Theor Appl Genet (2013) 126:1321–1336
DOI 10.1007/s00122-013-2055-3

ORIGINAL PAPER

Genetic characterization and mapping of the *Rht-1* homoeologs and flanking sequences in wheat

Edward P. Wilhelm · Rhian M. Howells · Nadia Al-Kaff · Jizeng Jia · Catherine Baker · Michelle A. Leverington-Waite · Simon Griffiths · Andy J. Greenland · Margaret I. Boulton · Wayne Powell

Received: 17 October 2012 / Accepted: 20 January 2013 / Published online: 5 February 2013

Abstract The introgression of *Reduced height* (*Rht*)-*B1b* and *Rht-D1b* into bread wheat (*Triticum aestivum*) varieties beginning in the 1960s led to improved lodging resistance and yield, providing a major contribution to the 'green revolution'. Although wheat *Rht-1* and surrounding sequence is available, the genetic composition of this region has not been examined in a homoeologous series. To determine this, three *Rht-1*-containing bacterial artificial chromosome (BAC) sequences derived from the A, B, and D genomes of the bread wheat variety Chinese Spring (CS) were fully assembled and analyzed. This revealed that *Rht-1* and two upstream genes were highly conserved among the homoeologs. In contrast, transposable elements (TEs) were not conserved among homoeologs with the exception of intronic miniature inverted-repeat TEs (MITEs). In relation to the *Triticum urartu* ancestral line, CS-A genic sequences were highly conserved and several colinear TEs were present. Comparative analysis of the CS wheat BAC sequences with assembled *Poaceae* genomes showed gene synteny and amino acid sequences were well preserved. Further 5′ and 3′ of the wheat BAC sequences, a high degree of gene colinearity is present among the assembled *Poaceae* genomes. In the 20 kb of sequence flanking *Rht-1*, five conserved non-coding sequences (CNSs) were present among the CS wheat homoeologs and among all the *Poaceae* members examined. *Rht-A1* was mapped to the long arm of chromosome 4 and three closely flanking genetic markers were identified. The tools developed herein will enable detailed studies of *Rht-1* and linked genes that affect abiotic and biotic stress response in wheat.

Introduction

Global food shortages prevalent in the 1960s were greatly abated by the development of high yielding bread wheat (*Triticum aestivum*) varieties containing either *Reduced height* (*Rht*)-*B1b* or *Rht-D1b* in what is commonly known as

Communicated by M. Sorrells.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-013-2055-3) contains supplementary material, which is available to authorized users.

E. P. Wilhelm (✉) · R. M. Howells · A. J. Greenland · W. Powell
National Institute of Agricultural Botany, Huntingdon Rd., Cambridge CB3 0LE, UK
e-mail: edwilhelm@yahoo.com

E. P. Wilhelm · N. Al-Kaff · C. Baker · M. A. Leverington-Waite · S. Griffiths · M. I. Boulton
John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK

Present Address:
N. Al-Kaff
Taibah University, The Road of Universities, P.O. Box 344, Al-Madina Al-Munwarah, Kingdom of Saudi Arabia

J. Jia
Key Laboratory of Crop Germplasm Resources and Utilization, Ministry of Agriculture, National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081, People's Republic of China

Present Address:
W. Powell
Institute of Biological, Environmental, and Rural Sciences, Aberystwyth SY23 3DA, UK

the 'green revolution'. *Rht-B1b* and *Rht-D1b* each confer a shortened stature to wheat, which is key to producing varieties with increased lodging resistance and greater harvestable yields under intensive agricultural practices. The *Rht-B1b* and *Rht-D1b* alleles along with the *Rht-B1a* and *Rht-D1a* wild-type alleles have been sequenced (Peng et al. 1999), physically mapped to the short arms of 4B (*Rht-B1*) and 4D (*Rht-D1*) (McVittie et al. 1978), and genetically mapped in several studies (Borner et al. 1997; Peng et al. 1999; Ellis et al. 2002; Somers et al. 2004; Quarrie et al. 2005; Draeger et al. 2007; Cuthbert et al. 2008; Srinivasachary et al. 2008, 2009; Cao et al. 2009; Raquin et al. 2008). The wild-type alleles encode DELLA proteins that repress plant growth in the absence of gibberellin (GA), but when GA is present, DELLA is degraded, thereby removing this growth restriction. *Rht-B1b* and *Rht-D1b* encode altered forms of the DELLA protein that result in reduced GA sensitivity and decreased stalk length (Peng et al. 1999). In addition to *Rht-B1* and *Rht-D1*, a third homoeolocus, *Rht-A1*, was recently identified using the bread wheat variety Chinese Spring (CS; Febrer et al. 2009). *Rht-A1* was also fully sequenced in the bread wheat variety Cadenza and shown to express the DELLA protein (Pearce et al. 2011). *Rht-A1* maps to chromosome 4A (Pearce et al. 2011), but linked genetic markers have not been identified and the physical location on 4A is not known.

Recent studies indicate that DELLA proteins play a role in abiotic and biotic stress tolerance in plants. DELLA proteins appear to be directly associated with changes in tolerance to pathogens in *Arabidopsis thaliana*, barley (*Hordeum vulgare*), and wheat (Navarro et al. 2008; Saville et al. 2012) and with changes in tolerance to cold and salinity in *Arabidopsis* (Achard et al. 2006, 2008) and potassium deprivation in wheat (Moriconi et al. 2012). *Rht-D1b* and *Rht-B1b* have both been associated with changes in the levels of resistance to Fusarium Head Blight (FHB), a major disease in wheat. A quantitative trait locus (QTL) that encompassed *Rht-D1b* and was associated with increased susceptibility to FHB has been identified in several studies (Hilton et al. 1999; Draeger et al. 2007; Srinivasachary et al. 2008, 2009). A QTL containing *Rht-B1b* was associated with increased Type I (initial infection) susceptibility to FHB and, conversely, increased Type II (fungal growth within the spike) resistance to FHB (Srinivasachary et al. 2009). The association of the semi-dwarfing alleles with changes in FHB resistance may be the result of changes in plant height (Yan et al. 2011); however, height was not correlated with FHB susceptibility when *Rht-D1a* and *Rht-D1b* wheat doubled haploid sub-populations arising from a single cross were analyzed separately (Draeger et al. 2007). In addition, Srinivasachary et al. (2009) reported that although the QTL associated with FHB susceptibility contained *Rht-D1b*, the peak was consistently placed not at *Rht-D1b*, but at a nearby marker. Identification of the loci neighboring the wheat *Rht-1* homoeologs should aid in identifying the underlying cause of deleterious or beneficial traits linked to *Rht-1*.

It is well established that gene synteny in the A, B, and D wheat genomes is highly conserved and that this colinearity extends to other members of the grass (*Poaceae*) family (Moore et al. 1995; Gale and Devos 1998; Kumar et al. 2009). However, gene micro-colinearity among the grasses may often be disrupted by gene duplications and chromosomal rearrangements (Gale and Devos 1998). The synteny of the hexaploid wheat genomes has been affected by several translocations, inversions, and deletions (Naranjo et al. 1987; Mickelson-Young et al. 1995; Devos et al. 1995; Nelson et al. 1995; Miftahudin et al. 2004). In cases where multiple copies of a single gene are present due to polyploidization or segmental doubling, duplicate genes may be lost due to redundancy or there may be a change in function (Kashkush et al. 2002; Lai et al. 2004). Transposable elements (TEs), relative to genes, evolve faster and are not as well conserved among species (Devos et al. 2008). Among the A, B, and D wheat genomes, which diverged approximately two to four million years ago (MYA; Huang et al. 2002; Dvorak and Akhunov 2005; Chalupska et al. 2008), studies of homoeologous regions have shown no apparent conservation of TEs with the exception of miniature inverted-repeat TEs (MITEs), which sometimes are conserved between genomes (Anderson et al. 2002; Gu et al. 2004; Chalupska et al. 2008; Ragupathy and Cloutier 2008). Similarly, between *Brachypodium distachyon* and *Brachypodium sylvaticum*, which diverged an estimated 1.7 to 2.0 MYA, TEs were not conserved except for a low proportion of MITEs (Buchmann et al. 2012). Transposon excisions can also result in deletions or insertions in flanking sequence that can be hundreds of bases in length, further eroding genome colinearity of intragenic regions (Buchmann et al. 2012). In wheat, a large genome size and the lack of an assembled sequence make comparative analysis with assembled and annotated *Poaceae* genomes a particularly useful tool for identification of genes, structural domains, and conserved non-coding sequences (CNSs).

To determine the genes and sequence immediately surrounding *Rht-1*, wheat bacterial artificial chromosome (BAC) clones containing *Rht-1* from the A, B, and D genomes of CS were isolated, sequenced, and assembled. In addition, BAC insert sequences from the D genome of Aibai/CS (Duan et al. 2012) and from the A genome ancestor *T. urartu* were examined in relation to the CS homoeologs. Comparative analysis of the wheat BAC genes was performed with respect to *Poaceae* orthologs from rice (*Oryza sativa*), *B. distachyon*, foxtail millet (*Setaria italica*), sorghum (*Sorghum bicolor*), maize (*Zea*

mays), and barley. Colinearity of the region that extends beyond the *Rht-1* BAC sequences was also examined in the assembled *Poaceae* genomes.

Materials and methods

BAC library screening and sequencing

CS wheat BAC clones 224-M10_CS-A and 1417-F16_CS-B were isolated as described in Febrer et al. (2009). BAC clone 155-I24_CS-D was isolated from the French component of the CS library created by Allouis et al. (2003) by screening colony filters using a 195 bp radiolabeled [^{32}P] probe, which spans the DELLA protein domain and is specific to *Rht-1* (CS *Rht-D1* nucleotide coordinates 49–243). The Rediprime II DNA labeling system (GE Healthcare Life Sciences) was used to radiolabel the probe. Southern hybridization of colony filters was carried out as described by Sambrook et al. (1989). Following hybridization, filters were exposed overnight on phosphor screens and visualized with a Typhoon phosphorimager (Amersham Biosciences). *Rht-1* presence on the inserts was confirmed by PCR using the Rht-F1 and Rht-R1 primers as described in Febrer et al. (2009). BAC 50–M3_CS-B was isolated from the French component of the CS BAC library held at the French Plant Genomic Resource Center (INRA-CNRGV, Toulouse, France). To isolate BAC 50-M3_CS-B, pooled BAC library DNA was PCR-screened by the CNRGV using primers Z-F02 (5′-GGG AGG CTA GCT CAT CAT CA-3′) and Z-R02 (5′-CAA CCG CAT ACA AAG CAA AA-3′), which are specific to a C3HC4 zinc finger (*ZnF*) gene that neighbors *Rht-1*. For the *Rht-1*-containing BAC sequence derived from *T. urartu* (Tu-JJ1), shotgun libraries for selected positive BACs were constructed as described by Kong et al. (2004).

BAC inserts from 224-M10_CS-A, 1417-F16_CS-B, and 155-I24_CS-D were sequenced to 8 × coverage by shotgun sequencing and assembled into four, three, and four contigs, respectively, as a service of The Genome Institute, Washington University, St. Louis, MO, USA. The order and orientation of contigs were determined using PCR with gaps between contigs estimated at 500 bp or less (pers. comm., W. Courtney, Washington University). BAC 50–M3_CS-B was sequenced by Eurofins MWG Operon (Ebersberg, Germany) using GS FLX Titanium series chemistry (Roche Applied Science) and assembled into 19 large contigs with orientation and order not determined. The *ZnF* gene sequence on BAC 50-M3_CS-B was determined by combining three overlapping contigs using sequence from BAC 155-I24_CS-D as template. BAC Tu-JJ1 was sequenced from forward and reverse directions using Applied Biosystems (ABI) BigDye 3.1 terminator chemistry and analyzed on an ABI 3730XL automated capillary sequencer. Assembly analyses were performed using PHRED (Ewing et al. 1998) and contigs assembled with Lasergene v7.10 software (http://www.dnastar.com/) with match size 40 and minimum match percentage 98. Gaps were closed and weak consensus regions strengthened by direct sequencing of subclones using primer walking with dGTP and DMSO added to the sequencing mix.

