Theor Appl Genet (2015) 128:745–755
DOI 10.1007/s00122-015-2469-1

ORIGINAL PAPER

Fine mapping of shattering locus *Br2* reveals a putative chromosomal inversion polymorphism between the two lineages of *Aegilops tauschii*

Zhengzhi Zhang · Huilan Zhu · Bikram S. Gill · Wanlong Li

Received: 4 August 2014 / Accepted: 26 January 2015 / Published online: 6 February 2015

Abstract

Key message **This work laid the foundation for cloning of shattering gene *Br2* and provided first line of evidence that two major *Aegilops tauschii* lineages are differentiated by an inversion polymorphism.**

Abstract Chromosome inversions often accompany population differentiation and capture local adaptation during speciation. *Aegilops tauschii*, the D-genome donor species of hexaploid wheat, consists of two genetically isolated lineages, L1 and L2, but little is known about the genetic mechanisms underlying the population differentiation in this diploid species. During fine mapping of the shattering gene *Br2* using a large F_2 population derived from a cross between TA1604 (an L1 accession) and AL8/78 (an L2 accession), we found contrasting patterns of crossover distribution in the *Br2* interval and neighboring regions despite the high local gene synteny with *Brachypodium distachyon* and rice. *Br*2 was localized in a 0.08-cM interval, and 13 marker loci formed a block, where single-crossovers were completely suppressed, but double-crossovers were enriched with a recombination rate of ~11 cM/Mb. In contrast, in a neighboring region no double-crossover was recovered, but single-crossover rate reached 24 cM/Mb, which is much higher than the genome-wide average. This result suggests a putative inversion polymorphism between the parental lines in the *Br2* region. Genotyping using the markers from the *Br2* region divided a collection of 55 randomly sampled *A. tauschii* accessions into two major groups, and they are largely genetically isolated. The two groups correspond to the L1 and L2 lineages based on their geographic distribution patterns. This provides first evidence that inversions may underlie the evolution of *A. tauschii* lineages. The presence of inter-lineage inversions may complicate map-based cloning in *A. tauschii* and transfer of useful traits to wheat.

Communicated by M. J. Sillanpaa.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-015-2469-1) contains supplementary material, which is available to authorized users.

Z. Zhang · H. Zhu · W. Li
Department of Biology and Microbiology, South Dakota State University, Brookings, SD 57007, USA

B. S. Gill
Department of Plant Pathology, Wheat Genetics Resource Center, Kansas State University, Manhattan, KS 66506, USA

W. Li (✉)
Department of Plant Science, South Dakota State University, Brookings, SD 57007, USA
e-mail: wanlong.li@sdstate.edu

Abbreviations

asl	Above sea level
BP	Before present
BAC	Bacterial artificial chromosome
CAPS	Cleaved amplified polymorphic sequence
cM	Centimorgan
dCAPS	Derived CAPS
indel	Small insertion/deletion
MTP	Minimal tiling path
RFLP	Restriction fragment length polymorphism
SNP	Single nucleotide polymorphism
UPGMA	Unweighted pair group method with arithmetic means
UTR	Untranslated region

Introduction

Modern crops were domesticated from their wild ancestors by artificial selection for traits suited to mass cultivation.

Among many characters distinguishing cultivated forms from their wild progenitors, loss of seed dispersal, ease of seed threshing, increase of seed size, synchronization of growth and flowering and changes in plant architecture are the most striking domestication targets and comprise the "domestication syndrome" (Gepts 2004; Harlan 1992). Reduction in grain shattering is critical to an effective harvest and has been viewed as the hallmark of cereal domestication. Several shattering genes have been identified and characterized in cereal crops. These include *SH4* (Li et al. 2006; Lin et al. 2007), *qSH1* (Konishi et al. 2006) and *SHAT1* (Zhou et al. 2012) of rice, and *Shattering1* of sorghum and maize (Lin et al. 2012).

Wheat (*Triticum*) was among the first wave of crops domesticated by humans in the Mideast of Asia: The diploid einkorn (*T. monococcum* subsp. *monococcum*, genome A^mA^m) was domesticated 11,000 years before present (BP), tetraploid emmer (*T. turgidum* subsp. *dicoccum*, genome AABB) domesticated 10,000 years BP, and hexaploid or common wheat (*T. aestivum*, genome AABBDD) originated in cultivation approximately 8,000 years BP (Salamini et al. 2002). In the *Triticum* species, spike shatters by spikelet disarticulation at every joint of the rachis at maturation. Spikelet disarticulation is divided into barrel-type and wedge-type based on the disarticulation occurring below or above the joint and the resulted shape of dissemination unit. The wedge-type disarticulation is found in diploid (*T. monococcum* subsp. *aegilopoides* and *T. urartu*, genome AA) and tetraploid wild wheat (*T. turgidum* subsp. *dicoccoides*, genome AABB) and hexaploid semi-wild Tibet wheat (*T aestivum* subsp. *tibetanum,* genome AABBDD) (reviewed in Li and Gill 2006). In polyploid wheat lineages, the wedge-type disarticulation is controlled by an orthologous locus *Br1* on the short arm of group-3 chromosomes (Li and Gill 2006; Nalam et al. 2006; Watanabe et al. 2006). The barrel-type disarticulation is only found in hexaploid spelt wheat (*T. aestivum* subsp. *spelta*, genome AABBDD), suggesting that the causal gene was derived from *Aegilops tauschii* Coss. (genome DD), the D-genome donor species (reviewed in Li and Gill 2006). Although the gene underlying the barrel-type disarticulation in spelt wheat remains to be mapped, we found that the *Br2* locus on the long arm of chromosome 3D (3DL) is responsible for the B-type shattering in *A. tauschii*, indicating that these two disarticulation types are controlled by different genetic pathways (Li and Gill 2006).

Shattering is crucial for wild species to survive in nature so that nonshattering mutations are rare. To date in *A. tauschii*, only one accession (TA1604) has been found having a tough rachis, which was collected in Afghanistan. In contrast, *A. tauschii* showed much higher level of variation in other traits such as disease resistance (Yildirim et al. 1995), insect resistance (Weng et al. 2005), glaucousness (Dudnikov 2011), growth habits (Yildirim et al. 1995), and spike morphology (Kimber and Feldman 1987). Based on spike morphology, *A. tauschii* was divided into two subspecies: subsp. *strangulata* and subsp. *tauschii* (Eig 1929; Hammer 1980). Subsp. *strangulata* has squared spikelets with wide and short glumes, whereas subsp. *tasuchii* has elongated spikelets (Kimber and Feldman 1987). Subsp. *tauschii* contains three variants, i.e., *anathera*, *meyeri*, and *typica*. Geographically, while subsp. *strangulata* is confined to Caucasus and southeast Caspian Sea coastal area, distribution of subsp. *tauschii* stretches from central China to west Turkey (Wang et al. 2013). Marker analysis of the diversity in *A. tauschii* population revealed discrepancy between the genetic categories and the botanical categories: Var. *meyeri* was genetically closer to subsp. *strangulata* than to subsp. *tauschii* (Lubbers et al. 1991). Subsequent analyses at greater scale corroborated this conclusion, and *A. tauschii* has been subdivided into two evolutionary lineages, i.e., L1 and L2 (Dvorak et al. 1998; Mizuno et al. 2010; Sohail et al. 2012; Wang et al. 2013). While L1 consists of part of the subsp. *tauschii*, L2 contains the whole of subsp. *strangulata* and the remaining part of subsp. *tauschii*. The paucity of intermediate types suggests that L1 and L2 are virtually isolated reproductively in nature (Wang et al. 2013). L1 and L2 accessions, however, cross readily, and F_1 hybrids are fully fertile. So what inhibits introgression between these two lineages? Is this due to their adaption to different habitats or chromosomal rearrangements that suppress recombination in the inter-lineage hybrids, or both? Answers to these questions are important for understanding of the speciation of wheat and its relatives and for effective use of these wild relatives for wheat improvement.

