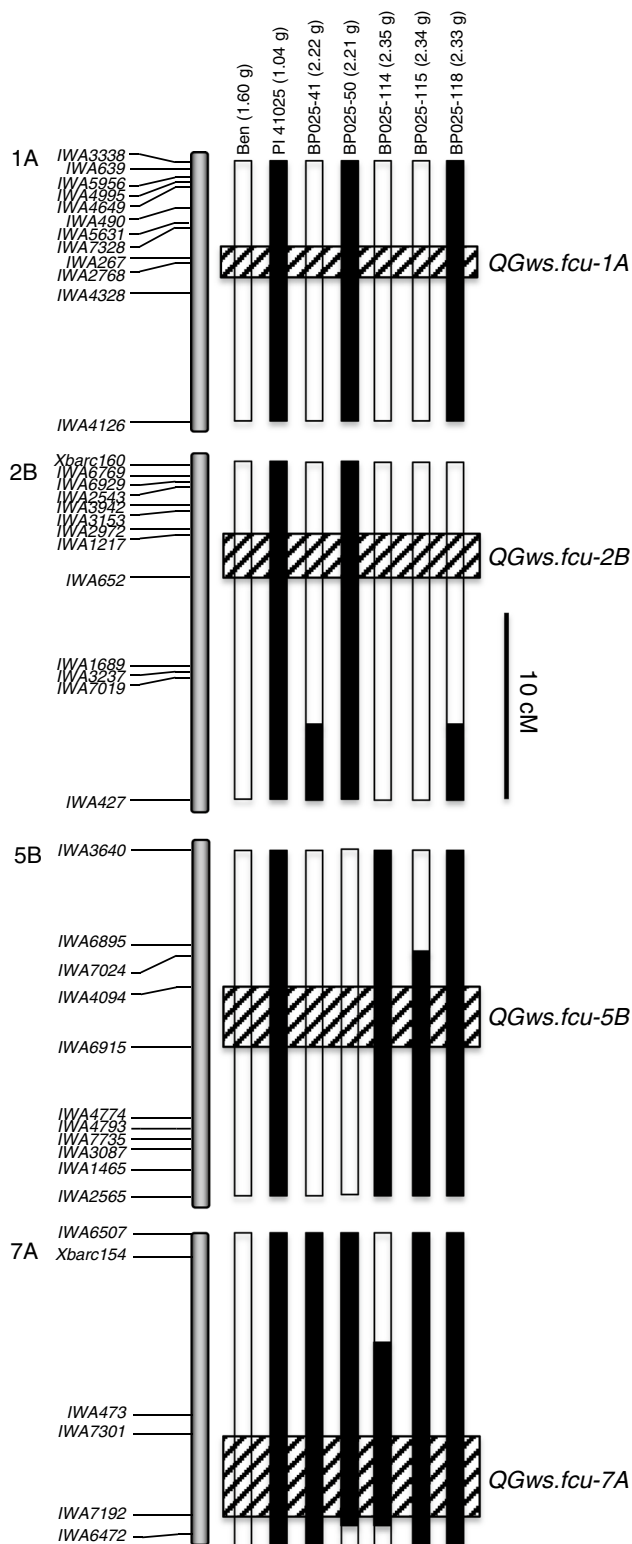
<sup>a</sup> A negative number value indicates that a higher value for a given trait is derived from PI 41025 and a positive number indicates that a higher value for the trait is derived from Ben

<sup>b</sup> The replicates of data for DTH were not homogeneous and therefore were analyzed separately. The three values presented for LOD,  $R^2 \times 100$ , and additive effects are the values derived from replicates 1, 2, and 3, respectively

Chromosomes 1A, 2B, and 3B carried QTL for TKW, and they were designated *QTKw.fcu-1A*, *QTKw.fcu-2B*, and *QTKw.fcu-3B*, respectively (Fig. 3; Table 4). The effects for increased TKW for all three loci were contributed by Ben,

and they had LOD and  $R^2$  values that ranged from 4.3 to 11.9 and 0.047 to 0.083, respectively.

QTL associated with FRA were identified on chromosomes 3B and 5A, and designated *QFra.fcu-3B* and



*QFra.fcu-5A*, respectively (Fig. 3; Table 4). *QFra.fcu-3B* had an LOD score of 7.1 and explained 6.8 % of the variation, whereas *QFra.fcu-5A*, for which the *Q* gene defined

**Fig. 4** Graphical genotypes of QTL regions of chromosomes 1A, 2B, 5B, and 7A, which harbor QTL for increased GWS. The graphical genotypes of Ben, PI 41025, and the RILs with significantly greater GWS than Ben including BP025-41, BP025-50, BP025-114, BP025-115, and BP025-118 are shown where *open bars* represent Ben chromatin and *black* represent PI 41025 chromatin. The QTL regions are indicated by the *hatched boxes*. A cM scale is indicated

the peak, had an LOD of 18.6 and explained 27.7 % of the variation in FRA. The effects of both QTL for rachis toughness were contributed by Ben.