Annotation of *Rht-1*-containing wheat BAC sequences

Locations and identities of repeat elements in the wheat BAC sequences were determined by BLASTn and tBLASTx searches of the Triticeae Repeat Sequence Database (http://wheat.pw.usda.gov/ITMI/Repeats) using an expectation value (*E*) threshold of e^{-10}. Retrotransposon long terminal repeats (LTR) and target site duplications were identified using the default settings of LTR Finder (http://tlife.fudan.edu.cn/ltr_finder; Xu and Wang 2007) and verified manually. Gene predictions were made across entire BAC inserts using Softberry FgeneSH software (http://linux1.softberry.com/berry.phtml). Nucleotide sequences of predicted genes were queried against the Triticeae Repeat Sequence Database, the National Center for Biotechnology Information (NCBI) nucleotide collection (http://blast.ncbi.nlm.nih.gov/Blast.cgi), The Institute for Genomic Research (TIGR) plant transcript assemblies (TAs) monocot collection (http://plantta.jcvi.org), and the Knowledge-based Oryza Molecular biological Encyclopedia rice cDNA collection (http://cdna01.dna.affrc.go.jp/cDNA) to identify regions of homology. Predicted genes were considered true genes if orthologous genes were identified in the database searches and were not considered true genes if the sequence had a partial or full match to a repeat element. If the predicted gene sequence did not have a match to an orthologous gene or to a repeat element, it remained a predicted gene. Intron–exon boundaries of genes were determined by manually aligning sequences with TIGR wheat TAs, TIGR barley TAs, and annotated orthologs in rice, *B. distachyon*, sorghum, and maize (B73 inbred) using GeneDoc v2.6.002 software (Nicholas and Nicholas 1997). The remaining BAC sequences that did not correspond to a TE, gene, or predicted gene were BLAST-searched against the NCBI database nucleotide collection to determine if any similarity existed. Annotation of BAC sequences was facilitated with Vector NTI v10.1.1 (Invitrogen) software.

Comparative analysis with *Poaceae* orthologs

Sequences of the *ZnF*, *EamA*, and *Rht-1* orthologs in rice, *B. distachyon*, sorghum, maize, and foxtail millet were

identified by nBLAST searches of the Phytozome v8.0 databases (Goodstein et al. 2012) using the wheat open reading frames (ORFs) as query. Primary transcript models from Phytozome were utilized in *ZnF*, *EamA*, and *Rht-1* ORF amino acid (AA) alignments, with the following two exceptions where alternate transcripts showed improved alignments with *Poaceae* orthologs: Foxtail millet *ZnF* (*Si034920m* with the 5′ end truncated) and maize *ZnF* from chromosome 1 (*GRMZM2G704032*). To identify orthologs in the unassembled barley genome, wheat gene ORFs were used in nBLAST searches of the NCBI nucleotide collection, the NCBI barley expressed sequenced tag (EST) collection, and the TIGR barley TA collection. Orthologs in *Arabidopsis*, grape (*Vitis vinifera*), *Eragrostis tef*, and sugarcane (*Saccharum officinarum*) were found by nBLAST searches of the NCBI nucleotide collection. To examine synteny among the assembled *Poaceae* genomes in the region that extends beyond that orthologous to the wheat BAC regions, predicted gene ORFs from rice were used in nBLAST searches of the *B. distachyon*, foxtail millet, sorghum, and maize genomes using Phytozome v8.0.

AA alignments of orthologs were made using ClustalX v2.0.12 (Larkin et al. 2007) and manually adjusted where necessary to improve alignments using GeneDoc v2.6.002. Locations of protein motifs on the *EamA* and *ZnF* AA sequences were determined using the European Bioinformatics Institute InterPro scan (http://www.ebi.ac.uk/Tools/InterProScan) with AA sequences serving as query. Phylogenetic relationships among orthologous genes were determined using the following applications of PHYLIP v3.6 software (Felsenstein 2004): 'Seqboot' was used to calculate bootstrap values (1,000 replicates); 'Protdist' was used to calculate genetic distance between species (10,000 multiple data set values); 'Neighbor', which employs the neighbor-joining method, was used to construct trees; 'Consense' was used to determine the consensus tree.

Identification of CNS regions

CNS regions were determined with the NCBI alignment tool using the BLASTn program with an *E* threshold of e^{-10}. CNSs were aligned with ClustalX v2.0.12 and manually edited, as needed, using GeneDoc v2.6.002.

Rht-A1 mapping

Aneuploid stocks were supplied by S. Reader, John Innes Centre (JIC), Norwich, UK. CS deletion stocks used were obtained from the Wheat Genetics Resource Center, Kansas State University, Manhattan, KS, USA. The biparental F_5 population 'Sears Synthetic 7010073 × Paragon' ('SS7010073 × Paragon') was developed by the Department for Environment Food and Rural Affairs and funded by the Wheat Genetic Improvement Network. 'SS7010073' is a synthetic hexaploid of *Triticum dicoccum* × *Aegilops tauschii* created by E. Sears, University of Missouri, Columbia, MO, USA (S. Reader, JIC, pers. comm.). Paragon is an elite UK spring wheat. Seed of SS7010073 was received from the JIC and seed of Paragon was received from the National Institute of Agricultural Botany (Cambridge, UK). Genomic DNA (gDNA) of aneuploid stocks, deletion lines, 'SS7010073', and 'Paragon' was extracted using a modification of the method described by Fulton et al. (1995). gDNA of the 'SS7010073 × Paragon' F_5 population was extracted using the DNeasy kit (Qiagen) according to manufacturer instruction.

For the physical mapping of *Rht-1*, gDNA extracts of aneuploid and deletion lines were used as templates in PCRs containing the *Rht-1* generic primer Rht-ABD-R6 (5′-TGC ATC CCC TGC TTG ATG-3′) along with one of the following locus specific primers: Rht-A-F3 (5′-GAT GCC GTC TCG CAA TCT-3′) for *Rht-A1*; Rht-B-F1 (5′-AGG CAA GCA AAA GCT TGA GA-3′) for *Rht-B1*; Rht-D-F1 (5′-CGA GGC AAG CAA AAG CTT C-3′) for *Rht-D1*. PCRs were performed in 10 µl volumes containing 1× Green GoTaq Reaction Buffer (Promega), 3 % glycerol, 0.2 mM of each dNTP, 2 mM $MgCl_2$, 1 µM forward primer, 1 µM reverse primer, 0.25 µl *Taq* polymerase, and 20 ng DNA. The PCR profile consisted of 5 min at 95 °C, followed by 40 cycles of 95 °C for 30 s, 64 °C anneal temp. for 30 s, and extension of 72 °C for 90 s, and concluded with 72 °C for 5 min. Amplified products were separated in a 1.5 % agarose gel in 1× TBE buffer and visualized under UV light with ethidium bromide. Expected product sizes were 1,304 bp (*Rht-A1*), 1,105 bp (*Rht-B1*), and 1,118 bp (*Rht-D1*).

For mapping *Rht-A1*, a genetic marker was created based on a 3 bp nucleotide deletion in SS7010073 relative to the intact Paragon allele that occurs approximately 150 bp upstream of the *Rht-1* start codon. The primers Rht-A-F3 and PS-Rht-R2 (5′-GGA GGA AGA AGG AGG AAG AAT A-3′) amplify a 120 bp product in Paragon and no product in SS7010073. Primers Rht-A-F3 and PS-Rht-R4 (5′-GGA GGA AGA AGG AGG AAG AAT G-3′) amplify a 120 bp product in SS7010073 and no product in Paragon. Reaction mixtures and conditions were as described above with an annealing temperature of 60 °C and a 30 s extension time for both primer pair mixes.

Simple sequence repeat (SSR) primer sequences were acquired from the Graingenes database (http://wheat.pw.usda.gov). Forward primers were labeled with the dyes FAM, VIC, NED, or PET (ABI) according to Schuelke (2000). PCR mixes were in 6.25 µl volumes that consisted of 3.125 µl Hotstar Taq Master Mix (Qiagen),

0.75 μM of each primer, and 12.5 ng gDNA. The PCR profile consisted of 15 min at 95 °C, followed by 35 cycles of 95 °C for 1 min, a primer pair-dependent annealing temperature according to the Graingenes website for 1 min, and 72 °C for 1 min, and concluded with 72 °C for 10 min. Products were measured on an ABI 3730 DNA Analyzer with a POP-7^{TM} polymer column. Linkage maps were based on *Rht-A1* and SSR marker scores of 94 individuals from the ('SS7010073 × Paragon') biparental F_5 population and created using JoinMap version 4.0 with a log of odds threshold of 5.0 (Van Ooijen 2006).

Accession numbers for the wheat BAC sequences are: JX978692 (224–M10_CS-A), JX978693 (1417–F16_CS-B), JX978694 (155–I24_CS-D), JX978691 (50–M3_CS-B), and JX978695 (Tu-JJ1).

Results

Composition of wheat BACs

Three of the *Rht-1*-containing clones identified in the CS BAC library representing the A genome (224-M10_CS-A), B genome (1417-F16_CS-B), and D genome (155-I24_CS-D) were sequenced and assembled. *Rht-1* was centrally located in each insert and the sequenced length of inserts ranged from 164 to 213 kb (Fig. 1; Table 1). BAC Tu-JJ1, derived from *T. urartu*, was approximately 100 kb in length with *Rht-1* positioned near the middle. BAC 1J9_D, derived from the 'D' genome of the near-isogenic line Aibai/CS (GenBank acc. no. HQ435325.1; Duan et al. 2012), was approximately 207 kb in length with *Rht-1* located just under 10 kb from the 3′ end. BAC clones 224-M10_CS-A, 155-I24_CS-D, Tu-JJ1, and 1J9_D each contained the same three genes in conserved order and orientation. From 5′ to 3′, these were: a zinc finger C3HC4 type domain containing protein (rice ortholog *Loc_Os03g49900*; herein referred to as *ZnF*), an integral membrane protein containing an EamA-like transporter (rice ortholog *Loc_Os03g49940*; previously known as Domain of Unknown Function 6 protein; herein referred to as *EamA*), and *Rht-1* (rice ortholog *Loc_Os03g49990*). BAC clone 1417–F16_CS-B contained *EamA* and *Rht-B1*, but did not contain sufficient sequence 5′ of *EamA* to determine if *ZnF* was the next gene upstream. The B genome copy of *ZnF* from CS was subsequently identified on BAC 50–M3_CS-B. This BAC clone was sequenced and assembled into 19 contigs (order and orientation not determined) with total sequence length estimated at 170 kb. Annotation of BAC 50-M3_CS-B indicated there was no overlapping sequence with BAC 1417-F16_CS-B. *ZnF* was the only gene present on BAC 50-M3_CS-B, residing on a 31 kb region assembled from three overlapping contigs. BACs 1J9_D and 0155-I24_CS-D have 105 kb of overlapping sequence that has 99.83 % nucleotide identity and contains the three identified genes. Between the two D genome BACs, 100 % nucleotide identity is retained between the two *ZnF* ORFs and also between the *EamA* ORFs, while the *Rht-D1* ORFs differ by a single nucleotide polymorphism (SNP). The *Rht-D1* SNP in 1J9_D, previously identified by Duan et al. (2012), results in a premature stop codon characteristic of the *Rht-D1b* semi-dwarfing allele (Peng et al. 1999), whereas CS contains the *Rht-D1a* allele. The combined length of the two D genome BACs spans approximately 316 kb.