We are isolating the shattering gene *Br2* in *A. tauschii* using a map-based cloning approach. In the present study, we developed a high-density linkage map for the *Br2* region, aligned it with genomic sequences of model grasses rice and *Brachypodium distachyon* and the bacterial artificial chromosome (BAC) contigs of *A. tauschii*, and detected a putative paracentric inversion in the *Br2* block. We subsequently genotyped 55 additional accessions of *A. tauschii* using 31 markers in the *Br2* region, which confirmed the differentiation of *A. tauschii* into two lineages and detected local genetic isolation. Here we report the results and implications on the evolution of the D-genome, map-based cloning of the D-genome genes and impact on transfer of useful traits from *A. tauschii* into wheat.

Materials and methods

Plant materials

The shattering *A. tauschii* accession AL8/78 and the nonshattering accession TA1604 were crossed manually for

generating an F_2 mapping population. TA1604 belongs to the L1 and AL8/78 to the L2 lineage. AL8/78 was supplied by Dr. Jan Dvorak (University of California, Davis, CA, USA), and remaining accessions are maintained by the Wheat Genetics Resource Center at Kansas State University. The Accession numbers and collection sites are listed in Table S1. A large F_2 population was developed from a separate cross between TA1604 and AL8/78 and planted in 12 × 6 Rootrainer (Beaver Plastics, Acheson, AB, Canada), from which recombinants were transplanted into 4″ square pots and allowed to grow to maturity. All plant materials were grown in a greenhouse at South Dakota State University, in which temperature was 22 °C during the day (16 h) and 17 °C at night (8 h). Leaf tissue was collected for DNA isolation as described by Li et al. (2008), and polymerase chain reaction (PCR) and gel electrophoresis were conducted following the procedure described by Li et al. (2013).

Marker development

Rice coding sequences were retrieved from the Rice Genome Annotation Project Database (http://rice.plantbiology.msu.edu) and used as queries for searching the Wheat Gene Index database (http://compbio.dfci.harvard.edu/tgi/). The best-hit sequences were retrieved and used to search the Rice Genome Annotation Database (http://rice.plantbiology.msu.edu) to assure that they hit back to the original rice genes. Verified wheat EST sequences were used for designing primers by Primer3 (http://primer3.wi.mit.edu) (Koressaar and Remm 2007) with estimated PCR product size ranging from 500 bp to 1500 bp. The unique amplicons from TA1604 and AL8/78 were purified and sequenced. The sequences were trimmed, assembled, and aligned with BioEdit (http://www.mbio.ncsu.edu/bioedit/bioedit.html) (Hall 1999) to identify single nucleotide polymorphisms (SNPs) and small insertion/deletion (indel) polymorphisms between TA1604 and AL8/78. Cleaved amplified polymorphic sequence (CAPS) markers were designed with SNP2CAPS (http://pgrc.ipk-gatersleben.de/snp2caps/) (Thiel et al. 2004), and derived CAPS (dCAPS) marker was developed with dCAPS Finder 2.0 (http://helix.wustl.edu/dcaps/dcaps.html) (Neff et al. 1998) to target the SNPs. The CAPS and dCAPS primers were used for PCR amplification of TA1604 and AL8/78, and the amplicons were digested with a specific restriction enzyme. The digestions were separated in 2.0 % agarose gel or 6.0 % polyacrylamide gel electrophoresis to confirm the expected polymorphisms. The polymorphic markers were then used for genotyping the segregation populations and a collection of *A. tauschii* accessions (Table S1). The primer sequences and restriction enzymes used are listed in Table S2.

Molecular mapping

CAPS or dCAPS markers were initially used to genotype a population of 93 F_2 individuals, *Br2* genotypes of which were deduced from the phenotypes of their F_3 families (Li and Gill 2006). Two markers flanking the *Br2* locus were chosen for screening the recombinants in a newly constructed large F_2 population. Recombinants were sorted out, phenotyped, and genotyped using additional markers that were mapped to the *Br2* interval. Linkage map was constructed by Mapmaker/Exp V3.0 (Lincoln and Lander 1992). The Kosambi (1944) mapping function was used for converting the recombination frequencies into genetic distances in terms of centimorgan (cM). The deviation of genotypes at marker loci from the 1:2:1 ratio was examined by Chi-square test using the Microsoft Excel function.

For alignment of the genetic map of the *Br2* region with the physical maps of the model grass genomes, the homologs were retrieved from the rice and *B. distachyon* genomes in the Rice Genome Annotation Database (http://rice.plantbiology.msu.edu/) and Phytozome database (http://www.phytozome.net), respectively, and used for constructing the physical maps based on their base pair positions. For physical mapping of the *Br2* region, the marker sequences were used as queries to search the D-genome marker database (http://probes.pw.usda.gov/WheatDMarker/) using the BLASTN algorithm. The remaining markers were further mapped by PCR using the BACs as templates. The AL8/78 BACs were obtained from University of California (Davis, CA, USA). The BAC contig size (kb) was estimated from the number of consensus fingerprinting bands by multiplying coefficient of 1.5 (Luo et al. 2013).

Phylogenetic analysis

CAPS or dCAPS markers were scored as 1 for presence or 0 for absence of the specific allele for each accession of the *A. tauschii* collection. Genetic distance matrix among the accessions was calculated as Jaccard's coefficient, and the distance coefficients were used to construct UPGMA (unweighted pair group method with arithmetic means) dendrogram by MEGA 5.0 (http://www.megasoftware.net) (Tamura et al. 2011). The reliability and goodness of dendrogram were tested through bootstrapping based on 100 samples.

Results

Targeted development of the CAPS and dCAPS markers

Previously, *Br2* was mapped to an interval delimited by restriction fragment length polymorphism (RFLP) markers

Xmwg2013 and *Xpsr170* on chromosome arm 3DL (Li and Gill 2006). BLAST search of rice genome database with the sequences of these RFLP probes showed that *Xmwg2013* and *Xpsr170* were homologs of rice loci *Os01g63980* and *Os01g67860*, respectively. The coding sequences of the rice genes from *Os01g64650* through *Os01g68324* were selected to search their homologs in wheat, and appropriate primers were designed. A total of 46 DNA fragments were sequenced from the parental lines. From 39,024-bp sequences obtained, we discovered 160 SNPs between TA1604 and AL8/78 with an average of 4.1 SNPs/kb and 19 indels with a frequency of ~ 0.5 indels/kb. This SNP frequency is 3.6-fold higher than the gene-space-wide average, 1.14 SNPs/kb, between AL8/78 and AS75 (You et al. 2011). Targeting the SNPs and indels, we developed 90 pairs of primers for CAPS and dCAPS assays, and 35 pairs successfully detected 37 polymorphisms between AL8/78 and TA1604 (Table S2).