The *Q* locus represented by marker *Xfcp650(Q)* was also the most significant marker associated with THR and defined the peak of the QTL *QFt.fcu-5A* on chromosome 5A (Fig. 3; Table 4). In addition to this QTL, two QTL associated with THR were identified on chromosome 2A (*QFt.fcu-2A.1* and *QFt.fcu-2A.2*) and one QTL each on chromosomes 2B (*QFt.fcu-2B*) and 3A (*QFt.fcu-3A*). The 5A QTL had the largest effects on THR with an LOD of 28.9 and explained 30.9 % of the variation in THR. The 2B QTL was the second largest QTL associated with THR with an LOD of 16.9 and explained 17.2 % of the variation. The 2A and 3A QTL had LOD values ranging from 4.6 to 9.7 and explained from 2.2 to 5.7 % of the phenotypic variation. The parent Ben contributed effects for ease of threshability for all five of these QTL.

#### Overlapping QTL

Evaluation of the nine traits described in this research led to the identification of a total of 37 QTL. However, the 37 QTL were represented by 24 unique loci on 12 chromosomes due to some loci likely having pleiotropic effects on multiple traits or being tightly linked (Fig. 3). QTL for GWS (*QGws.fcu-1A*) and TKW (*QTKw.fcu-1A*) co-localized on chromosome 1A. On homoeologous group 2 chromosomes, QTL for CMP and THR on 2A (*QCmp.fcu-2A.1* and *QFt.fcu-2A.2*) and 2B (*QCmp.fcu-2B* and *QFt.fcu-2B*) co-localized. Also, QTL for KPS (*QKps.fcu-2B*), GWS (*QGws.fcu-2B*) and TKW (*QTKw.fcu-2B*) coincided on chromosome 2B. The chromosome 3B QTL *QSpn.fcu-3B*, *QCmp.fcu-3B*, and *QFra.fcu-3B* all co-localized to the 58–62 cM position (Fig. 3; Table 4). QTL for CMP and SPN (*QCmp.fcu-7A* and *QSpn.fcu-7A*) co-localized on chromosome 7A, and QTL for KPS and GWS coincided at the same positions on chromosomes 5B (*QKps.fcu-5B* and *QGws.fcu-5B*) and 7A (*QKps.fcu-7A* and *QGws.fcu-7A*). Most remarkably, QTL for FRA, THR, CMP, and SL (*QFra.fcu-5A*, *QFt.fcu-5A*, *QCmp.fcu-5A*, and *QEl.fcu-5A*) all co-localized at the *Xfcp650(Q)* locus on chromosome 5A demonstrating the pleiotropic effects of the *Q* gene.

## Discussion

### Map construction

Most previous genetic linkage maps developed for durum or other tetraploid wheat subspecies have employed molecular marker platforms such as restriction fragment length polymorphisms (RFLPs), amplified fragment length polymorphisms (AFLPs), SSRs, diversity array technology (DArT), SNPs, or combinations thereof (Blanco et al. 1998; Elouafi and Nachit 2004; Mantovani et al. 2008; Chu et al. 2010; Buerstmayr et al. 2012; Marone et al. 2012; Colasuonno et al. 2013; Thanh et al. 2013). vanPoecke et al. (2013) recently conducted genotype-by-sequencing (GBS) demonstrating the utility of next generation sequencing for *de novo* SNP marker discovery and genotyping in durum wheat. The wheat iSelect 9 K array was developed for genotyping in hexaploid wheat (Cavanagh et al. 2013) and its use in tetraploid wheat has not been reported other than for marker discovery and linkage mapping of the short arm of chromosome 2B in a population of RILs derived from a cross between durum wheat and a durum-*T. turgidum* ssp. *dicoccoides* disomic chromosome 2B substitution line (Faris et al. 2014a). Here, we showed that the iSelect 9 K array originally developed for hexaploid wheat works very well for the development of high-density whole-genome linkage maps in tetraploid wheat as well.

### Days to heading

The primary objective of this work was to evaluate domestication-related traits, i.e., rachis fragility and threshability, to gain insights regarding the transition events that occurred during the evolution of cultivated emmer to durum wheat, but also to evaluate spike morphology and yield-related traits as well as heading time. There are three genetic systems in wheat that affect heading/flowering time in wheat: vernalization, photoperiod sensitivity, and earliness per se (see Kamran et al. 2014 for review). Of the four QTL identified for DTH, two were on the long arms of chromosomes 5A and 5B in the vicinity of the vernalization genes *Vrn-A1* and *Vrn-B1*, respectively. Therefore, it is probable that the effects of *QEet.fcu-5A* and *QEet.fcu-5B* are due to allelic variation between Ben and PI 41025 within the *Vrn-A1* and *Vrn-B1* loci. The QTL for DTH (*QEet.fcu-2B*) mapped to the short arm of 2B, where the photoperiod sensitivity gene *Ppd-B1* is known to be located in common wheat (Mohler et al. 2004) making it possible that the earliness effects contributed by PI 41025 at *QEet.fcu-2B* are due to allelic variation in that gene. Finally, the DTH QTL on 1B (*QEet.fcu-1B*) mapped near the distal end of the long arm where Lin et al. (2008) detected a QTL for flowering time

in common wheat thus suggesting the possibility that the 1BL QTL detected in our research is the same.