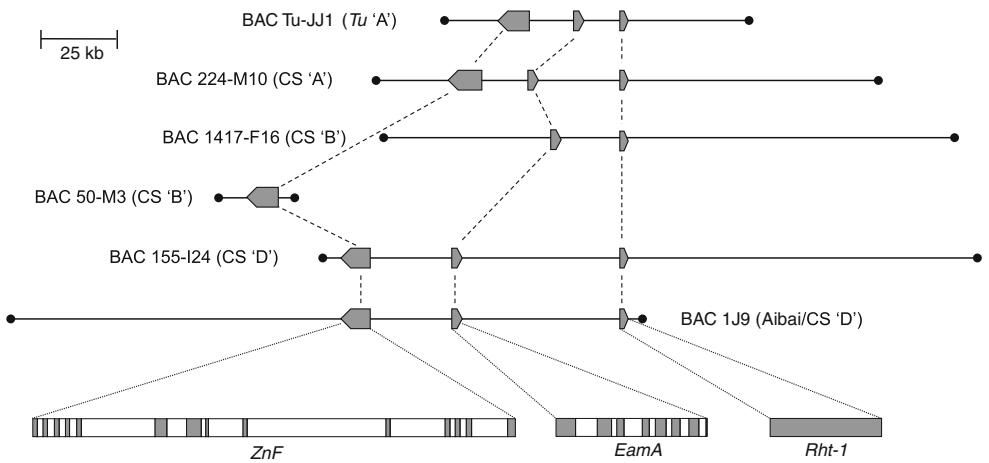


Fig. 1 Gene content of wheat BAC clones. Genes and direction of transcription are shown on each BAC insert. BACs are identified by clone address with genetic background and genome shown in parentheses (*Tu*, *Triticum urartu*; CS, Chinese Spring). *Filled circles* denote the boundaries of assembled BAC sequence. *Dashed lines* connect homoeologs. Exon–intron structure of the C3HC4 zinc finger (*ZnF*), *EamA*, and *Rht-1* genes are shown in enlarged images. BAC 50-M3 was not fully assembled and only the portion containing a gene is shown. BAC 1J9 is from Duan et al. (2012), GenBank acc. no. HQ435325.1

Table 1 Genetic composition of wheat BAC inserts

Clone address[a]	Genetic background	Base pairs	TE (%)	Gene (%)[b]	No similarity (%)[c]
224M-10	*Ta* CS (A)	164,257	55.9	8.7	35.4
1417-F16	*Ta* CS (B)	187,310	75.6	2.3	22.1
155-I24	*Ta* CS (D)	213,794	71.5	7.7	20.9
Tu-JJ1	*T. urartu* (A)	100,141	43.4	13.5	43.1
1J9	*Ta* Aibai/CS (D)	207,530	73.3	6.1	20.5
50-M3	*Ta* CS (B)	170,073	65.5	5.3	29.2

Ta Triticum aestivum, *CS* Chinese Spring with genome in parenthesis, *TE* transposable element

[a] The clone address for *T. urartu* is not available and is designated Tu-JJ1. BAC 1J9 is from Duan et al. (2012), GenBank acc. no. HQ435325.1

[b] Includes introns and exons

[c] Indicates sequence with no similarity to TEs or genes

TEs constitute from 43.4 to 75.6 % of each BAC insert, while sequence with homology to annotated genes (including introns and exons) ranged from 2.3 to 13.5 % of each insert (Table 1). While the *ZnF-EamA-Rht-1* gene synteny is highly conserved among the wheat homoeologs, there is no apparent conservation of TEs among the homoeologs with the exception of several MITEs located within gene introns (Online Resources 1 to 6). Between CS-A and its ancestral line *T. urartu*, intergenic TEs are partially conserved, with several corresponding TEs present in both BACs upstream of *EamA* and downstream of *Rht-A1* (Online Resources 1 and 4). Between 20.5 and 43.1 % of each BAC sequence did not have significant ($E < e^{-10}$) homology to annotated genes or TEs in the searched databases. One of these regions of "no similarity" occurs immediately upstream of the *Rht-1* ORF in each of the five *Rht-1*-containing BACs and is 8 kb or more in length. The largest region of "no similarity" occurs on 224-M10_CS-A in the region from approximately 37 to 51 kb downstream of *Rht-A1*. This region of CS-A contains the only predicted gene among the BAC inserts that did not have homology to TEs or annotated genes. The predicted gene on CS-A consists of a single exon of 582 bp, which has full-length matches to wheat ESTs CK209908 ($E = 5e^{-26}$) and CK209889 ($E = 9e^{-23}$). CK209908 is annotated as a homeobox protein DLX-2 related cluster and CK209889 is a predicted signal transduction protein containing EAL and modified HD-GYP domains.

Comparative analysis of *Rht-1*, *ZnF*, and *EamA* ORFs

The *Rht-1* homoeologs in CS are composed of a single exon with lengths of 1,863 bp (620 residues) in *Rht-A1*, 1,866 bp (621 residues) in *Rht-B1*, and 1,872 bp (623 residues) in *Rht-D1*. Among all three CS homoeologs, 94.1 % of the nucleotide identities and 96.8 % of the AA identities are shared. CS *Rht-A1* and CS *Rht-D1* share 96.4 % nucleotide identity (differing by 40 SNPs and the need to insert 27 gaps for alignment) and 98.1 % residue identity (seven AA substitutions and five gaps). CS *Rht-A1* and CS *Rht-B1* have 96.0 % nucleotide identity (61 SNPs and 15 gaps) and 98.1 % residue identity (five AA substitutions and seven gaps). CS *Rht-B1a* and CS *Rht-D1a* have 95.6 % nucleotide identity (56 SNPs and 26 gaps) and 97.4 % residue identity (eight AA substitutions and eight gaps). The *T. urartu Rht-A1* AA sequence is identical to CS *Rht-A1*. The CS *Rht-1* AA sequences were aligned to the *Arabidopsis* ortholog *Gibberellic Acid Insensitive* (*GAI*; *At1g14920.1*) and to the following *Poaceae* orthologs: *Slender1* (barley), *Bradi1g11090.1* (*B. distachyon*), *Slender1* (rice), *Si039400m* (foxtail millet), *Sb01g010660* (sorghum), GenBank acc. no. JN793959 (*E. tef*), *Dwarf8* (maize chromosome 1), *Dwarf 9* (maize chromosome 5), and GenBank acc. no. DQ062091.1 (sugarcane) (Online Resource 7). Among the *Poaceae* members, the DELLA protein consisted of a single exon containing between 618 and 630 AAs. The motifs of the DELLA protein were well conserved among all *Poaceae*, with the exceptions of approximately 20 bp of the PFYRE domain and the end residues of the poly S/T/V domain. In the regions of the protein motifs that are highly conserved among the *Poaceae*, the AA sequences of the CS homoeologs are identical to one another with the exception of an L in *Rht-D1* to F substitution in *Rht-A1* and *Rht-B1* in the LHR1 domain (Online Resource 7, position (pos.) 269). *Arabidopsis GAI* consisted of a single exon, but only contained 533 AAs. The protein motifs of *GAI* were less well conserved relative to the *Poaceae* DELLA orthologs and, in particular, the poly S/T/V domain has few AAs in common with any *Poaceae* member. The high level of conservation in the functional domains and the lack of a stop codon or frameshift mutation suggests that the *Poaceae* DELLA orthologs are functional and may share a similar role. Among the *Poaceae* members, barley *Slender1* and *Bradi1g11090.1* were the most closely related to wheat *Rht-1*, sharing an average of 95 and 88 % of the residues,

respectively, across the three wheat homoeologs. The remaining *Poaceae* sequences shared from 80 to 84 % of the residues with the wheat homoeologs, while *GAI* shared only 52 % of the residues.

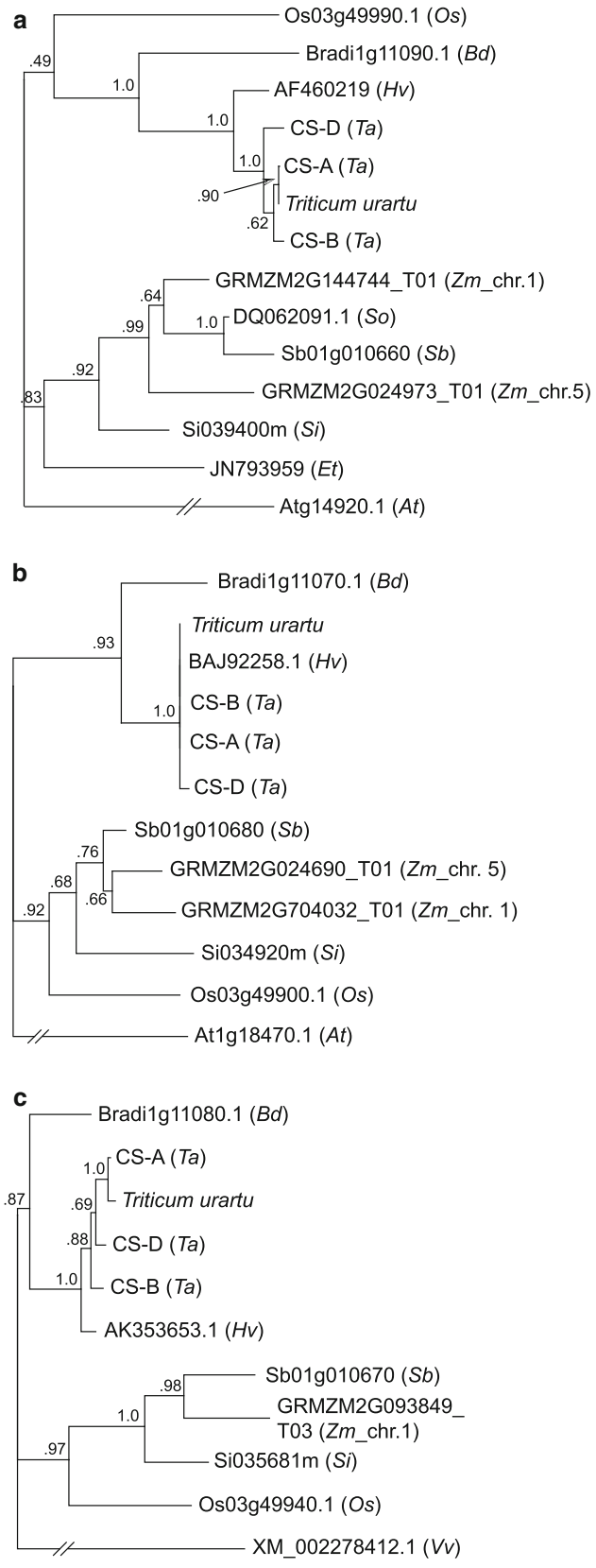
The AA sequences of the C3HC4 *ZnF* wheat homoeologs were aligned with the *Arabidopsis* ortholog *At1g18470.1* and with the following *Poaceae* orthologs: GenBank acc no. BAJ92258.1 (barley), *Bradi1g11070.1* (*B. distachyon*), *Loc_Os03g49900.1* (rice), *Si034920m* (foxtail millet), *Sb01g010680.1* (sorghum), *GRMZM2G704032_T01* (maize chromosome 1), and *GRMZM2G024690_T01* (maize chromosome 5) (Online Resource 8). The entire *ZnF* AA sequence was very highly conserved among all species including *Arabidopsis*. In each ortholog, *ZnF* consisted of 14 exons and 13 introns. Predicted peptide lengths among the *Poaceae* ranged from 472 to 474 AAs, and the *Arabidopsis* ortholog consisted of 467 AAs. The *ZnF* proteins contain a region with similarity to a C3HC4 domain, which is a cysteine-rich domain that coordinates two zinc ions. Using the CS-D *ZnF* as query, sequence with similarity ($E = 1.9e^{-12}$) to a C3HC4 domain was identified near the C terminus (Online Resource 8, pos. 424 to 467). A region with similarity ($E = 2.8e^{-73}$) to a Transmembrane Fragile X-F protein domain was also indentified, which spans the majority of the ORF (Online Resource 8, pos. 35 to 299). The absence of frameshift mutations, the lack of predicted stop codons, and the high level of conservation among the AA sequences of the *ZnF* orthologs suggest that these genes are active and likely play a similar role. Among the CS *ZnF* protein sequences, only two AA changes differentiate the homoeologs (Online Resource 8, pos. 315, CS-D = L, CS-A and CS-B = V; pos. 338, CS-B = Y, CS-A and CS-D = C). Both changes occur between the Transmembrane Fragile X-F and C3HC4 zinc finger domains. The CS-A and *T. urartu ZnF* AA sequences are identical. The barley *ZnF* ortholog had the greatest AA identity relative to the CS wheat *ZnF* sequences, sharing an average of 99 % of the residues with the three homoeologs. The *B. distachyon ZnF* ortholog shared 95 % of the residues with the CS homoeologs. The remaining *Poaceae ZnF* orthologs shared 90–92 % of the identities with the CS homoeologs and the *Arabidopsis ZnF* ortholog shared 67 % residue identity relative to the CS homoeologs.