As marker development continued, we determined the order of these CAPS and dCAPS marker loci in the *Br2* region using 93 F_2 individuals, *Br2* genotypes of which were determined by the previous study (Li and Gill 2006). We mapped 18 markers to the *Br2* region, which spanned 11.3 cM (Fig. 1). Overall, this chromosome interval showed good synteny to a 1.81-Mb interval from *Bradi2g56020* to *Bradi2g58370* on chromosome 2 of *B. distachyon* and to a 2.18-Mb interval from *Os01g64650* to *Os01g68320* on chromosome 1 of rice. Gene distributions are highly conserved among these three species except for the two inversions, which occurred in *A. tauschii* lineage because rice and *B. distachyon* are collinear in these regions. Distal to the *Br2* locus, a large inversion was detected in *A. tauschii* corresponding to the segment between *Bradi2g57220* and *Bradi2g58040* in *B. distachyon*, and the portion between *Os01g66300* and *Os01g67850* in rice. Proximal to the *Br2* locus, a relatively small inversion was found between two loci homologous to *Bradi2g56020* and *Bradi2g56070* in *B. distachyon,* and *Os01g64650* and *Os01g64700* in rice (Fig. 1). The relative gene density and intergenic spaces showed good correspondence between *B. distachyon* and rice, and they are roughly proportional to the genetic distances between mapped markers in *A. tauschii*.

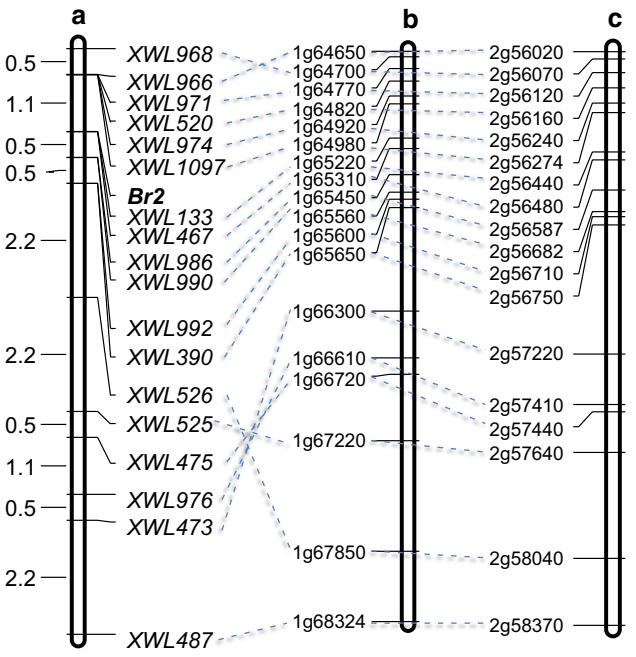


Fig. 1 Frame mapping of the *Br2* locus with newly developed *markers* and alignment with the homologs of rice and *B. distachyon.* **a** A linkage map of chromosome 3DL of *A. tauschii* constructed from 93 F_2 individuals with the *Br2* genotypes deduced from the F3 families. The *markers* are listed on the right side of map, and genetic distances (cM) between the marker loci are indicated on the *left* side of the map. The *Br2* locus is indicated in *bold*. The *top* of the map is toward the centromere, and the *bottom* is toward the telomere. **b** The physical map of rice chromosome 1 and (**c**) the physical map of *B. distachyon* chromosome 2. The *dashed lines* link the shared homologs in the *A. tauschii*, rice, and *B. distachyon* genomes

Fine mapping of *Br2*

Flanking markers WL968 and WL992, 1.6 cM proximal and 1.0 cM distal to the *Br2* locus, respectively, were selected for screening a population of 3,421 F_2 individuals (6,842 gametes). Chi-square test detected segregation distortion at the marker loci *XWL968* ($P = 8.16481E-42$) and *XWL992* ($P = 8.41824E-39$), and the gametes carrying the TA1604 allele were preferentially transmitted to the progeny at twice the frequency of those carrying the AL8/78 allele. Based on the *P* values, a gene causing segregation distortion is proximal to *XWL992*. A total of 104 recombinants were recovered from this screen. These recombinants were first genotyped using six of the 10 markers that were mapped to this interval in Fig. 1 and then by 19 newly developed markers. When maker loci *XWL1373* and *XWL1375* were mapped distal to the *Br2* locus at 0.12 cM, we focused our effort on increasing marker density in the *Br2* region. The fine-scale genetic map spans a 0.4-cM interval, from *XWL968* to *XWL1375*. Marker loci *XWL1445* and *XWL3041* further delimited the *Br2* locus to a 0.08-cM interval, and a block of nine marker loci, from *XWL3037* through *XWL1557,* co-segregated with the *Br2* locus (Fig. 2c). The order of the markers in this block is determined based on their location in BAC contig ctg3280 (Fig. 3). On the fine scale, a high colinearity was maintained among *A. tauschii*, *B. distachyon*, and rice except the marker locus *XWL1443*, which is transposed to a distal location close to *XWL1359* in *A tauschii* (Figs. 2c, 3). The sequence length corresponding to this interval was ~131 kb in *B. distachyon* (Fig. 2a) and ~145 kb in rice (Fig. 2b).