### Spike morphology

QTL affecting SL were identified on 4A, 4B, and 5A. The QTL on 5A was due to the effects of the *Q* gene, which was expected because it is well known that the *Q* allele confers a shorter spike relative to the *q* allele (see Faris et al. 2005 for review). QTL affecting spike length have been previously reported in the vicinity of *QEl.fcu-4A* on chromosome 4A (Borner et al. 2002; Jantasuriyarat et al. 2004) and of *QEl.fcu-4B* on chromosome 4B (Wang et al. 2010) making it possible that the ones identified in this research are the same as those previously reported. It is interesting to note that although the effects for a shortened spike were contributed by Ben for all three QTL, Ben and PI 41025 did not differ significantly in SL and the RIL population exhibited strong transgressive segregation for this trait. This would suggest that PI 41025 harbors genes that contribute to a shortened spike that differ from those of Ben, and that they went undetected. Reasons for this are unclear, but it is possible that PI 41025 contains numerous genes with minor effects on SL. In this case it is worthy to note that QTL with effects for reduced SL contributed by PI 41025 that approached the significance threshold (LOD > 2.5) were detected on chromosomes 2A, 2B, 5A, and 6A, but were not considered because they did not reach the critical LOD threshold of 3.2. It is possible that these represent real QTL with effects contributed by PI 41025. It is also possible that epistatic interactions exist that contributed to the transgressive segregation.

Although SPS relates directly to yield, it also influences spike morphology. The number of SPS and other yield components have been evaluated in many previous studies leading to reports of QTL associated with SPS on most of the wheat chromosomes. Therefore, it is likely that the four QTL associated with SPS in the BP025 population have been reported before. For example, Peng et al. (2003) reported a QTL on 1BS associated with SPS that may be the same as *QSpn.fcu-1B*, and Jantasuriyarat et al. (2004) and Ma et al. (2007) have reported SPS QTL in the same vicinity of *QSpn.fcu-7A*. Many yield and yield-related QTL have been reported on 3B (Borner et al. 2002; Groos et al. 2003; Elouafi and Nachit 2004; Cuthbert et al. 2008; Graziani et al. 2014). However, to our knowledge, QTL associated with SPS have not previously been reported on 7B.

QTL for spike compactness have been reported on nearly all chromosomes as well, and they coincide many times with QTL for SL, but also coincide sometimes with QTL for SPS or both SL and SPS (Wang et al. 2010; Faris et al. 2014b). In this work, we found that the CMP QTL on chromosomes 3B and 7A coincided with SPS QTL,

whereas the 5A CMP QTL coincided with a SL QTL due to the underlying effects of *Q*. The remaining six CMP QTL did not coincide with either SPS or SL QTL, however it is interesting to note that a CMP QTL on 2B was very near the 2B QTL for THR—which was most likely due to the *Tg-B1* gene (see below)—and a CMP QTL on 2A nearly coincided with a 2A THR QTL, which could perhaps represent *Tg-A1*. It may be possible that the *Tg* genes on group 2 chromosomes have slight, albeit significant, pleiotropic effects on spike compactness as well as threshability.

#### Yield-related traits

QTL for yield and yield components have been identified on most, if not all, wheat chromosomes (Marza et al. 2005; Edae et al. 2014), and some potentially coincide with the QTL identified in this research for KPS, GWS, and TKW. For example, Borner et al. (2002) reported QTL for GWS on 1AS, 2BS, and 7AS, and a QTL on 3B for TKW that may correspond to *QGws.fcu-1A*, *QGws.fcu-2B*, *QGws.fcu-7A*, and *QTKw.fcu-3B*, respectively. Chromosome 5B has been reported to harbor QTL for various yield components as well (Gonzalez-Hernandez et al. 2004; Edae et al. 2014), which may indicate the previous identification of QTL associated with KPS and GWS as identified in this research.

The QTL on chromosomes 7A and 5B was associated with both KPS and GWS indicating that the underlying genes contributed by PI 41025 contribute primarily to increase the number of kernels per spike, and this increase leads to an increase in the weight of the spike. On the contrary, the 1A QTL is associated with GWS and TKW, but not KPS, indicating that the gene(s) underlying the 1A QTL and contributed by Ben enhance seed weight but not the number of kernels. The chromosome 2B QTL was associated with all three traits (KPS, GWS, and TKW) and had the largest effects. Therefore, the gene(s) underlying the 2B QTL contributed by Ben has pleiotropic effects on the number of seeds produced per spike and the weight of the seed.