The AA sequences of the wheat *EamA* homoeologs were compared to the following *Poaceae* orthologs: GenBank acc. no. AK353653.1 (barley), *Bradi1g11080.1* (*B. distachyon*), *Loc_Os03g49940.1* (rice), *Si035681m* (foxtail millet), *Sb01g010670.1* (sorghum), and *GRMZM2G093849_T03* (maize chromosome 1). No *EamA* ortholog was present on maize chromosome 5. There was no significant match to the wheat *EamA* gene in *Arabidopsis*, therefore a putative ortholog in grape (*Vitis vinifera*; GenBank acc. no. XM_002278412.1) was used to represent a non-*Poaceae* species. Relative to *ZnF* and *Rht-1*, *EamA* showed reduced conservation of AAs among the CS wheat homoeologs and among the *Poaceae* orthologs (Online Resource 9). In each *Poaceae* ortholog, *EamA* consists of eight exons, while the grape *EamA* ortholog consisted of seven predicted exons. The EamA proteins of the *Poaceae* contain between 430 and 454 AAs with 461 present in grape. The *EamA* proteins contain two regions with similarity to an EamA type domain (Online Resource 9, pos. 156 to 255, $E = 7e^{-7}$; pos. 345 to 415, $E = 7.1e^{-13}$). No predicted frameshift mutations or stop codons are present among the *Poaceae* species studied, suggesting that those proteins are functional. However, the high level of sequence divergence among the *Poaceae EamA* orthologs suggests the possibility of functional alterations among the proteins. Among the CS wheat *EamA* homoeologs, the CS-A and CS-D genome AA sequences are the most similar having 96.8 % identity (14 substitutions). The CS-B and CS-A genome *EamA* homoeologs have 95.2 % AA identity (17 substitutions and a 4 bp deletion in CS-B *EamA*, which occurs at pos. 34 to 37, Online Resource 9). The CS-B and CS-D *EamA* homoeologs have 94.8 % AA identity (19 substitutions and a 4 bp deletion in CS-B *EamA*). The *EamA* genes from CS-A and *T. urartu* have three AA substitutions (99.3 % identity). Relative to the CS *EamA* homoeologs, the barley *EamA* ortholog shared an average of 91 % AA identity and *B. distachyon EamA* shared an average of 85 % AA identity. The remaining *Poaceae* orthologs shared from 72 to 75 % of the residue identities with the CS *EamA* homoeologs and grape *EamA* shared 48 % AA identity with the CS *EamA* homoeologs.

Phylogenetic comparisons of the *Rht-1*, *ZnF*, and *EamA* orthologs reflect the close genetic similarity among the CS homoeologs and *T. urartu* (Fig. 2). Relative to the *Triticum* genomes, the barley orthologs of each gene are the most closely related followed by the *B. distachyon* orthologs, which had a closer similarity to wheat than did rice or the remaining *Poaceae*.

Conserved regions 5′ and 3′ of *Rht-1* among the *Poaceae*

Non-coding sequence 5′ and 3′ of the *Rht-1* ORF was examined to determine if there were any conserved regions among the three CS wheat genomes. It was hypothesized that regions of high conservation close to the ORF could reflect regulatory regions for expression of *Rht-1*. For each CS *Rht-1* homoeolog, similarity to TEs predominated beyond 10 kb of either end of the ORF (Online Resources 1, 2, and 3). TEs are unlikely to be involved in gene regulation; therefore, the analysis was restricted to the regions from 0 to 10 kb upstream and 0 to 10 kb downstream of the ORF. In



◀**Fig. 2** Unrooted neighbor-joining trees of (**a**) *Rht-1*, (**b**) *ZnF* and (**c**) *EamA* orthologs in wheat and *Poaceae* orthologs. CS-A (*Ta*), CS-B (*Ta*) and CS-D (*Ta*) represent, respectively, the Chinese Spring *Triticum aestivum* A, B, and D homoeologs of each gene. *Poaceae* orthologs are shown with locus name or GenBank accession number followed by genus and species initials in parenthesis. *At*, *Arabidopsis thaliana*; *Bd*, *Brachypodium distachyon*; *Et*, *Eragrostis tef*; *Hv*, *Hordeum vulgare*; *Os*, *Oryza sativa*; *Sb*, *Sorghum bicolor*; *Si*, *Setaria italica*; *So*, *Saccharum officinarum*; *Vv*, *Vitis vinifera*; *Zm*, *Zea mays*. Branch lengths indicate relative genetic distance. Bootstrap frequencies are shown at each node

each comparison, CS-D sequence from these regions served as the query with an *E* threshold of e^{-10}. Among the wheat *Rht-1* homoeologs, for the regions and criteria defined above, CS-D and CS-A had 6.4 kb of similar upstream sequence and 3.2 kb of similar downstream sequence. CS-D and CS-B were similar across 3.6 kb of upstream and 3.8 kb of downstream sequence. An additional comparison of this region was carried out between CS-A and CS-B and this revealed that 3.4 kb of sequence upstream of the ORF and 3.6 kb of downstream sequence were similar.

Using the same criteria described above, five CNS regions surrounding *Rht-1* that were common to CS-A, CS-B, CS-D, *B. distachyon*, rice, foxtail millet, sorghum, maize chromosome 1, maize chromosome 5, barley, and *E. tef* were identified (Fig. 3). The CNS locations and lengths matched those recently identified by Duan et al. (2012) when examining this region in the wheat D genome, *A. tauschii*, *B. sylvaticum*, *B. distachyon*, rice, maize chromosome 1, and sorghum. In our analysis, CNS regions 1, 2, 3, and 5 had 67 % or greater nucleotide identity between any two species, while CNS4 had a nucleotide identity of 60 % or greater between any two species. CNS1 in foxtail millet is bisected by a 344 bp segment of low complexity sequence, which is similar ($E = 9e^{-76}$) to a Tourist TSI-1 MITE. A predicted gene (*Si040419m*) also exists in foxtail millet between CNS3 and CNS4, but does not overlap the conserved regions. No other predicted genes were present in the examined regions of the *Rht-1* ORF in any of the species examined. Similar to Duan et al. (2012), the longest continuous string of invariant residues (37 residues) present in all species was found in CNS5, however, a SNP present in foxtail millet at the fifth nucleotide shortened the length of invariant nucleotides to 32.

Microsynteny in the *Rht-1* region among *Poaceae*

The wheat BACs were aligned with fully assembled members of the *Poaceae* (rice, *B. distachyon*, foxtail millet, sorghum, and maize) to determine if gene synteny was preserved (Fig. 4; Online Resource 10). A single