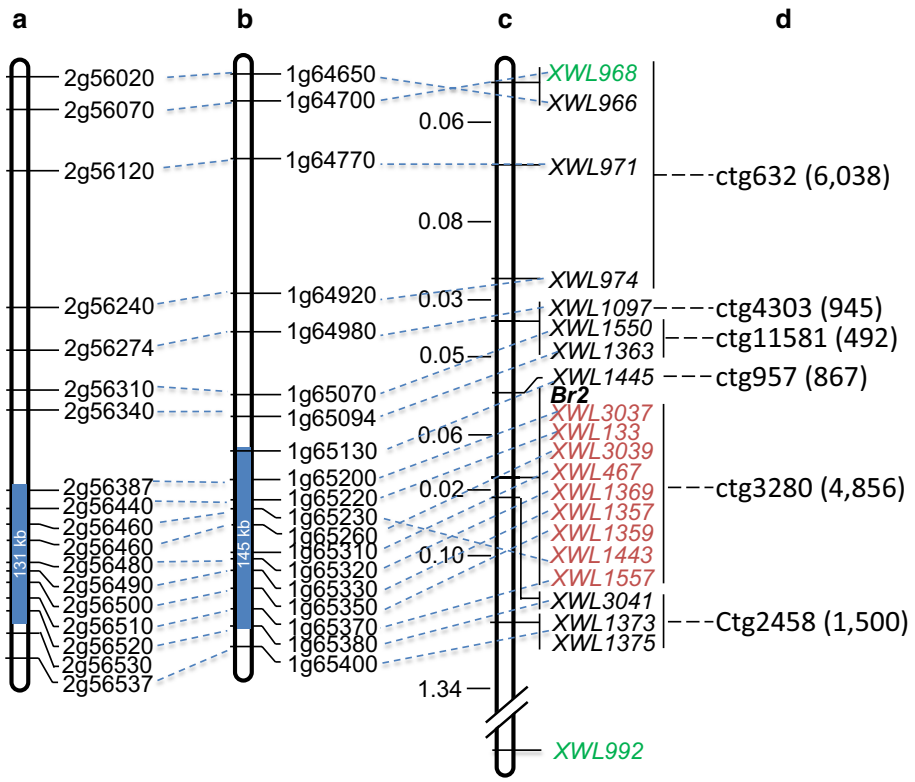


Fig. 2 Fine mapping of the *Br2* locus and alignment with the homologs of rice and *B. distachyon*. **a** The physical map of rice chromosome 1 and (**b**) the physical map of *B. distachyon* chromosome 2. The segments corresponding to the *Br2* region in *A. tauschii* are marked in *blue* color, and their sizes are indicated. **c** A linkage map constructed from 3,421 F_2 individuals. The *markers* are listed on the right side of map, and genetic distances (cM) between the marker loci are indicated on the *left* side of the map. The markers used for screening the whole population are labeled in *green* color, and the marker loci that co-segregated with the *Br2* locus are in orange color. The top of the map is toward the centromere, and the bottom is toward the telomere. The *dashed lines* link the homologs in the *A. tauschii*, rice, and *B. distachyon* genomes. **d** Anchoring BAC contigs to the linkage map. The *dash lines* connect the markers and the BAC contigs containing the corresponding marker sequences, and the *numbers in parentheses* indicate the sizes (kb), which are derived from multiplication of consensus fingerprint band number by a conversion factor of 1.5 kb/consensus band (Luo et al. 2013) (color figure online)

The linkage map of the *Br2* region was anchored to BAC contigs ctg632, ctg4303, ctg11581, ctg3280, and ctg2458 based on search of the D-genome marker database using the marker sequences as queries (Fig. 2d). BAC contig ctg957 was anchored by PCR assay. The block of markers that co-segregated with the *Br2* locus is included in BAC contig ctg3280, the minimal tiling path (MTP) of which contains 47 BACs and is 4,855.5 kb long. Mapping of the CAPS and dCAPS markers to MTP of ctg3280 resolved the order of markers in the co-segregating block and revealed the *XWL1443* transposition event in *A. tauschii* (Figs. 2c, 3). While relatively high level of colinearity is maintained between *A. tauschii* and the model grasses, physical mapping of the collinear genes suggested local genomic expansion in *A. tauschii*. WL1557 and WL3041 correspond to two neighboring genes 15,572 bp apart in the rice genome (*Os01g65370* and *Os01g65380*) and 12,315 bp apart in the *B. distachyon* genome (*Bradi2g56520* and *Bradi2g56530*). Marker WL1557 is located in the middle of the BAC contig ctg3280, but WL3041 is located in BAC contig ctg2458. If gap between the two contigs is ignored, these two genes are ~2 Mb apart *in A. tauschii*. This could be either due to the amplification of transposable elements or insertion of other chromosome blocks by transposition in *A. tauschii* after divergence from *B. distachyon*. Search of the draft genome sequences of *A. tauschii* located marker WL1445 in the genomic scaffold 191134 (GenBank accession KD689925), from which marker WL3373 was developed, and WL3373 in turn identified a positive BAC at the end of the contig ctg957 (Fig. 3).

Distribution of crossovers in the *Br2* region

To measure the genetic-to-physical distance ratio in the *Br2* region, we searched the draft genomic sequences of *A. tauschii* (Jia et al. 2013) and found that WL467 and WL1369 are located in GenBank accession KD511786, and WL1373 and WL3041 are in GenBank accession

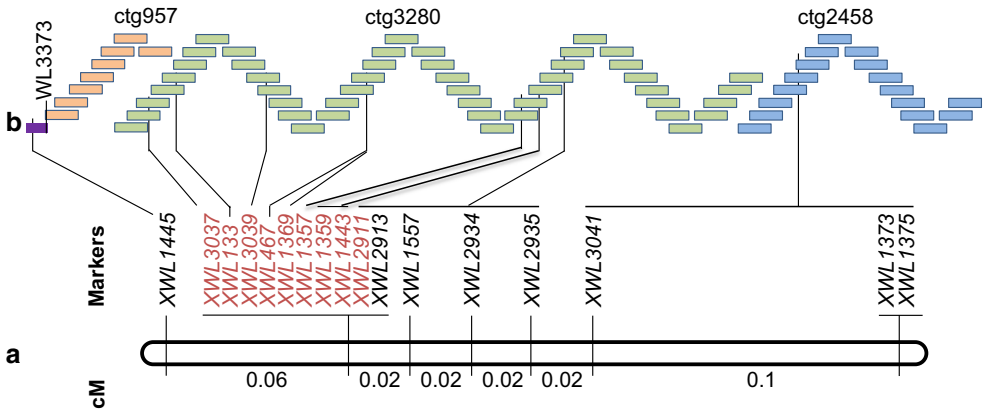


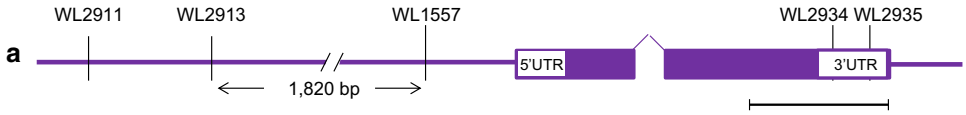
Fig. 3 Finer-scale mapping of the *Br2* locus and anchoring of BAC contigs. **a** Dissection of the *Br2* region by double-crossovers. Mapping of four newly developed markers, WL2911, WL2913, WL2934 and WL2935, detected two double-crossovers (also see Fig. 4b). **b** A diagram of the MTPs of the *A. tauschii* BAC contigs ctg957, ctg3280 and ctg2458. The *color horizontal bars* present individual BACs in the MTPs, and the *fine lines* connected the marker loci with the BACs containing the markers. The *purple bar* indicates genomic scaffold KD689925, which contains marker WL1445 and overlaps with BAC contig ctg957 (color figure online)

KD549154. Distal to the *Br2* locus, marker loci *XWL3041* and *XWL1373* span a distance of 0.1 cM on the genetic map and are 4,254 bp apart in the genomic sequence, estimating a genetic-to-physical distance ratio of 23.5 cM/Mb for this interval. By contrast, no recombination was detected for a cluster of nine marker loci that co-segregated with *Br2*. This cluster of marker loci spans 2,910 kb long. Comparing with the neighboring interval *XWL3041–XWL1373*, recombination is obviously suppressed in the *Br2*-containing interval.