Based on the results of this research it is evident that PI 41025 harbors alleles that confer an increased number of SPS, Ben contains alleles that contribute an increase in seed weight (TKW), and both Ben and PI 41025 harbor alleles that contribute to the number of KPS. It should therefore be possible to combine alleles for increased SPS and KPS from PI 41025 with alleles for increased KPS and TKW from Ben to achieve genotypes with significant yield increases compared to either parent. Indeed, the measure of GWS in the population revealed five RILs with GWS values significantly better than Ben. The best RIL, BP025-114, had a GWS value of 2.35 g/spike, which was 0.75 g/spike more than Ben (a BP025-114 spike is shown in

Fig. 1). It is also interesting to note that, because BP025-50 had significantly higher GWS than Ben but carried alleles for increased GWS at only one of the four detected QTL (*QGws.fcu-7A*), it is likely that other QTL affecting GWS went undetected. It would be interesting to conduct large field-based experiments to evaluate more accurately yield and yield-related traits in this population, because some RILs, perhaps such as BP025-114, may be useful for incorporation into breeding programs to achieve advances in durum wheat yields.

#### Rachis fragility and seed threshability

The pleiotropic effects of the *Q* gene on morphological and domestication traits, including rachis fragility, are well documented (Watkins 1940; Sears 1956; Muramatsu 1963, 1979, 1985, 1986; Kato et al. 1999, 2003; Faris and Gill 2002; Faris et al. 2003, 2005; Simons et al. 2006; Zhang et al. 2011). Here, the profound effects of *Q* on FRA and THR were fully expected, because cultivated emmer wheat is known to possess the *q* allele and durum wheat the *Q* allele. The other QTL associated with FRA was on chromosome 3B. Whereas the effects of *QFra.fcu-3B* were relatively minor in comparison to those of the QTL governed by the *Q* locus (*QFra.fcu-5A*), it is interesting to note that the position of *QFra.fcu-3B* corresponds well with the map position of the brittle rachis gene *Br-B1* (Nalam et al. 2006), which is considered to have undergone mutation along with other brittle rachis genes during the transition of wild emmer wheat (*T. turgidum* ssp. *dicoccoides*) to cultivated emmer wheat (*T. turgidum* ssp. *dicoccum*). It is possible that some residual effects of *Br-B1* remain and yet contribute slightly to rachis fragility in cultivated emmer. But, it is also interesting to note that *QFra.fcu-3B* coincided with QTL for SPS and CMP, suggesting that, whatever the gene responsible for *QFra.fcu-3B*, it likely has pleiotropic effects on all three traits.

As expected, the *Q* gene had the largest impact on THR explaining over 30 % of the variation. The QTL with the second largest impact was *QFt.fcu-2B*, which corresponds well with the position of *Tg-B1* determined by Faris et al. (2014a), who used a segregating mapping population derived from a Langdon durum–*T. turgidum* ssp. *dicoccoides* chromosome 2B disomic chromosome substitution line crossed with Langdon to map *Tg-B1*. In both this work and that of Faris et al. (2014a), the locus governing threshability mapped very near the SSR marker *Xwmc154*, indicating they are likely the same gene. The observance of *Tg-B1* contributed by PI 41025 in this research thus indicates that the mutation in *Tg-B1* giving rise to the free-threshing *tg-B1* allele did not occur during the transition of wild to cultivated emmer, but instead must have happened during the transition of cultivated

emmer to durum or another fully domesticated tetraploid subspecies.

Dvorak et al. (2012) and Faris et al. (2014a) both presented evidence of the existence of factors inhibiting threshability on chromosome 2A in polyploid wheat and speculated that it was most likely the homoeologue of *Tg-B1* and *Tg-D1*, which would be referred to as *Tg-A1*. It is likely that one of the two THR QTL on chromosome 2A in the BP025 population represents *Tg-A1*. The genetic position of *QFt.fcu-2A.1* along chromosome 2A is similar to that of the *Tg-B1* locus (*QFt.fcu-2B*) on chromosome 2B, which might suggest that *QFt.fcu-2A.1* is associated with *Tg-A1*. However, it is interesting to note that the second QTL on chromosome 2A (*QFt.fcu-2A.2*) is associated with CMP (*QCmp.fcu-2A*) as well as THR, and the *Tg-B1* locus on 2B is also associated with both of these traits. Although the *Tg* loci have not previously been reported to be associated with spike compactness, it is possible that they might influence the trait in certain genetic backgrounds. If this was the case, then *QFt.fcu-2A.2* might represent *Tg-A1*. In either case, more work is needed to determine which 2A QTL (*QFt.fcu-2A.1* or *QFt.fcu-2A.2*) is homoeologous to *QFt.fcu-2B* and therefore representative of *Tg-A1*.

#### Insights into wheat domestication events

It has long been known that cultivated emmer harbors the *q* allele because the seed is hulled and the rachis somewhat fragile, however, it had not been determined if cultivated emmer possessed *Tg-B1*, *Tg-A1*, neither gene, or perhaps both genes. The results of the current study indicate that the cultivated emmer accession PI 41025 clearly harbors *Tg-B1*, and suggest that it contains *Tg-A1* as well. Whereas mutations in genes conferring a brittle rachis led to the transition of wild emmer to cultivated emmer wheat, seeds of the latter remained hulled and difficult to thresh due to the likelihood of having the genotype *Tg-A1Tg-A1/Tg-B1Tg-B1/qq*. The evolution of the first free-threshing tetraploid wheat then required mutations at all three loci, i.e., *Tg-A1*, *Tg-B1*, and *q* to give rise to the genotype *tg-A1tg-A1/tg-B1tg-B1/QQ*. The mutations in these three loci likely occurred during the transition from cultivated emmer to the first free-threshing tetraploid. However, other cultivated emmer accessions should be evaluated for threshability QTL to further validate this hypothesis.