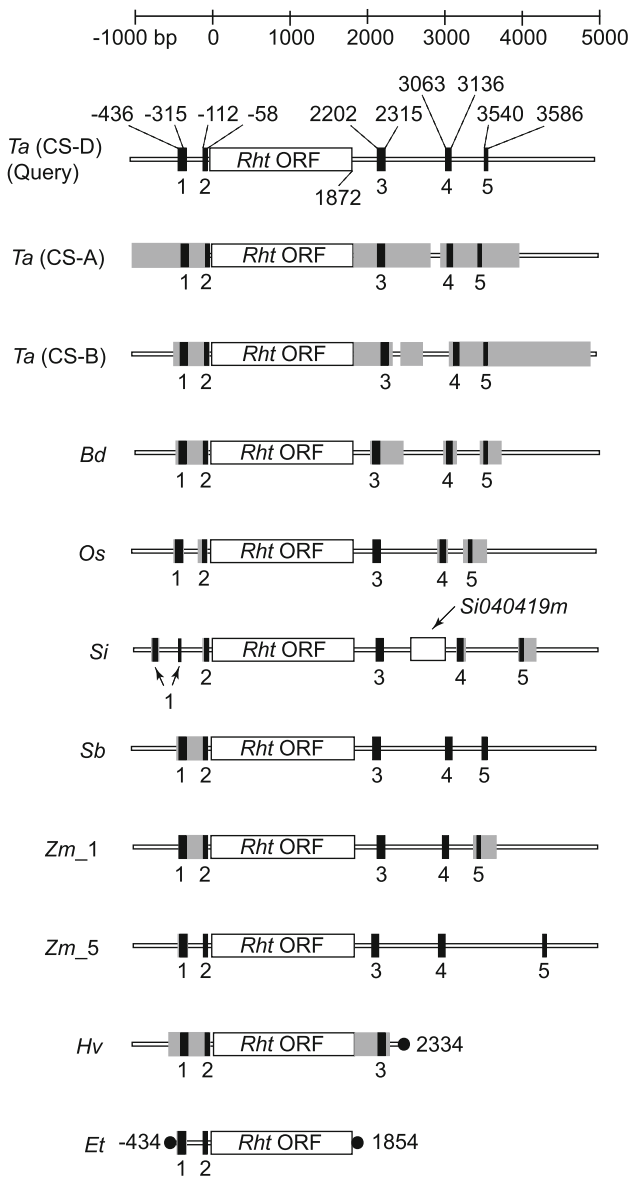


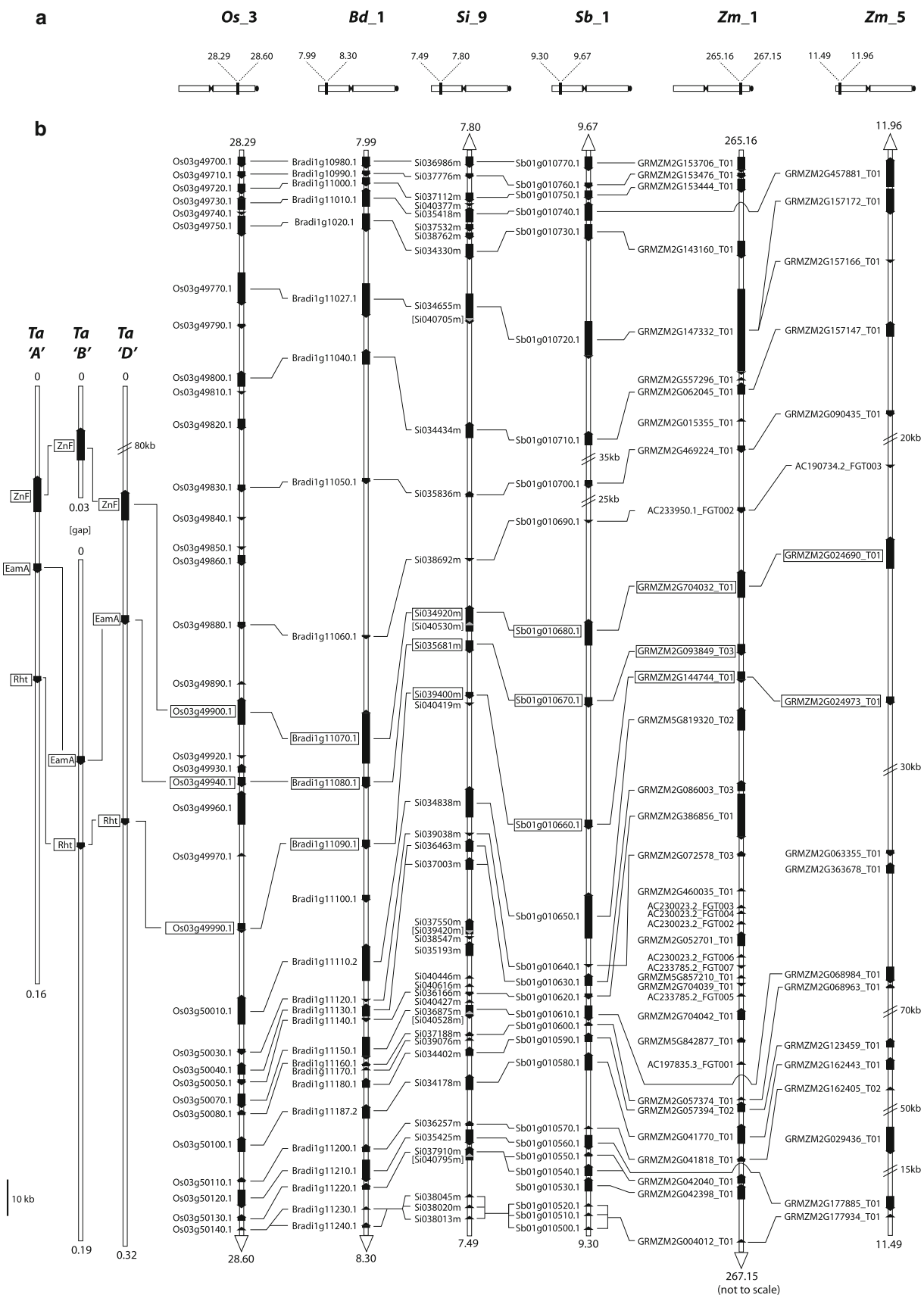
Fig. 3 *Rht-1* conserved non-coding sequences (CNSs) in the *Poaceae*. *Triticum aestivum* Chinese Spring D genome (*Ta* CS-D) sequence (10 kb 5′ of the *Rht-1* ORF start and 10 kb 3′ of the ORF end) served as query to identify homologous sequence in the *Poaceae* species (*Bd*, *Brachypodium distachyon*; *Os*, *Oryza sativa*; *Si*, *Setaria italica*; *Sb*, *Sorghum bicolor*; *Zm*_1, *Zea mays* chromosome 1; *Zm*_5, *Zea mays* chromosome 5; *Hv*, *Hordeum vulgare*; *Et*, *Eragrostis tef*). Regions conserved with respect to CS-D are indicated by *gray rectangles*. Regions conserved among all *Poaceae* are indicated by *black rectangles* numbered 1–5, which all occurred in the region from 1,000 bp upstream (negative numbers) to 5,000 bp downstream relative to the *Rht-1* start nucleotide. Nucleotide coordinates of the CNSs in reference to CS-D are indicated. *Si040419m* is a predicted gene. *Black dots* represent truncated sequences with start and end base pair coordinates shown

orthologous region containing *ZnF-EamA-Rht-1* was identified in rice (chr. arm 3L), *B. distachyon* (chr. arm 1S), foxtail millet (chr. arm 9S), and sorghum (chr. arm 1S). In maize, two orthologous regions were identified, located on chromosome arms 1L (*Zm*_1) and 5S (*Zm*_5). In each orthologous region, *ZnF-EamA-Rht-1* synteny is retained with the exception of *Zm*_5, in which *EamA* is absent. In rice, four additional predicted genes that are expressed exist within the *ZnF-EamA-Rht-1* linkage group, but only one has a functional annotation and none of the predicted genes has an ortholog in the *Poaceae* species examined. Foxtail millet also contained an additional predicted gene (*Si040530m*) located in intron 5 of the *ZnF* gene, but it has no functional annotation nor does it have a known *Poaceae* ortholog.

Colinearity and gene density of the region flanking *Rht-1* from 17 genes upstream of *ZnF* to 11 genes downstream of *Rht-1* in rice was then examined in rice, *B. distachyon*, foxtail millet, sorghum, and maize chromosomes 1 and 5 (Fig. 4; Online Resource 10). A high degree of colinearity in the *Rht-1* region had been previously reported among rice, sorghum, maize chromosome 1, and the D genome of wheat (Duan et al. 2012). In our study, a high degree of synteny in the *Rht-1* region was present in each of the species examined; however, on maize chromosome 5, several genes that were present in the other *Poaceae* members were absent. Across the entire region, there are 23 genes with orthologs in each of rice, *B. distachyon*, foxtail millet, sorghum, and maize. One gene (*Bradi1g11180.1*) has orthologs in all species except rice, two genes have two orthologs each, and 46 genes have no orthologs (12 in rice; one in *B. distachyon*; 15 in foxtail millet; 18 in maize, with 15 of these on *Zm*_1 and three on *Zm*_5). Among the maize genes with a known ortholog in another *Poaceae* member, 11 occur on both *Zm*_1 and *Zm*_5, 11 occur on only *Zm*_1, and three exist only on *Zm*_5. The only rearrangement of gene order in the *Rht-1* region occurred on *Zm*_1 where the position of *GRMZM2G072578* is two genes more distal to the *Rht-1* ortholog than in the other *Poaceae*. Four of the 23 genes contained apparent duplicates in at least one species. The genetic region examined encompasses approximately 300 kb in each of *B. distachyon* (one gene per 11.2 kb), rice (one gene per 8.6 kb), and foxtail millet (one gene per 7.0 kb). The region totals 362 kb in sorghum (one gene per 12.9 kb), 470 kb in *Zm*_5 (one gene per 26.1 kb) and 1,979 kb in *Zm*_1 (one gene per 52.1 kb). In wheat, 316 kb of CS-D BAC sequence contained only three genes (1 gene per 105.3 kb), 164 kb of CS-A BAC sequence contained only three genes and one predicted gene (1 gene per 41 kb), and 366 kb of CS-B BAC sequence contained only three genes (1 gene per 122 kb).

Rht-1 mapping in wheat

The approximate physical locations of *Rht-1* on the A, B, and D genomes of wheat were determined using gDNA of


a
Os_3
Bd_1
Si_9
Sb_1
Zm_1
Zm_5
28.29
28.60
7.99
8.30
7.49
7.80
9.30
9.67
265.16
267.15
11.49
11.96
b
Ta 'A'
Ta 'B'
Ta 'D'
ZnF
EamA
Rht
80kb
0.03
[gap]
0.16
0.19
0.32
10 kb
35kb
25kb
20kb
30kb
70kb
50kb
15kb
Os03g49700.1
Os03g49710.1
Os03g49720.1
Os03g49730.1
Os03g49740.1
Os03g49750.1
Os03g49770.1
Os03g49790.1
Os03g49800.1
Os03g49810.1
Os03g49820.1
Os03g49830.1
Os03g49840.1
Os03g49850.1
Os03g49860.1
Os03g49880.1
Os03g49890.1
Os03g49900.1
Os03g49920.1
Os03g49930.1
Os03g49940.1
Os03g49960.1
Os03g49970.1
Os03g49990.1
Os03g50010.1
Os03g50030.1
Os03g50040.1
Os03g50050.1
Os03g50070.1
Os03g50080.1
Os03g50100.1
Os03g50110.1
Os03g50120.1
Os03g50130.1
Os03g50140.1
Bradi1g10980.1
Bradi1g10990.1
Bradi1g11000.1
Bradi1g11010.1
Bradi1g1020.1
Bradi1g11027.1
Bradi1g11040.1
Bradi1g11050.1
Bradi1g11060.1
Bradi1g11070.1
Bradi1g11080.1
Bradi1g11090.1
Bradi1g11100.1
Bradi1g11110.2
Bradi1g11120.1
Bradi1g11130.1
Bradi1g11140.1
Bradi1g11150.1
Bradi1g11160.1
Bradi1g11170.1
Bradi1g11180.1
Bradi1g11187.2
Bradi1g11200.1
Bradi1g11210.1
Bradi1g11220.1
Bradi1g11230.1
Bradi1g11240.1
Si036986m
Si037776m
Si037112m
Si040377m
Si035418m
Si037532m
Si038762m
Si034330m
Si034655m
[Si040705m]
Si034434m
Si035836m
Si038692m
Si034920m
[Si040530m]
Si035681m
Si039400m
Si040419m
Si034838m
Si039038m
Si036463m
Si037003m
Si037550m
[Si039420m]
Si038547m
Si035193m
Si040446m
Si040616m
Si036166m
Si040427m
Si036875m
[Si040528m]
Si037188m
Si039076m
Si034402m
Si034178m
Si036257m
Si035425m
Si037910m
[Si040795m]
Si038045m
Si038020m
Si038013m
Sb01g010770.1
Sb01g010760.1
Sb01g010750.1
Sb01g010740.1
Sb01g010730.1
Sb01g010720.1
Sb01g010710.1
Sb01g010700.1
Sb01g010690.1
Sb01g010680.1
Sb01g010670.1
Sb01g010660.1
Sb01g010650.1
Sb01g010640.1
Sb01g010630.1
Sb01g010620.1
Sb01g010610.1
Sb01g010600.1
Sb01g010590.1
Sb01g010580.1
Sb01g010570.1
Sb01g010560.1
Sb01g010550.1
Sb01g010540.1
Sb01g010530.1
Sb01g010520.1
Sb01g010510.1
Sb01g010500.1
GRMZM2G153706_T01
GRMZM2G153476_T01
GRMZM2G153444_T01
GRMZM2G143160_T01
GRMZM2G147332_T01
GRMZM2G557296_T01
GRMZM2G062045_T01
GRMZM2G015355_T01
GRMZM2G469224_T01
AC233950.1_FGT002
GRMZM2G704032_T01
GRMZM2G093849_T03
GRMZM2G144744_T01
GRMZM5G819320_T02
GRMZM2G086003_T03
GRMZM2G386856_T01
GRMZM2G072578_T03
GRMZM2G460035_T01
AC230023.2_FGT003
AC230023.2_FGT004
AC230023.2_FGT002
GRMZM2G052701_T01
AC230023.2_FGT006
AC233785.2_FGT007
GRMZM5G857210_T01
GRMZM2G704039_T01
AC233785.2_FGT005
GRMZM2G704042_T01
GRMZM5G842877_T01
AC197835.3_FGT001
GRMZM2G057374_T01
GRMZM2G057394_T02
GRMZM2G041770_T01
GRMZM2G041818_T01
GRMZM2G042040_T01
GRMZM2G042398_T01
GRMZM2G004012_T01
GRMZM2G457881_T01
GRMZM2G157172_T01
GRMZM2G157166_T01
GRMZM2G157147_T01
GRMZM2G090435_T01
AC190734.2_FGT003
GRMZM2G024690_T01
GRMZM2G024973_T01
GRMZM2G063355_T01
GRMZM2G363678_T01
GRMZM2G068984_T01
GRMZM2G068963_T01
GRMZM2G123459_T01
GRMZM2G162443_T01
GRMZM2G162405_T02
GRMZM2G029436_T01
GRMZM2G177885_T01
GRMZM2G177934_T01
7.80
9.67
11.96
7.99
265.16
28.29
28.60
8.30
7.49
9.30
11.49
267.15
(not to scale)