To distinguish between the mechanisms suppressing the recombination, local epigenetic modification, or chromosome structural changes, we developed more markers surrounding the existing ones. Markers WL2911, WL2913, WL1557, WL2934, and WL2935 were developed from a 3,866-bp genomic sequence of *A. tauschii* (Fig. 4a), and this segment contained a gene encoding an MYB transcription factor, homologous to *Os01g65370* of rice and *Bradi2g56520* of *B. distachyon.* With these newly developed markers, two double-crossovers were detected in two recombinants. In recombinant 50, the double-crossovers surrounded *XWL2935*, and in recombinant 59, the double-crossovers flanked *XWL1557* (Fig. 4b), which expanded the interval to 0.06 cM (Fig. 3). Three crossovers occurred within this 3,866-bp interval, estimating a recombination rate of ~ 11 cM/Mb, which is 6.4-fold higher than the genome-wide average of 1.73 cM/Mb (Luo et al. 2013). As a result, the *Br2* region showed a unique pattern of crossover distribution: Single-crossovers only detected in the marker intervals on either side of the *Br2* locus, and double-crossovers occurred only in the marker interval that co-segregated with the *Br2* locus. These results indicated that single-crossovers among the *Br2*-co-segregating marker loci were strongly suppressed, but the double-crossovers survived, suggesting a scenario that the two parental genomes differ for an inversion polymorphism in the *Br2*-containing interval. This result also narrowed down the *Br2* locus to a 0.08-cM interval between *XWL1445* and *XWL1557* (Figs. 3, 4b).

Haplotype analysis of *Br2* region in *A. tauschii*

Upon encountering suppression of recombination in the vicinity of the *Br2* locus, we haplotyped a collection of 55 additional *A. tauschii* accessions in an attempt to localize *Br2* using an approach similar to association mapping. The collection was genotyped with 29 markers mapped to the *Br2* region. Together with AL8/78 and TA1604, 57 accessions were collected from fourteen countries ranging from west Turkey and the Middle East to the Central Asia and China (Table S1) and included several morphological variants: two *anathera*, three *meyeri*, six *strangulata*, 25 *tauschii*, and 21 *typica*. Two alleles were detected at all the marker loci in the collection except *XWL971*, *XWL1097*, and *XWL986* loci, at which three alleles were detected, and the *XWL475* locus, at which five alleles were detected (Table S4). The TA1604 allele appeared at higher frequencies compared with the AL8/78 allele in this panel. Based on marker genotypes, these 57 accessions were grouped into 21 haplotypes, *H1* through *H21*. *H2* is the largest haplotype consisting of 13 accessions, and ten haplotypes, *H1*, *H4*, *H5*, *H6*, *H8*, *H9*, *H11*, *H16*, *H17*, and *H18*, are singletons containing only one accession (Table S1). Cluster analysis further grouped the 21 haplotypes into two haplotype groups: *H1* through *H13* in the group 1 and *H14* through *H21* in the group 2 (Fig. 5). The group-1 accessions carry the TA1604 alleles, and the group-2 accessions harbor AL8/78 alleles at most marker loci in the *Br2*



b

Rec.	1445	3037	133	3039	467	1369	1357	1359	1443	2911	2913	1557	2934	2935	3041	1373	*Br2*	pheno
4	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	B
5	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	2	B
7	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	B
11	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0	2	B
16	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	B
22	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0	2	B
50	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	T
56	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	1	2	B
59	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	2	1	B
66	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	2	B
73	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0	2	B
92	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0	2	B
103	0	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	B

Fig. 4 Distribution of double-crossovers. **a** A diagram of promoter, coding region, intron, and the 5′ and 3′ untranslated regions (UTRs) of a gene encoding an MYB transcription factor. The positions of the marker loci are indicated by the *vertical bars*. The *scale bar* indicates 500 bp. **b** Mapping data showing position of crossovers. "*0*" stands for the TA1604 allele, "*1*" for the AL8/78 allele, and "*2*" for the heterozygote. For phenotypes, "*B*" represents breaking, and "*T*" tough rachis. The double-crossovers are *underlined*, and the shaded marker loci co-segregated with the *Br2* locus

interval except *XWL1557*, where all the accessions except TA1604 carried the AL8/78 allele (Fig. 6; Table S4), indicating that the TA1604 allele originated from a recent point mutation. But allele distribution at the marker loci outside of the *Br2* interval is not lineage-specific, particularly in the region distal to the *Br2* interval (Fig. 6). Morphologically, *anathera* is only found in the group 1, *meyeri* and *strangulata* only in the group 2, but the *typica* was included in both haplotype groups. Geographically, the group-1 accessions were widely scattered, from western Turkey to Transcaucasia, central Asia and China, but the group-2 accessions were concentrated along the south and west Caspian Sea coastal region with only two accessions scattered away from these areas (Fig. 7). The data suggest that the group 1 and 2 in the present study correspond to L1 and L2 lineages in the study by Wang et al. (2013), respectively, and that L1 and L2 are genetically isolated in the *Br2* interval.

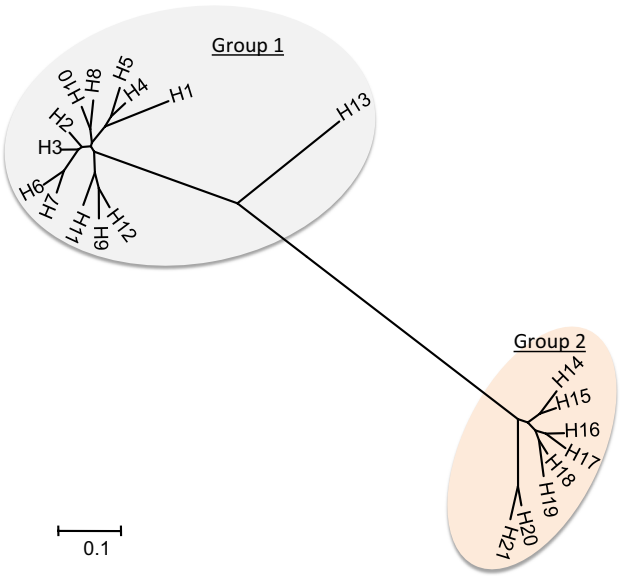


Fig. 5 Cluster analysis of 57 accessions of *A. tauschii* based on 29 CAPS and dCAPS markers from the *Br2* region. The unrooted tree was constructed with MEGA5 using UPGMA method, and 100 bootstrap replicates were conducted. *Scale bar* indicates the distance unit calculated with Jaccard coefficient. A list of these accessions, morphological classification, collection countries and sites, and their haplotypes are shown in Table S1

Discussion

A. tauschii, the D-genome donor species of hexaploid wheat, is an important genetic resource for wheat improvement and the diploid model for wheat genome sequencing. The rich genetic variation (Dvorak et al. 1998; Lubbers et al. 1991; Mizuno et al. 2010; Sohail et al. 2012; Wang et al. 2013) and genomic resources such as BAC-based physical maps, high-density genetic maps (Luo et al. 2013), the draft genome sequences (Jia et al. 2013), and soon-to-be-finished genome sequences will facilitate transfer of agronomically important genes from *A. tauschii* to common wheat and empower genetic, genomic, and molecular biology studies of wheat, particularly in a map-based cloning approach. One example is the recent isolation of