With regards to the formation of hexaploid common wheat (*T. aestivum*), it is most probable that a free-threshing tetraploid subspecies with genotype *tg-A1tg-A1/tg-B1tg-B1/QQ* was involved in the amphiploidization event (Matsuoka and Nasuda 2004; Dvorak et al. 2012; Faris 2014; Faris et al. 2014a) with *A. tauschii* in the Middle East. The first hexaploid amphiploid would have been

non-free-threshing due to the acquisition of *Tg-D1* from *A. tauschii* (McFadden and Sears 1946; Kerber and Rowland 1974). The amphiploid must have acquired a mutation in *Tg-D1* very rapidly to give rise to the *tg-D1* allele and fully free-threshing wheat, because true non-free-threshing hexaploid ancestors have not been observed in the archeological record (Nesbitt and Samuel 1996). This fact lends further support to the argument that the tetraploid parent of *T. aestivum* was free-threshing with genotype *tg-A1tg-A1/tg-B1tg-B1/QQ*, because if it was not, then it would have been necessary for the original hexaploid to undergo mutations at multiple loci to become free-threshing, which would be expected to take a significant amount of time. Therefore, it is likely that the original hexaploid wheat plant had the genotype *tg-A1tg-A1/tg-B1tg-B1/Tg-D1Tg-D1/QQ* and underwent a mutation in *Tg-D1* to give rise to the fully domesticated *T. aestivum* ssp. *aestivum* with genotype *tg-A1tg-A1/tg-B1tg-B1/tg-D1tg-D1/QQ*.

#### Conclusions

Although cultivated emmer wheat was instrumental to the Agricultural Revolution of the Neolithic times, which led to the rise of human civilization, today it is the fully domesticated durum wheat that has replaced cultivated emmer as the primary form of tetraploid wheat that is grown worldwide and provides a significant source of calories for humans. Pronounced differences in agronomic, morphological, productivity-associated, and domestication-related traits exist between cultivated emmer and modern domesticated tetraploid wheat, and the current research provides insights regarding the transitional events that led not only to a fully tough rachis and free-threshing seed, but also to alterations in spike architecture and increased seed weight and yield. Although modern durum wheat is superior to cultivated emmer in agronomic performance and productivity, cultivated emmer provides a rich pool of genetic diversity from which to obtain genes for potentially desirable traits such as enhanced yield, quality, disease and pest resistance, and agronomic performance for the improvement of durum wheat varieties.

**Author contributions** JF, QZ, SC, ZZ, and SX performed experiments and collected data. JF and SX analyzed and interpreted the data. JF, SC, and SX wrote the paper.

**Acknowledgments** The authors thank Ms. Breanna Oldenburg and Ms. Megan Stoley for technical assistance. This research was supported by USDA-ARS CRIS project 5442-22000-037-00D.