◀**Fig. 4** Comparative analysis of *Rht-1* and surrounding loci in *Triticum aestivum* (*Ta*) relative to the assembled *Poaceae* genomes. **a** Orthologous regions with chromosomal start and end points in Megabases (Mb) are shown as *black rectangles*. The telomere of the long arm is shown as a *filled black circle* and the centromere as a constriction. *Os*, *Oryza sativa*; *Bd*, *Brachypodium distachyon*; *Si*, *Setaria italica*; *Sb*, *Sorghum bicolor*; *Zm*_1, *Zea mays* chromosome 1; *Zm*_5, *Zea mays* chromosome 5. **b** Summary of the *Rht-1* region in the *Ta* BAC sequences relative to the *Poaceae* with sequence length indicated by *long open rectangles*. A sequence gap of unknown length is indicated in the *Ta* 'B' sequence. For the *Poaceae*, chromosomal start and end points are shown in Mb and an *open arrowhead* indicates direction to the long arm telomere. For all sequences, predicted loci and direction of transcription are indicated by *black arrows* and orthologous loci are connected by lines. Genes identified on *Ta* BAC sequences and their orthologs are *boxed*. Intergenic space in *Zm*_5 is 10 × the indicated scale. *Double hash marks* indicate omissions of intergenic sequence with the number of omitted kilobases (kb) indicated. Loci shown in *gray with bracketed names* are contained within the intron of the locus named immediately above. Gene locations and additional data are provided in Online Resource 10

CS aneuploid lines in PCR mixes that contained locus-specific primer pairs. For *Rht-B1* physical mapping, the presence of a locus associated with sterility on 4BS required use of the CS 4B/Ditelo(Dt)4BL heteromorphic line. Using gDNA extracted from progeny of self-pollinated 4B/Dt4BL plants in PCRs, a product was not amplified in approximately 10 % of the reactions. These individuals likely represent Dt4BL homozygotes, as approximately 10 % of 4B/Dt4BL progeny would be expected to be Dt4BL homozygotes (S. Reader, JIC, pers. comm.), thereby indicating that *Rht-B1* is present on the short arm of 4B. The wheat B genome *ZnF* gene was also mapped to 4BS using gDNA from Dt4BL homozygous plants and locus-specific primers. For *Rht-D1*, a locus-specific primer pair was utilized in PCR mixes that contained gDNA of CS, Nulli(N)4A-Tetra(T)4D, N4DT4B, or Dt4DL plants. The predicted products were amplified in CS and N4AT4D, but not in N4DT4B or Dt4DL, indicating that *Rht-D1* is on 4DS. To further delineate the physical location of *Rht-D1*, gDNA extracts of CS deletion lines with breakpoints of 0.53, 0.67 or 0.82 were used in PCRs with an *Rht-D1* specific primer pair. No product was amplified in any instance, indicating *Rht-D1* is distal to breakpoint 0.82 on 4DS.

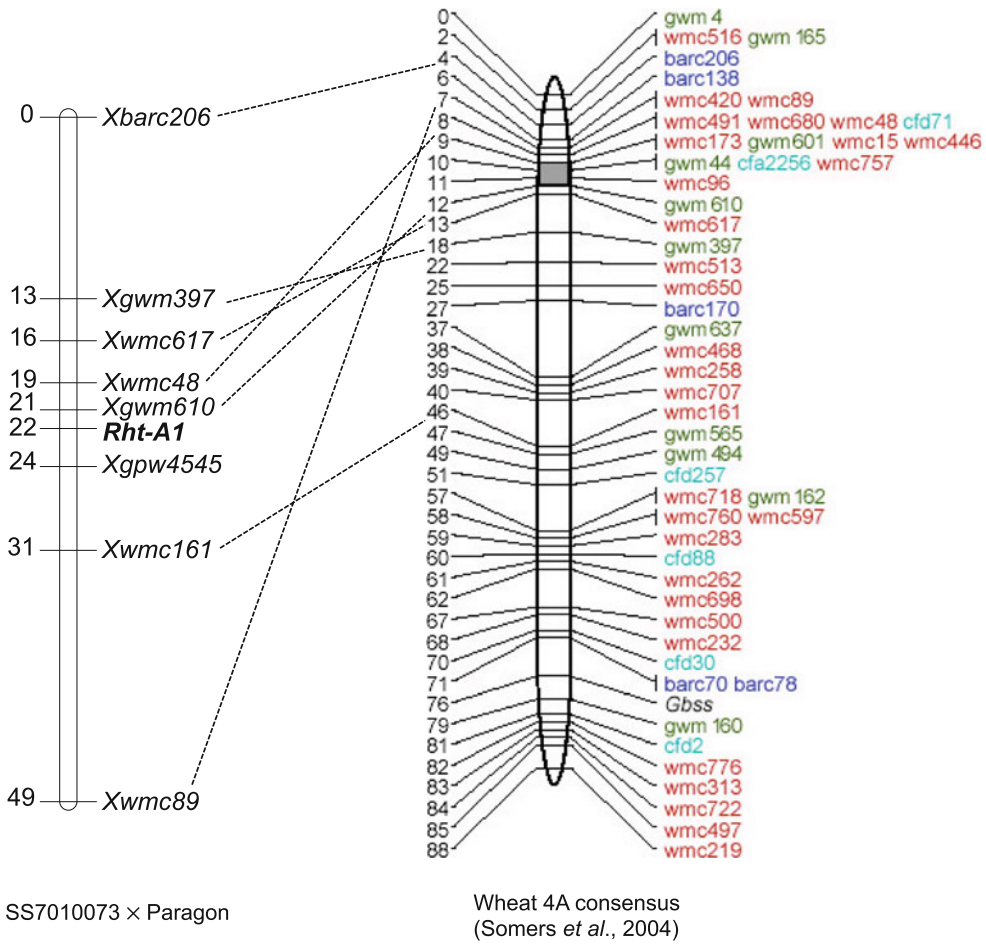


Fig. 5 'SS7010073 × Paragon' F_5 wheat linkage map showing *Rht-A1* and alignment with the wheat 4A consensus map. Marker names are shown to the *right* of the linkage groups with distance in centimorgans indicated to the *left*. The centromere is indicated by a *gray rectangle* on the 4A consensus map. Wheat 4A consensus map reprinted with permission of D. Somers

For *Rht-A1* physical mapping, gDNA of CS, N4AT4D, Dt4AS, and Dt4AL were used in separate PCRs along with an *Rht-A1* specific primer pair. PCR products were amplified in CS and Dt4AL, but not in N4AT4D and Dt4AS indicating that *Rht-A1* is located on the long arm of 4A. An attempt to further delineate the location of *Rht-A1* on 4AL using CS deletion lines was inconclusive due to conflicting results, because a PCR product was amplified using gDNA extracted from 4AL-4 (breakpoint 0.80), 4AL-5 (breakpoint 0.66), and 4AL-12 (breakpoint 0.43), but not from 4AL-13 (breakpoint 0.59). To determine genetic markers linked to *Rht-A1*, a polymorphism was identified between the lines SS7010073 and Paragon that was approximately 150 bp upstream of *Rht-A1*, which enabled a marker to be created. The *Rht-A1* marker and chromosome 4A SSR markers were used to screen 94 individuals of the 'SS7010073 × Paragon' F_5 population. A linkage map was constructed containing *Rht-A1* and eight SSR markers (Fig. 5). Among these SSR markers, three flanking markers (*Xwmc48-4A*, *Xgwm610-4A*, and *Xgpw4545-4A*) were located within 3 cM of *Rht-A1*. On the Somers et al. (2004) consensus map, *Xwmc48-4A* and *Xgwm610-4A* have centromeric locations with *Xwmc48*-4A located on 4AS and *Xgwm610*-4A on 4AL. Overall, the mapping results indicate the most likely physical location of *Rht-A1* is on the long arm of chromosome 4A near the centromere.

Discussion

Analyses of BAC insert sequences from the A, B, and D genomes of CS demonstrated the presence of *Rht-1* along with a *ZnF* and *EamA* gene in each genome. These same three genes were also present on BAC inserts from *T. urartu* (A genome) and, as also previously reported by Duan et al. (2012), the D genome of the near-isogenic line Aibai/CS. Comparative analyses with the assembled *Poaceae* genomes identified an orthologous region in each species in which *ZnF-EamA-Rht-1* synteny remains intact and gene order is highly conserved well beyond the *ZnF-EamA-Rht-1* linkage group. In addition, five CNS regions adjacent to the *Rht-1* ORF were identified in the three wheat genomes and all *Poaceae* examined.

The *ZnF-EamA-Rht-1* linkage block was wholly contained on all BAC inserts examined with the exception of the CS 'B' genome inserts, of which 1417–F16_CS-B contained *EamA* and *Rht-1* and 50-M3_CS-B contained *ZnF*. As no overlap was detected between the two CS 'B' genome BACs, it is not certain that *ZnF-EamA-Rht-1* synteny remains intact; however, the physical mapping of both *Rht-B1* and *ZnF-B* to 4BS and the strong conservation of this linkage block among the *Poaceae* suggest that these three genes retain synteny on the wheat B genome. Based on the locations of *EamA* and *ZnF* on the B genome BACs, these genes are separated by at least 65 kb, which is greater than the interval that exists between these two genes on the bread wheat A (14 kb) or D (34 kb) genomes. The relatively large interval between *EamA* and *ZnF* on the B genome appears to result from a proliferation of TEs, notably WHAM retrotransposon elements, in this region. To fully determine if *ZnF-EamA-Rht-1* gene synteny is retained on the B genome, sequencing and assembly of additional BACs may be required.

The microsynteny of genes surrounding the *ZnF-EamA-Rht-1* linkage group was also well conserved among the assembled *Poaceae* genomes examined (rice, *B. distachyon*, foxtail millet, sorghum, and maize). Of 35 genes identified in the *Slender1* (the rice *Rht-1* ortholog) region of rice, 23 were present in each of the *Poaceae* members (Fig. 4). Similarly, Duan et al. (2012) reported that gene synteny in the *Rht-1* region was well conserved among the wheat D genome, rice, sorghum, and maize chromosome 1. In our study, a single orthologous region containing the 23 genes was present in rice, *B. distachyon*, foxtail millet, and sorghum, whereas two orthologous regions were identified in maize that collectively contained the 23 genes. One orthologous region in maize was on chromosome 1 and contains the *Rht-1* ortholog *Dwarf8* and the second region was identified on chromosome 5 where the *Dwarf9* ortholog is present. The presence of two paralogous *Rht-1* regions in maize is likely the result of a recent (five to twelve MYA) whole-genome duplication (Swigonova et al. 2004). Only 11 of the 23 genes were duplicated on both chromosomes 1 and 5. Duplicates of the remaining genes have apparently been eliminated since the whole-genome duplication and most of these have been lost on chromosome 5. Similarly, Lai et al. (2004) reported that approximately 50 % of the duplicate genes resulting from the polyploidization of maize have been lost. Immediately upstream of the *ZnF* orthologs in the *Poaceae* are orthologs of *Teosinte branched (Tb)1*. *Tb1* was also found to be the next gene upstream of *ZnF* on the D genome of wheat (Duan et al. 2012). *Tb1* is a key domestication gene on maize chromosome 1 that controls branch number (Doebley et al. 2006) and in barley, the *Tb1* ortholog *Intermedium-C* has been associated with lateral spikelet fertility and plant tillering (Ramsay et al. 2011). The close proximity of *Tb1* to *Rht-1* likely means that *Tb1* alleles linked to *Rht-B1b* and *Rht-D1b* have been incorporated into many modern wheat varieties, which could impact plant architecture and wheat productivity.