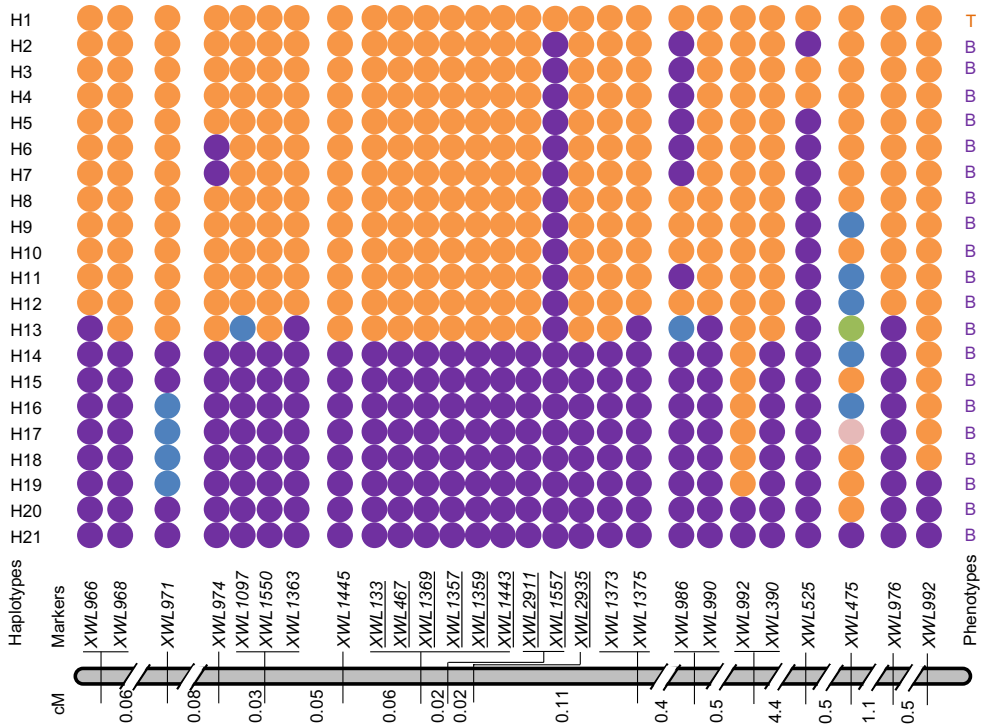


Fig. 6 Marker allele distribution among haplotypes. The marker loci are arranged in the orders as they were in the linkage map at the *bottom*. The *Br2*-co-segregating markers are *underlined*. The marker alleles were *color* coded, *orange* for the TA1604 allele, *purple* for the AL8/78 allele, and *blue*, *green*, and *ballet pink* for other alleles. Haplotypes are indicated on the *left*, and rachis disarticulation phenotype is indicated on the *right*. "*B*" represents the breaking, and "*T*" represents the tough rachis (color figure online)

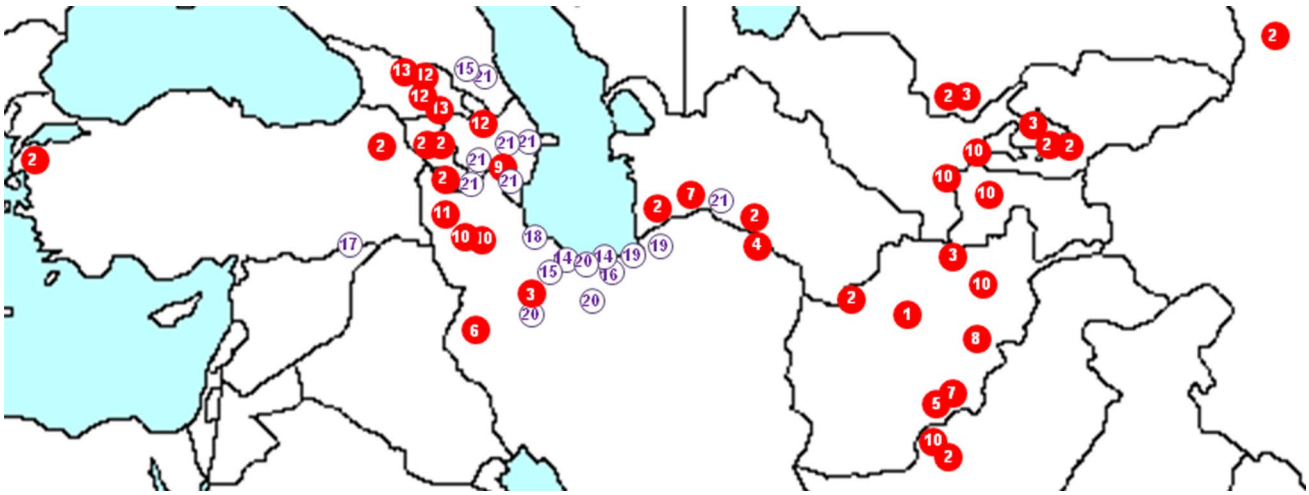


Fig. 7 Geographical distribution of *A. tauschii* haplotypes. Each spot in the map represents the collection site of an accession, and its haplotype number is indicated in the spot. The *solid dots* in *red* are for the L1 lineage, and the *purple open circles* for the L2 lineage (color figure online)

stem rust resistance gene *Sr33* from *A. tauschii* (Periyannan et al. 2013). Success of map-based cloning, however, depends not only on the physical map, but also strongly on the high resolution of the genetic map encompassing the gene of interest, which in turn depends on the local recombination rate and colinearity between the parents of the mapping population. Suppression of local recombination is a major cause for linkage drag. Fine mapping of the *Br2* locus revealed a unique distribution pattern of crossovers in the targeted genomic region and in the *A. tauschii* population. These results provide insight into the genetic mechanism underlying the *A. tauschii* population differentiation

and call attention to parent selection for constructing mapping populations and map-based cloning.

Recombination suppression and population differentiation

Several studies suggested the two lineages of *A. tauschii*, L1 and L2, are genetically isolated (Lubbers et al. 1991; Mizuno et al. 2010; Sohail et al. 2012; Wang et al. 2013). Genetic isolation can be achieved by geological and ecological separation, or reproductive barriers. Although the L1 and L2 lineages can be readily crossed and produce fertile hybrids, they occupy different geographic territories, as shown in Fig. 7. In addition to distinctive geographic distribution patterns, the lineages and sub-lineages showed adaptation to different elevations. L1 accessions are mainly collected from the high elevation sites, 400 to 3,000 m above sea level (asl), the accessions of L2 W sub-lineage were located at elevation 400 to 1,500 m asl in Transcaucasia, and accessions of L2E sub-lineage were found in Caspian sea coast of Azerbaijan and Iran at elevation not higher than 25 m asl (Wang et al. 2013).