**Conflict of interest** The authors declare no conflict of interest.

## References

- Abeysekara NS, Friesen TL, Liu ZH, McClean PE, Faris JD (2010) Marker development and saturation mapping of the tan spot PtrToxB sensitivity locus *Tsc2* in hexaploid wheat. *Plant Genome* 3:179–189
- Blake NK, Leheldt BR, Lavin M, Talbert LE (1999) Phylogenetic reconstruction based on low copy DNA sequence data in an allopolyploid: the B genome of wheat. *Genome* 42:351–360
- Blanco A, Bellomo MP, Cenci A, De Giovanni C, D'Ovidio R, Iacono E, Laddomada B, Pagnotta MA, Porceddu E, Sciancalepore A, Simeone R, Tanzarella OA (1998) A genetic linkage map of durum wheat. 97:721–728
- Borner A, Schumann E, Furste A, Coster H, Leithold B, Roder MS, Weber WE (2002) Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (*Triticum aestivum* L.). *Theor Appl Genet* 105:921–936
- Buerstmayr M, Huber K, Heckmann J, Steiner B, Nelson JC, Buerstmayr H (2012) Mapping of QTL for Fusarium head blight resistance and morphological and developmental traits in three backcross populations derived from *Triticum dicoccum* × *Triticum durum*. *Theor Appl Genet* 125:1751–1765
- Cavanagh C, Chao S, Huang BE, Wang S, Stephen S, Forrest K, Kiani S, Sainetnac C, Brown-Guedira G, Akhunova A, See D, Bai G, Pumphrey M, Tomar L, Wong D, Kong S, Reynolds M, da Silva ML, Bockelman H, Talbert L, Anderson J, Dreisigacker S, Baezinger S, Carter A, Korzun V, Morrell P, Dubcovsky J, Morell M, Sorrells M, Hayden M, Akhunov E (2013) Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landrace and cultivars. *Proc Natl Acad Sci USA* 110:8057–8062
- Chalupska D, Lee HY, Faris JD, Evrard A, Chalhoub B, Haselkorn R, Gornicki P (2008) *Acchomoeoloci* and the evolution of the wheat genomes. *Proc Natl Acad Sci USA* 105:9691–9696
- Chu C-G, Chao S, Friesen TL, Faris JD, Zhong S, Xu SS (2010) Identification of novel tan spot resistance QTLs using an SSR-based linkage map of tetraploid wheat. *Mol Breeding* 25:327–338
- Colasuonno P, Mastrangelo AM, Blanco A, Gadaleta A (2013) Description of durum wheat linkage map and comparative sequence analysis of wheat mapped DArT markers with rice and Brachypodium genomes. *BMC Genet* 14:114
- Cuthbert JL, Somers DJ, Brule-Babel AL, Brown PD, Crow GH (2008) Molecular mapping of quantitative trait loci for yield and yield components in spring wheat (*Triticum aestivum* L.). *Theor Appl Genet* 117:595–608
- Dvorak J, Zhang HB (1990) Variation in repeated nucleotide sequences sheds light on the phylogeny of the wheat B and G genomes. *Proc Natl Acad Sci USA* 87:9640–9644
- Dvorak J, di Terlizzi P, Zhang H-B, Resta P (1993) The evolution of polyploid wheats: identification of the A genome donor species. *Genome* 36:21–31
- Dvorak J, Deal KR, Luo M-C, You FM, von Borstel K, Dehghani H (2012) The origin of spelt and free-threshing hexaploid wheat. *J Hered* 103:426–441
- Edae EA, Byrne PF, Haley SD, Lopes MS, Reynolds MP (2014) Genome-wide association mapping of yield and yield components of spring wheat under contrasting moisture regimes. *Theor Appl Genet* 127:791–807
- Elias EM, Miller JD (1998) Registration of 'Ben'durum wheat. *Crop Sci* 38:895
- Elouafi I, Nachit MM (2004) A genetic linkage map of the durum × *Triticum dicoccoides* backcross population based on SSRs and AFLP markers, and QTL analysis for milling traits. *Theor Appl Genet* 108:401–413
- Faris JD (2014) Wheat domestication: Key to agricultural revolutions past and future. In: Tuberosa R, Graner A, Frison E (ed) *Genomics of plant genetic resources*, vol. 1 Managing, sequencing and mining genetic resources. Springer, pp 439–464
- Faris JD, Gill BS (2002) Genomic targeting and high-resolution mapping of the domestication gene *Q* in wheat. *Genome* 45:706–718
- Faris JD, Fellers JP, Brooks SA, Gill BS (2003) A bacterial artificial chromosome contig spanning the major domestication locus *Q* in wheat and identification of a candidate gene. *Genetics* 164:311–321
- Faris JD, Simons KJ, Zhang Z, Gill BS (2005) The wheat super domestication gene *Q*. *Wheat Info Serv* 100:129–148
- Faris JD, Zhang Z, Chao S (2014a) Map-based analysis of the tenacious glume gene *Tg-B1* of wild emmer and its role in wheat domestication. *Gene* 542:198–208
- Faris JD, Zhang Z, Garvin DF, Xu SS (2014b) Molecular and comparative mapping of genes governing spike compactness from wild emmer wheat. *Mol Genet Genomics* 289:641–651
- Gonzalez-Hernandez JL, Elias EM, Kianian SF (2004) Mapping genes for grain protein concentration and grain yield on chromosome 5B of *Triticum turgidum* (L.) var. *dicoccoides*. *Euphytica* 139:217–225
- Graziani M, Maccaferri M, Royo C, Salvatorelli F, Tuberosa R (2014) QTL dissection of yield components and morpho-physiological traits in a durum wheat elite population tested in contrasting thermo-pluviometric conditions. *Crop Pasture Sci* 65:80–95
- Groos C, Robert N, Bervas E, Charmet G (2003) Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. *Theor Appl Genet* 106:1032–1040
- Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, Gornicki P (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. *Proc Natl Acad Sci USA* 99:8133–8138
- Jantasuriyarat C, Vales MI, Watson CJW, Riera-Lizarazu O (2004) Identification and mapping of genetic loci affecting free-threshing habit and spike compactness in wheat (*Triticum aestivum* L.). *Theor Appl Genet* 108:261–273
- Joehanes R, Nelson JC (2008) QGene 4.0, an extensible Java QTL-analysis platform. *Bioinformatics* 24:2788–2789
- Kamran A, Iqbal M, Spaner D (2014) Flowering time in wheat (*Triticum aestivum* L.): a key factor for global adaptability. *Euphytica* 197:1–26
- Kato K, Miura H, Sawada S (1999) QTL mapping of genes controlling ear emergence time and plant height on chromosome 5A of wheat. *Theor Appl Genet* 98:472–476
- Kato K, Sonokawa R, Miura H, Sawada S (2003) Dwarfing effect associated with the threshability gene *Q* on wheat chromosome 5A. *Plant Breed* 122:489–492
- Kerber ER, Dyck PL (1969) Inheritance in hexaploid wheat of leaf rust resistance and other characters derived from *Aegilops squarrosa*. *Can J Genet Cytol* 11:639–647
- Kerber ER, Rowland GG (1974) Origin of the free threshing character in hexaploid wheat. *Can J Genet Cytol* 16:145–154
- Kihara H (1944) Discovery of the DD-analyzer, one of the ancestors of *Triticum vulgare*. *Agric Hortic (Tokyo)* 19:13–14
- Kosambi DD (1944) The estimation of map distances from recombination values. *Ann Eugen* 12:172–175
- Kumar S, Gill BS, Faris JD (2007) Identification and characterization of segregation distortion loci along chromosome 5B in tetraploid wheat. *Mol Genet Genomics* 278:187–196
- Li WL, Gill BS (2006) Multiple pathways for seed shattering in the grasses. *Funct Integr Genomics* 6:300–309
- Lin F, Xue SL, Tian DG, Li CJ, Cao Y, Zhang ZZ, Zhang CQ, Ma ZQ (2008) Mapping chromosomal regions affecting flowering time in a spring wheat RIL population. *Euphytica* 164:769–777