Gene density in the *Rht-1* region was low in wheat relative to the other *Poaceae* members. In the A, B, and D genome regions covered by the five bread wheat BAC sequences, there is an average of one gene per 94 kb of sequence. In contrast, gene densities in the *Rht-1* regions of

the *Poaceae* examined here are much greater. For the regions defined in Fig. 4, the average number of kb per gene was ranged from 7 to 13 among rice, *B. distachyon*, foxtail millet, and sorghum. For this region in maize, there was an average of one gene per 25 kb on chromosome 5 and one gene per 52 kb on chromosome 1. In a study that examined 63 genes surrounding *Rht-1*, gene density on the D genome of wheat (one gene per 28.6 kb of sequence) was considerably reduced relative to orthologous regions in rice (one gene per 10 kb) and sorghum (one gene per 10.3 kb) (Duan et al. 2012). In another study that utilized five randomly chosen bread wheat BACs, one gene occurred per 75 kb of sequence (Devos et al. 2005), which is similar to the gene density of the bread wheat BACs examined here. Thus, gene density in wheat is reduced relative to the other *Poaceae*, which is most likely the result of the high proportion (80 %) of wheat sequence composed of repetitive elements (Gupta et al. 2008).

In contrast to genes, which were highly conserved among the CS homoeologous BAC inserts, TEs (with the exception of intronic MITEs) were not. High conservation of genic sequence relative to TE sequence was also reported for the *Acetyl-CoA carboxylase* homoeoloci (Chalupska et al. 2008) and *Hardness* homoeoloci (Ragupathy and Cloutier 2008) in bread wheat and for the high molecular weight glutenin (*Glu-1*) loci from *Triticum durum* (A and B genomes) and *A. tauschii* (D genome) (Gu et al. 2004). Intronic MITEs in the *ZnF* and *EamA* genes were partially conserved among the wheat homoeologs. Partial conservation of intronic MITEs was previously reported for the *Glu-1* homoeoloci in wheat (Anderson et al. 2002). The *ZnF* and *EamA* intronic MITEs may be conserved due to their proximity to functional genes, or because MITEs may affect gene function and regulation (Kuang et al. 2009). In contrast to the wheat homoeologous regions, several intergenic TEs are conserved between the CS-A and *T. urartu* BAC sequences. *T. urartu* is the ancestor of the bread wheat A genome with divergence of the two estimated at less than 0.5 MYA, while the A, B, and D genomes have estimated divergence dates of 2–4 MYA (Huang et al. 2002; Dvorak and Akhunov 2005; Chalupska et al. 2008), which likely explains the greater conservation of TEs present between the two A genomes. Similar conservation of TEs between bread wheat genomes and corresponding ancestral lines was also reported at the *Glu-1* loci (Gu et al. 2006).

The *ZnF*, *EamA*, and *Rht-1*, ORFs were well conserved among the wheat homoeologs and *Poaceae* species examined. The resulting proteins of the three genes are expressed in rice (Rice Genome Annotation Project; http://rice.plantbiology.msu.edu) and no stop codons or frameshift mutations that are predicted to cause a loss of function are present in any of the *Poaceae* AA sequences. The nucleotide sequence of CS *Rht-A1* is identical to the *Rht-A1* allele from variety Cadenza (GenBank acc. no. JF930277), which was previously shown to express *Rht-A1* at levels similar to the wild-type *Rht-B1a* and *Rht-D1a* alleles (Pearce et al. 2011). Phylogenetic comparisons of the *Rht-1*, *ZnF*, and *EamA* AA sequences among the *Poaceae* indicate that for each gene, barley and then *B. distachyon* have the closest genetic similarity to wheat. The phylogenies of the *Poaceae Rht-1* orthologs presented here closely resemble the phylogenies shown by Duan et al. (2012) and Smith et al. (2012) when examining similar sets of *Poaceae Rht-1* orthologs. The phylogenies of the *ZnF*, *EamA*, and *Rht-1* ORFs also closely match that of the *Poaceae* genomes (Paterson et al. 2009). These results also support previous work indicating that *B. distachyon* serves as a close genetic model for wheat (International Brachypodium Initiative 2010). In the sequence surrounding *Rht-1*, five CNSs were identified that were common to the three bread wheat homoeologs and to *B. distachyon*, rice, foxtail millet, sorghum, maize (chromosomes 1 and 5), barley, and *E. tef*. The same five CNSs were also identified by Duan et al. (2012) in *A. tauschii*, *B. distachyon*, *B. sylvaticum*, rice, maize (chromosome 1), sorghum, and the D genome of bread wheat. The five CNSs represent potential *Rht-1* regulatory sites as CNSs that have a regulatory role have been previously identified in plants (Uchida et al. 2007). CNS1 of foxtail millet contains a 344 bp insertion, but otherwise the CNS regions are well conserved among the *Poaceae*, giving no clear indication that *Rht-1* regulation is extensively altered.

In the *ZnF-EamA-Rht-1* homoeologous regions, the B genome was the most diverged of the three wheat genomes in CS. In regards to the *Rht-1* flanking regions of the wheat homoeologs, the A and D genomes of bread wheat have a much larger proportion of sequence in common with each other than either have with the B genome flanking sequence. Similarly, the B genome nucleotide sequences of *Rht-1* and *EamA* are the most diverged relative to the A and D genome sequences, while the *ZnF* sequences of all three are nearly identical. In previous studies of wheat homoeologs, the B genome copies were the most divergent in *Knotted1-like homeobox* (Morimoto et al. 2005), *Starch branching enzyme IIa* (Botticella et al. 2012), and *Storage protein activator* (Salse et al. 2008). An analysis of bread wheat ESTs also found the B genome to be more differentiated from the A and D genomes than the A and D genomes are from one another (Akhunov et al. 2003a).

Physical mapping placed *Rht-A1* on the long arm of chromosome 4A, whereas *Rht-B1* and *Rht-D1* were mapped to 4BS and 4DS, respectively. The short arm locations of *Rht-B1* and *Rht-D1* are in agreement with previous mapping using telocentric analyses (McVittie et al. 1978; Izumi et al. 1983). Further mapping of *Rht-D1* using CS

deletion lines places the locus between breakpoint 0.82 and the 4DS telomere. In contrast, *Rht-B1* appears to be proximal to the centromere with telocentric mapping placing it 13 cM from the centromere (McVittie et al. 1978) and genetic mapping indicating a location near the centromere (Borner et al. 1997; Ellis et al. 2002; Somers et al. 2004). The cause of the apparent difference in *Rht-B1* and *Rht-D1* map location is not clear. A pericentric inversion involving chromosome 4B (Mickelson-Young et al. 1995) could account for some of the discrepancy in physical location between these two loci. For *Rht-A1*, the physical location on the long arm relative to the short arm locations of the *Rht-B1* and *Rht-D1* homoeologs is likely the result of an ancient pericentric inversion that resulted in the majority of the native 4AS arm being transferred to 4AL (Miftahudin et al. 2004). Further physical mapping of these loci or assemblies of the group 4 chromosomes in wheat are needed to more accurately determine the physical locations of the *Rht-1* loci. Genetic mapping of *Rht-A1* indicates that it is located within 3 cM of SSR markers *Xwmc48-4A*, *Xgwm610-4A*, and *Xgpw4545-4A*. This presents a valuable marker resource for identifying traits linked with *Rht-A1* or as an aid in backcrossing for wheat improvement. *Xwmc48-4A* and *Xgwm610-4A* were previously mapped near the centromere (Somers et al. 2004), suggesting *Rht-A1* is likely to also be located near the centromere. Loci located near the centromere have lower recombination rates and reduced rates of duplication relative to distal regions (Akhunov et al. 2003b), which suggests that recombination may be reduced around *Rht-B1* and *Rht-A1* relative to *Rht-D1*. A telomeric *Rht-D1* location may partially explain the presence of the *Rht-D1c* dwarf allele, which results from a segmental duplication (Pearce et al. 2011; Li et al. 2012), while no *Rht-A1* or *Rht-B1* alleles resulting from segmental duplication have been reported.

The *Rht-1* homoeoloci have played a key role in improving wheat productivity worldwide, beginning with the green revolution. More recently, the association of *Rht-1* with changes in FHB susceptibility in wheat and the broad association of DELLA proteins with altered tolerances to abiotic and biotic stresses in plants emphasize that the *Rht-1* loci (along with linked loci) remain an important target for further research and breeding efforts. The tools developed herein will enable more precise manipulation of the dwarfing alleles and neighboring loci, particularly through the use of genome-specific markers. This will enable detailed studies of pleiotropy and causative effects of dwarfing genes and linked loci in relation to FHB and other stresses.

Acknowledgments We thank the NIAB trust and the Biotechnology and Biological Sciences Research Council for funding Ed Wilhelm's research as part of a PhD. We thank Paul Bailey (JIC) for providing guidance in phylogenic comparisons and Steve Reader (JIC) for advice regarding the CS aneuploid stocks.

References

Achard P, Cheng H, De Grauwe L, Decat J, Schoutteten H, Moritz T, Van Der Straeten D, Peng JR, Harberd NP (2006) Integration of plant responses to environmentally activated phytohormonal signals. Science 311:91–94

Achard P, Gong F, Cheminant S, Alioua M, Hedden P, Genschik P (2008) The cold-inducible CBF1 factor-dependent signalling pathway modulates the accumulation of the growth-repressing DELLA proteins via its effect on gibberellin metabolism. Plant Cell 20:2117–2129

Akhunov ED, Akhunova AR, Linkiewicz AM, Dubcovsky J, Hummel D, Lazo GR, Chao S, Anderson OD, David J, Qi L, Echalier B, Gill BS, Miftahudin, Gustafson JP, La Rota M, Sorrells ME, Zhang D, Nguyen HT, Kalavacharla V, Hossain K, Kianian SF, Peng J, Lapitan NLV, Wennerlind EJ, Nduati V, Anderson JA, Sidhu D, Gill KS, McGuire PE, Qualset CO, Dvorak J (2003a) Synteny perturbations between wheat homoeologous chromosomes caused by locus duplications and deletions correlate with recombination rates. Proc Natl Acad Sci USA 100:10836–10841

Akhunov ED, Goodyear AW, Geng S, Qi LL, Echalier B, Gill BS, Dvorak J, Miftahudin, Gustafson JP, Lazo G, Chao S, Anderson OD, Linkiewicz AM, Dubcovsky J, La Rota M, Sorrells ME, Zhang D, Nguyen HT, Kalavacharla P, Hossain K, Kianian SF, Peng J, Lapitan N, Gonzalez-Hernandez JL, Anderson JA, Choi DW, Close TJ, Dilbirligi M, Gill KS, Walker-Simmons MK, Steber C, McGuire PE, Qualset CO, Dvorak J (2003b) The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. Genome Res 13:753–763

Allouis S, Moore G, Bellec A, Sharp R, Faivre Rampant P, Mortimer K, Pateyron S, Foote TN, Griffiths S, Caboche M, Chalhoub B (2003) Construction and characterisation of a hexaploid wheat (*Triticum aestivum* L.) BAC library from the reference germplasm 'Chinese Spring'. Cer Res Comm 31:331–338

Anderson OD, Larka L, Christoffers MJ, McCue KF, Gustafson JP (2002) Comparison of orthologous and paralogous DNA flanking the wheat height molecular weight glutenin genes: sequence conservation and divergence, transposon distribution, and matrix-attachment regions. Genome 45:367–3890