Studies in diverse organisms demonstrated that inversions can capture the locally adapted alleles during population differentiation and speciation (Ayala et al. 2013; Guerrero et al. 2012; Lowry and Willis 2010; McGaugh and Noor 2012). To date, chromosome inversions have not been observed in *A. tauschii*. In this study, we found two double-crossovers, but no single-crossovers were detected in the 3,866-bp *XWL2911–XWL2935* interval, the proximal part of which co-segregated with *Br2*. Outside of the *Br2*-containing interval, an opposite scenario was observed in the 1,274-bp *XWL3041–XWL1373* interval, where six single-crossovers, but no double-crossovers, were detected. The contrasting crossover patterns suggest that AL8/78 and TA1604 differ by a paracentric inversion in the *Br2*-containing region, which suppressed the single-crossovers but enriched the double-crossovers. This is because a single-crossover in the inverted region would produce a dicentric chromosome and an acentric fragment, which would cause abortive gametes and could not be transmitted to the next generation in a diploid species like *A. tauschii*. Within this tentative paracentric inversion, lineage-specific distribution of marker alleles was observed at eight of nine marker loci in the *Br2*-containing region except for the allele *XWL1557*–TA1604 which is rare and originated recently (Fig. 6). Outside of the inversion region, 0.11 cM both proximal and distal, the L2-dominant allele was detected in L1 accessions and vice versa, suggesting possible introgression between the two lineages. This lineage-specific distribution pattern of alleles in the *Br2* genomic interval suggests that this inversion has probably been fixed in the two lineages of *A. tauschii* and contributed to local genetic isolation.

The size of this inversion may be ~5 Mb, which would be difficult for application in situ hybridization in a metaphase or pachytene chromosome of ~600 Mb long. Considering that the MTPs of AL8/78 BAC contigs are now being sequenced, re-sequencing of multiple genomes from the two lineages would facilitate to identify more small inversions using a comparative genomics approach. These inversions would provide a road map to identify genes under natural selection for local adaption.

Map-based cloning in *A. tauschii*

Map-based cloning is a straightforward approach for isolation of genes with unknown products, and majority of agriculturally important genes fall in this category. High polymorphism and high recombination rate are the two prerequisites for success of map-based cloning. In most cases of intraspecific crosses, the levels of recombination and polymorphism are positively correlated, and both are high in the distal regions and low in the proximal regions of the wheat genome (Qi et al. 2004). But this correlation can be reversed by chromosome inversions. In the inverted region, the polymorphism can be higher than the genome-wide average, but recombination is largely suppressed. This is what exactly found in the *Br2* region, where RFLP frequency between the two parental lines was as high as 70 % (Li and Gill 2006), and SNP frequency is 3.6-fold higher than the genome-wide average, but recombination is significantly suppressed. To take advantage of the high inter-lineage sequence polymorphism and to avoid the disadvantage caused by potential inversions, multiple mapping populations should be constructed. Populations derived from the inter-lineage crosses would be desirable for frame mapping, and those from crosses between sub-lineages may be used for fine mapping. For cloning of *Br2*, based on the allele distribution, crosses between TA1604 (*H1*) and two *H13* accessions (TA2853 and TA10104), which were collected from Georgia and belong to L1 W sub-lineage, will be made. *H1* and *H13* haplotypes are polymorphic at marker loci *XWL1097*, *XWL1363*, *XWL1557*, and *XWL1375* (Fig. 6). Combining the mapping data from the current population, the new populations derived from the intra-lineage crosses are expected to increase the map resolution. At the same time, the draft genome sequences of AL8/78 and sequences from wheat 3DL BAC contigs are expected to bridge the gap between ctg3280 and ctg957 for a continuum of DNA sequence. We have recently sequenced TA1604 genome (Wanlong Li and Ghana S. Challa, unpublished). The comparative sequence data of these two parental genomes in the *Br2* region will help identify *Br2* candidate genes based on allelic variations, and tissue-specific gene expression will narrow down the number of the candidate genes.

Transfer of useful genes from *A. tauschii* into wheat

A frequently encountered problem in transferring useful traits from wild relatives into crops is the close linkage of genes for unwanted traits with the genes of interest, i.e., linkage drag. Marker-assistant selection in a large segregating population can efficiently recover the recombination between the target gene and the drag, but the inversion polymorphism between the L1 and L2 lineages may cause linkage drag when transferring genes from *A. tauschii* to wheat. Because the D-genome of common wheat was derived from the L2 W sub-lineage (Wang et al. 2013), an L2 accession should be considered if a trait of interest is available in both L1 and L2 lineages. A successful example is the transfer of the *Lr21* resistance gene from the L2 accessions into common wheat (Huang et al. 2009). If the target trait is only available in L1 lineage and located in a rearranged region, fine mapping of a large population may be used to reduce the size of the introgressed fragment. Additionally, the unwanted gene in the transferred *A. tauschii* fragment may be knocked out by mutagenesis or by RNA-guided gene editing (Shan et al. 2013) if its sequence is known.

Author contribution statement WL conceived and supervised the project; ZZ, HZ, and WL conducted the research and analyzed data; BSG provided new materials; and WL, BSG, and ZZ wrote the paper.

Acknowledgments We thank Dr. Jan Dvorak for supplying the BAC clones and the seeds of *A. tauschii* accession AL8/78. This project was supported by USDA Hatch program through South Dakota Agricultural Experiment Station (WL) and NSF WGRC I/UCRC Award Number 1338897 (BSG).

Conflict of interest The authors declare that they have no conflict of interest.

References

Ayala D, Guerrero RF, Kirkpatrick M (2013) Reproductive isolation and local adaptation quantified for a chromosome inversion in a malaria mosquito. Evolution 67:946–958

Dudnikov AJ (2011) Waxiness in *Aegilops tauschii*: its occurrence in natural habitats of the species. Cereal Res Commun 39:283–288

Dvorak J, Luo M-C, Yang Z-L, Zhang H-B (1998) The structure of Aegilops tauschii genepool and the evolution of hexaploid wheat. Theor Appl Genet 97:657–670

Eig A (1929) Monographisch-Kritische Uebersicht der Gatung *Aegilops*. Verlag des Repertoriums, Dahlem bei Berlin

Gepts P (2004) Crop domestication as a long-term selection experiment. Plant Breed Rev 24:1–44

Guerrero RF, Rousset F, Kirkpatrick M (2012) Coalescent patterns for chromosomal inversions in divergent populations. Philos Trans R Soc Lond Ser B Biol Sci 367:430–438

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series, pp 95–98

Hammer K (1980) Vorarbeiten zur monographischen Darstellung von Wildpflanzensortimenten: *Aegilops* L. Kulturpflanze 28:33–180

Harlan JR (1992) Crops and man, 2nd edn. Am Soc, Agronomy

Huang L, Brooks S, Li W, Fellers J, Nelson JC, Gill B (2009) Evolution of new disease specificity at a simple resistance locus in a crop-weed complex: reconstitution of the Lr21 gene in wheat. Genetics 182:595–602