- Lorieux M (2012) MapDisto: fast and efficient computation of genetic linkage maps. *Mol Breeding* 30:1231–1235
- Ma Z, Zhao D, Zhang C, Zhang Z, Xue S, Lin F, Kong Z, Tian D, Luo Q (2007) Molecular genetic analysis of five spike-related traits in wheat using RIL and immortalized F<sub>2</sub> populations. *Mol Genet Genomics* 277:31–42
- Mantovani P, Maccaferri M, Sanguineti MC, Tuberosa R, Catizone I, Wenzl P, Thomson B, Carling J, Huttner E, DeAmbrogio E, Kilian A (2008) An integrated DAR-T-SSR linkage map of durum wheat. *Mol Breed* 22:629–648
- Marone D, Laido G, Gadaleta A, Colusuonno P, Ficco DBM, Giancaspro A, Giove S, Panio G, Russo MA, De Vita P, Cattivelli L, Papa R, Blanco A, Mastrangelo AM (2012) A high-density consensus map of A and B wheat genomes. *Theor Appl Genet* 125:1619–1638
- Marza F, Bai G-H, Carver BF, Zhou W-C (2005) Quantitative trait loci for yield and related traits in the wheat population Ning7840 × Clark. *Theor Appl Genet* 112:688–698
- Matsuoka Y (2011) Evolution of polyploid *Triticum* wheats under cultivation: the role of domestication, natural hybridization and allopolyploid speciation in their diversification. *Plant Cell Physiol* 52:750–764
- Matsuoka Y, Nasuda S (2004) Durum wheat as a candidate for the unknown female progenitor of bread wheat: an empirical study with a highly fertile F<sub>1</sub> hybrid with *Aegilopstauschii* Coss. *Theor Appl Genet* 109:1710–1717
- McFadden ES, Sears ER (1946) The origin of *Triticum spelta* and its free-threshing hexaploid relatives. *J Hered* 37(81–89):107–116
- Mohler V, Lukman R, Ortiz-Islas S, William M, Worland AJ, Van Beem J, Wenzel G (2004) Genetic and physical mapping of photoperiod insensitive gene *Ppd-B1* in common wheat. *Euphytica* 138:33–40
- Muramatsu M (1963) Dosage effect of the *spelta* gene *q* of hexaploid wheat. *Genetics* 48:469–482
- Muramatsu M (1979) Presence of the *vulgare* gene, *Q*, in a dense-spike variety of *Triticum dicoccum* Schübl. Report of the Plant Germ-Plasm Institute, Kyoto Univ 4:39–41
- Muramatsu M (1985) Spike type in two cultivars of *Triticum dicoccum* with the *spelta* gene *q* compared with the *Q*-bearing variety *liguliforme*. *Jpn J Breed* 35:255–267
- Muramatsu M (1986) The *vulgare* super gene, *Q*: its universality in durum wheat and its phenotypic effects in tetraploid and hexaploid wheats. *Can J Genet Cytol* 28:30–41
- Nalam VJ, Vales MI, Watson CJW, Kianian SF, Riera-Lizarazu O (2006) Map-based analysis of genes affecting the brittle rachis character in tetraploid wheat (*Triticum turgidum* L.). *Theor Appl Genet* 112:373–381
- Nalam VJ, Vales MI, Watson CJW, Johnson EB, Riera-Lizarazu O (2007) Map-based analysis of genetic loci on chromosome 2D that affect glume tenacity and threshability components of free-threshing habit in common wheat (*Triticum aestivum* L.). *Theor Appl Genet* 116:135–145
- Nesbitt M, Samuel D (1996) From staple crop to extinction? The archaeology and history of hulled wheats. In: Padulosi S, Hammer K, Heller J (eds) Hulled wheats, Promoting the conservation and use of underutilized and neglected crops 4: Proceedings of the first international workshop on hulled wheats. Castelveccchio Pascoli, Tuscany, Italy, pp 41–100
- Peleg Z, Fahima T, Korol AB, Abbo S, Saranga Y (2011) Genetic analysis of wheat domestication and evolution under domestication. *J Exp Botany* 62:5051–5061
- Peng JH, Ronin Y, Fahima T, Röder MS, Li YC, Nevo E, Korol A (2003) Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat. *Proc Natl Acad Sci USA* 100:2489–2494
- Pestsova E, Ganal MW, Röder MS (2000) Isolation and mapping of microsatellite markers specific for the D genome of bread wheat. *Genome* 43:689–697
- Röder MS, Korzun V, Wendehake K, Plaschke J, Tixier M-H, Leroy P, Ganal MW (1998) A microsatellite map of wheat. *Genetics* 149:2007–2023