Jia J, Zhao S, Kong X, Li Y, Zhao G, He W, Appels R, Pfeifer M, Tao Y, Zhang X, Jing R, Zhang C, Ma Y, Gao L, Gao C, Spannagl M, Mayer K, Li D, Pan S, Zheng F, Hu Q, Xia X, Li J, Liang Q, Chen J, Wicker T, Gou C, Kuang H, He G, Luo Y, Keller B, Xia Q, Lu P, Wang J, Zou H, Zhang R, Xu J, Gao J, Middleton C, Quan Z, Liu G, Wang J, Consortium IWGS, Yang H, Liu X, He Z, Mao L, Wang J (2013) Aegilops tauschii draft genome sequence reveals a gene repertoire for wheat adaptation. Nature 496:91–95

Kimber G, Feldman M (1987) Wild wheat, an introduction. Columbia, MO

Konishi S, Izawa T, Lin S, Ebana K, Fukuta Y, Sasaki T, Yano M (2006) An SNP caused loss of seed shattering during rice domestication. Science 312:1392–1396

Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289–1291

Kosambi DD (1944) The estimation of map distances from recombination values. Ann Eugen 12:172–175

Li W, Gill BS (2006) Multiple genetic pathways for seed shattering in the grasses. Funct Integr Genomics 6:300–309

Li C, Zhou A, Sang T (2006) Rice domestication by reducing shattering. Science 311:1936–1939

Li W, Huang L, Gill BS (2008) Recurrent deletions of puroindoline genes at the grain hardness locus in four independent lineages of polyploid wheat. Plant Physiol 146:200–212

Li W, Zhu H, Challa GS, Zhang Z (2013) A non-additive interaction in a single locus causes a very short root phenotype in wheat. Theor Appl Genet 126:1189–1200

Lin Z, Griffith ME, Li X, Zhu Z, Tan L, Fu Y, Zhang W, Wang X, Xie D, Sun C (2007) Origin of seed shattering in rice (*Oryza sativa* L.). Planta 226:11–20

Lin Z, Li X, Shannon LM, Yeh CT, Wang ML, Bai G, Peng Z, Li J, Trick HN, Clemente TE, Doebley J, Schnable PS, Tuinstra MR, Tesso TT, White F, Yu J (2012) Parallel domestication of the Shattering1 genes in cereals. Nature Genet 13:720–724

Lincoln S, Lander ES (1992) Systematic detection of errors in genetic linkage data. Genomics 14:604–610

Lowry DB, Willis JH (2010) A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. PLoS Biol 8:e1000500

Lubbers L, Gill KS, Cox TS, Gill BS (1991) Variation of molecular markers among geographically diverse accessions of Triticum tauschii. Genome 34:354–361

Luo MC, Gu YQ, You FM, Deal KR, Ma Y, Hu Y, Huo N, Wang Y, Wang J, Chen S, Jorgensen CM, Zhang Y, McGuire P, Pasternak S, Stein J, Ware D, Kramer M, McCombie WR, Kianian SF, Martis MM, Mayer KF, Sehgal SK, Li W, Gill BS, Bevan MW, Simková H, Dolezel J, Weining S, Lazo G, Anderson OD, Dvorak J (2013) A 4-gigabase physical map unlocks the structure and evolution of the complex genome of Aegilops tauschii, the wheat D-genome progenitor. Proc Nat Acad Sci USA 110:7940–7945

McGaugh SE, Noor MA (2012) Genomic impacts of chromosomal inversions in parapatric Drosophila species. Philos Trans R Soc Lond Ser B Biol Sci 367:422–429

Mizuno N, Yamasaki M, Matsuoka Y, Kawahara T, Takumi S (2010) Population structure of wild wheat D-genome progenitor Aegilops tauschii Coss.: implications for intraspecific

lineage diversification and evolution of common wheat. Mol Ecol 19:999–1013

Nalam VJ, Vales MI, Watson CJ, 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

Neff MM, Neff JD, Chory J, Pepper AE (1998) dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in Arabidopsis thaliana genetics. Plant J 14:387–392

Periyannan S, Moore J, Ayliffe M, Bansal U, Wang X, Huang L, Deal K, Luo M, Kong X, Bariana H, Mago R, McIntosh R, Dodds P, Dvorak J, Lagudah E (2013) The gene Sr33, an ortholog of barley Mla genes, encodes resistance to wheat stem rust race Ug99. Science 341:786–788

Qi LL, Echalier B, Chao S, Lazo GR, Butler GE, Anderson OD, Akhunov ED, Dvorak J, Linkiewicz AM, Ratnasiri A, Dubcovsky J, Bermudez-Kandianis CE, Greene RA, Kantety R, La RCM, Munkvold JD, Sorrells SF, Sorrells ME, Dilbirligi M, Sidhu D, Erayman M, Randhawa HS, Sandhu D, Bondareva SN, Gill KS, Mahmoud AA, Ma X-F, Gustafson JP, Miftahudin, Wennerlind EJ, Nduati V, Gonzalez-Hernandez JL, Anderson JA, Peng JH, Lapitan NLV, Hossain KG, Kalavacharla V, Kianian SF, Pathan MS, Zhang DS, Nguyen HT, Choi D-W, Close TJ, McGuire PE, Qualset CO, Gill BS (2004) A chromosome bin map of 10,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. Genetics 168:701–712

Salamini F, Ozkan H, Brandolini A, Schäfer-Pregl R, Martin W (2002) Genetics and geography of wild cereal domestication in the near east. Nat Rev Genet 3:429–441

Shan Q, Wang Y, Li J, Zhang Y, Chen K, Liang Z, Zhang K, Liu J, Xi JJ, Qiu JL, Gao C (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotech 31:686–688

Sohail Q, Shehzad T, Kilian A, Eltayeb AE, Tanaka H, Tsujimoto H (2012) Development of diversity array technology (DArT) markers for assessment of population structure and diversity in *Aegilops tauschii*. Breed Sci 62:38–45

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol 28:2731–2739

Thiel T, Kota R, Grosse I, Stein N, Graner A (2004) SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. Nucleic Acids Res 32:e5

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

Watanabe N, Fujii Y, Kato N, Ban T, Martinek P (2006) Microsatellite mapping of the genes for brittle rachis on homoeologous group 3 chromosomes in tetraploid and hexaploid wheats. J Appl Genet 47:93–98

Weng Y, Li W, Devkota RN, Rudd JC (2005) Microsatellite markers associated with two Aegilops tauschii-derived greenbug resistance loci in wheat. Theor Appl Genet 110:462–469

Yildirim A, Jones SS, Murray TD, Cox TS, Line RF (1995) Resistance to stripe rust and eyespot disease of wheat in Triticum tauschii. Plant Dis 79:1230–1236

You FM, Huo N, Deal KR, Gu YQ, Luo MC, McGuire PE, Dvorak J, Anderson OD (2011) Annotation-based genome-wide SNP discovery in the large and complex Aegilops tauschii genome using next-generation sequencing without a reference genome sequence. BMC Genom 12:59

Zhou Y, Lu D, Li C, Luo J, Zhu BF, Zhu J, Shangguan Y, Wang Z, Sang T, Zhou B, Han B (2012) Genetic control of seed shattering in rice by the APETALA2 transcription factor shattering abortion1. Plant Cell 24:1034–1048