- Salse J, Chague V, Bolot S, Magdelenat G, Huneau C, Pont C, Belcram H, Couloux A, Gardais S, Evrard A, Segurens B, Charles M, Ravel C, Samain S, Charret G, Boudet N, Chalhou B (2008) New insights into the origin of the B genome of hexaploid wheat: evolutionary relationships at the *SPA* genomic region with the S genome of the diploid relative *Aegilops speltoides*. *BMC Genom* 9:555
- SAS Institute (2011) SAS/STAT 9.3 User's Guide. SAS Institute Inc., Cary, NC
- Sears ER (1956) The systematics, cytology and genetics of wheat. *Handb Pflanzenzücht*, vol. 2, 2nd edn, pp 164–187
- Simonetti MC, Bellomo MP, Laghetti G, Perrino P, Simeone R, Blanco A (1999) Quantitative trait loci influencing free-threshing habit in tetraploid wheats. *Genet Res Crop Evol* 46:267–271
- Simons KJ, Fellers JP, Trick HN, Zhang Z, Tai Y-S, Gill BS, Faris JD (2006) Molecular characterization of the major wheat domestication gene *Q*. *Genetics* 172:547–555
- Somers DJ, Isaac P, Edwards K (2004) A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theor Appl Genet* 109:1105–1114
- Song QJ, Shi JR, Singh S, Fickus EW, Costa JM, Lewis J, Gill BS, Ward R, Cregan PB (2005) Development and mapping of microsatellite (SSR) markers in wheat. *Theor Appl Genet* 110:550–560
- Sood S, Kuraparthi V, Bai GH, Gill BS (2009) The major threshability genes soft glume (*sog*) and tenacious glume (*Tg*), of diploid and polyploid wheat, trace their origin to independent mutations at non-orthologous loci. *Theor Appl Genet* 119:341–351
- Sourdille P, Cadalen T, Guyomarc'h H, Snape JW, Perretant MR, Charret G, Boeuf C, Bernard S, Bernard M (2003) An update of the Courtot × Chinese Spring intervarietal molecular marker linkage map for the QTL detection of agronomic traits in wheat. *Theor Appl Genet* 106:530–538
- Sourdille P, Singh S, Cadalen T, Brown-Guedira GL, Gay G, Qi L, Gill BS, Dufour P, Murigneux A, Bernard M (2004) Microsatellite-based deletion bin system for the establishment of genetic-physical map relationships in wheat (*Triticum aestivum* L.). *Funct Integr Genomics* 4:12–25
- Thanh PT, Vladutu CI, Kianian SF, Thanh PT, Ishii T, Nitta M, Nasuda S, Mori N (2013) Molecular genetic analysis of domestication traits in emmer wheat. I: map construction and QTL analysis using an F<sub>2</sub> population. *Biotechnology* 27:3627–3637
- Trebbi D, Maccaferri M, de Heer P, Sorensen A, Giuliani S, Salvi S, Sanguineti MC, Massi A, van der Vossen AG, Tuberosa R (2011) High-throughput SNP discovery and genotyping in durum wheat (*Triticum durum* Desf.). *Theor Appl Genet* 123:555–569
- Tzarfati R, Saranga Y, Barak V, Gopher A, Korol AB, Abbo S (2013) Threshing efficiency as an incentive for rapid domestication of emmer wheat. *Ann Bot* 112:829–837
- vanPoeck RMP, Maccaferri M, Tang J, Truong HT, Janssen A, van Orsouw NJ, Salvi S, Sanguineti MC, Tuberosa R, van der Vossen EAG (2013) Sequence-based SNP genotyping in durum wheat. *Plant Biotechnol J* 11:809–817
- Wang J, Liao X, Li Y, Zhou R, Yang X, Gao L, Jia J (2010) Fine mapping a domestication-related QTL for spike-related traits in a synthetic wheat. *Genome* 53:798–804
- Watanabe N, Ikebata N (2000) The effects of homoeologous group 3 chromosomes on grain colour dependent seed dormancy and brittle rachis in tetraploid wheat. *Euphytica* 115:215–220

- Watkins AE (1940) The inheritance of glume shape in *Triticum*. *J Genet* 39:249–264
- Zhang ZC, Belcram H, Gornicki P, Charles M, Just J, Huneau C, Magdelenat G, Couloux A, Samain S, Gill BS, Rasmussen JB, Barbe V, Faris JD, Chalhoub B (2011) Duplication and partitioning in evolution and function of homoeologous *Q* loci governing domestication characters in polyploid wheat. *Proc Natl Acad Sci USA* 108:18737–18742