Borner A, Roder M, Korzun V (1997) Comparative molecular mapping of GA insensitive *Rht* loci on chromosomes 4B and 4D of common wheat (*Triticum aestivum* L.). Theor Appl Genet 95:1133–1137

Botticella E, Sestili F, Lafiandra D (2012) Characterization of *SBEIIa* homoeologous genes in bread wheat. Mol Genet Genomics 287:515–524

Buchmann JP, Matsumoto T, Stein N, Keller B, Wicker T (2012) Interspecies sequence comparison of *Brachypodium* reveals how transposon activity corrodes genome colinearity. Plant J 71:550–563

Cao W, Somers DJ, Fedak G (2009) A molecular marker closely linked to the region of *Rht-D1c* and *Ms2* genes in common wheat (*Triticum aestivum*). Genome 52:95–99

Chalupska D, Lee HY, Faris JD, Evrard A, Chalhoub B, Haselkorn R, Gornicki P (2008) *Acc* homoeoloci and the evolution of wheat genomes. Proc Natl Acad Sci USA 105:9691–9696

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

Devos KM, Dubcovsky J, Dvorak J, Chinoy CN, Gale MD (1995) Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. Theor Appl Genet 91:282–288

Devos KM, Ma J, Pontaroli AC, Pratt LH, Bennetzen JL (2005) Analysis and mapping of randomly chosen bacterial artificial

chromosome clones from hexaploid bread wheat. Proc Natl Acad Sci USA 102:19243–19248

Devos KM, Costa de Oliveira A, Xu X, Estill JC, Estep M, Jogi A, Morales M, Pinheiro J, San Miguel P, Bennetzen JL (2008) Structure and organization of the wheat genome—the number of genes in the hexaploid wheat genome. In: Appels R, Eastwood R, Lagudah E, Langridge P, Mackay M, McIntyre L, Sharp P (eds) Proceedings of 11th international wheat geneticc symposium. Sydney University Press, Sydney, eO24, pp 1–5

Doebley JF, Gaut BS, Smith DS (2006) The molecular genetics of crop domestication. Cell 127:1309–1321

Draeger R, Gosman N, Steed A, Chandler E, Thomsett M, Srinivasachary, Schondelmaier J, Buerstmayr H, Lemmens M, Schmolke M, Mesterhazy A, Nicholson P (2007) Identification of QTLs for resistance to Fusarium head blight, DON accumulation and associated traits in the winter wheat variety Arina. Theor Appl Genet 115:617–625

Duan J, Wu J, Liu Y, Xiao J, Zhao G, Gu Y, Jia J, Kong X (2012) New *cis*-regulatory elements in the *Rht-D1b* locus region of wheat. Funct Integr Genomics 12:489–500

Dvorak J, Akhunov ED (2005) Tempos of gene locus deletions and duplications and their relationship to recombination rate during diploid and polyploidy evolution in the *Aegilops-Triticum* alliance. Genetics 171:323–332

Ellis MH, Spielmeyer W, Gale KR, Rebetzke GJ, Richards RA (2002) "Perfect" markers for the *Rht-B1b* and *Rht-D1b* dwarfing genes in wheat. Theor Appl Gen 105:1038–1042

Ewing B, Hillier L, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8:175–185

Febrer M, Wilhelm E, Al-Kaff N, Wright J, Powell W, Bevan MW, Boulton MI (2009) Rapid identification of the three homoeologs of the wheat dwarfing gene *Rht* using a novel PCR-based screen of three-dimensional BAC pools. Genome 52:993–1000

Felsenstein J (2004) PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Dept of Genome Sciences, University of Washington, Seattle

Fulton TM, Chunwongse J, Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol Biol Rep 13:207–209

Gale MD, Devos KM (1998) Comparative genetics in the grasses. Proc Natl Acad Sci USA 95:1971–1974

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS (2012) Phytozome: a comparative platform for green plant genomics, Nucl Acids Res 40(D1):D1178-–D1186

Gu YQ, Coleman-Derr D, Kong X, Anderson OD (2004) Rapid genome evolution revealed by comparative sequence analysis of orthologous regions from four *Triticeae* genomes. Plant Physiol 135:459–470

Gu YQ, Salse J, Coleman-Derr D, Dupin A, Crossman C, Lazo GR, Huo N, Belcram H, Ravel C, Charmet G, Charles M, Anderson OD, Chalhoub B (2006) Types and rates of sequence evolution at the high-molecular-weight glutenin locus in hexaploid wheat and its ancestral genomes. Genetics 174:1493–1504

Gupta PK, Mir RR, Mohan A, Kumar J (2008) Wheat genomics: present status and future prospects. Int J Plant Genomics, Art ID 896451. doi:10.1155/2008/896451

Hilton AJ, Jenkinson P, Hollins TW, Parry DW (1999) Relationship between cultivar height and severity of Fusarium ear blight in wheat. Plant Path 48:202–208

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

International Brachypodium Initiative (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature 463:763–768

Izumi N, Sawada S, Sasakuma T (1983) Genetic analysis of dwarfness in *Triticum aestivum* L. cv. 'Ai-bian 1'. Seiken Ziho 31:38–48

Kashkush K, Feldman M, Levy AA (2002) Gene loss, silencing, and activation in a newly synthesized wheat allotetraploid. Genetics 160:1651–1659

Kong X-Y, Gu YQ, You FM, Dubcovsky J, Anderson OD (2004) Dynamics of the evolution of orthologous and paralogous portions of a complex locus region in two genomes of allopolyploid wheat. Plant Mol Biol 54:55–69

Kuang H, Padmanabhan C, Li F, Kamei A, Bhaskar PB, Ouyang S, Jiang J, Buell CR, Baker B (2009) Identification of miniature inverted-repeat transposable elements (MITEs) and biogenesis of their siRNAs in the *Solanaceae*: new functional implications for MITEs. Genome Res 19:42–56

Kumar S, Mohan A, Balyan HS, Gupta PK (2009) Orthology between genomes of *B. distachyon*, wheat and rice. BMC Research Notes 2:93

Lai J, Ma J, Swigonova Z, Ramakrishna W, Linton E, Llaca V, Tanyolac B, Park Y-J, Jeong O-Y, Bennetzen JL, Messing J (2004) Gene loss and movement in the maize genome. Genome Res 14:1924–1931

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947–2948

Li Y, Xiao J, Wu J, Duan J, Liu Y, Ye X, Zhang X, Guo X, Gu Y, Zhang L, Jia J, Kong X (2012) A tandem segmental duplication (TSD) in green revolution gene *Rht-D1b* region underlies plant height variation. New Phytol 196:282–291

McVittie JA, Gale MD, Marshall GA, Westcott B (1978) The intrachromosomal mapping of the Norin 10 and Tom Thumb dwarfing genes. Heredity 40:67–70

Mickelson-Young L, Endo TR, Gill BS (1995) A cytogenetic ladder-map of the wheat homoeologous group-4 chromosomes. Theor Appl Genet 90:1007–1011

Miftahudin, Ross K, Ma X-F, Mahmoud AA, Layton J, Rodriguez M, Chikmawati T, Ramalingam J, Feril O, Pathan MS, Surlan Momirovic G, Kim S, Chema K, Fang P, Haule L, Struxness H, Birkes J, Yaghoubian C, Skinner R, McAllister J, Nguyen V, Qi LL, Gill BS, Linkiewicz AM, Dubcovsky J, Akhunov ED, Dvorak J, Dilbirligi M, Gill KS, Peng JH, Lapitan NLV, Bermudez-Kandianis CE, Sorrells ME, Hossain KG, Kalavacharla V, Kianian SF, Lazo GR, Chao S, Anderson OD, Gonzalez-Hernandez J, Wennerlind EJ, Anderson JA, Choi D-W, Fenton RD, Close TJ, McGuire PE, Qualset CO, Nguyen HT, Gustafson JP (2004) Analysis of EST loci on wheat chromosome group 4. Genetics 168:651–663

Moore G, Devos K, Wang Z, Gale M (1995) Cereal genome evolution: grasses, line up and form a circle. Curr Biol 5:737–739

Moriconi JI, Buet A, Simontacchi M, Santa-Maria GE (2012) Near-isogenic wheat lines carrying altered function alleles of the *Rht-1* genes exhibit differential responses to potassium deprivation. Plant Sci 185–186:199–207

Morimoto R, Kosugi T, Nakamura C, Takumi S (2005) Intragenic diversity and functional conservation of the three homoeologous loci of the *KN1*-type homeobox gene *Wknox1* in common wheat. Plant Mol Biol 57:907–924

Naranjo T, Roca P, Goicoechea PG, Giraldez R (1987) Arm homoeology of wheat and rye chromosomes. Genome 29: 873–882

Navarro L, Bari R, Achard P, Lison P, Nemri A, Harberd NP, Jones JDG (2008) DELLAs control plant immune responses by

modulating the balance and salicylic acid signaling. Curr Biol 18:650–655

Nelson JC, Sorrells ME, Van Deynze AE, Lu YH, Atkinson M, Bernard M, Leroy P, Faris JD, Anderson JA (1995) Molecular mapping of wheat: major genes and rearrangements of homoeologous groups 4, 5, and 7. Genetics 141:721–731

Nicholas KB, Nicholas HB (1997) GeneDoc: a tool for editing and annotation multiple sequence alignments. Distributed by the author (http://www.psc.edu/biomed/genedoc)

Paterson AH, Bowers JE, Feltus FA, Tang H, Lin L, Wang X (2009) Comparative genomics of grasses promises a bountiful harvest. Plant Physiol 149:125–131

Pearce S, Saville R, Vaughan SP, Chandler PM, Wilhelm EP, Sparks CA, Al-Kaff N, Korolev A, Boulton MI, Phillips AL, Hedden P, Nicholson P, Thomas SG (2011) Molecular characterisation of *Rht-1* dwarfing genes in hexaploid wheat. Plant Physiol 157:1820–1831

Peng JR, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F, Sudhakar D, Christou P, Snape JW, Gale MD, Harberd NP (1999) 'Green revolution' genes encode mutant gibberellin response modulators. Nature 400:256–261

Quarrie SA, Steed A, Calestani C, Semikhodskii A, Lebreton C, Chinoy C, Steele N, Pljevljakusic D, Waterman E, Weyen F, Schondelmaier J, Habash DZ, Farmer P, Saker L, Clarkson DT, Abugalieva A, Yessimbekova M, Turuspekov Y, Abugalieva S, Tuberosa R, Sanguineti M-C, Hollintong PA, Aragues R, Royo A, Dodig D (2005) A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring × SQ1 and its use to compare QTLs for grain yield across a range of environments. Theor Appl Genet 110:865–880

Ragupathy R, Cloutier S (2008) Genome organisation and retrotransposon driven molecular evolution of the endosperm *Hardness* (*Ha*) locus in *Triticum aestivum* cv Glenlea. Molec Genet Genomics 280:467–481

Ramsay L, Comadran J, Druka A, Marshall DF, Thomas WT, Macaulay M, MacKenzie K, Simpson C, Fuller J, Bonar N, Hayes PM, Lundqvist U, Franckowiak JD, Close TJ, Muehlbauer GJ, Waugh R (2011) *INTERMEDIUM-C*, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene *TEOSINTE BRANCHED 1*. Nature Genet 43:169–173

Raquin AL, Brabant P, Rhone B, Balfourier F, Leroy P, Goldringer I (2008) Soft selective sweep near a gene that increases plant height in wheat. Mol Ecol 17:741–756

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, Charmet G, Boudet N, Chalhoub 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 Genomics 9:555

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, New York

Saville RJ, Gosman N, Burt CJ, Makepeace J, Steed A, Corbitt M, Chandler E, Brown JKM, Boulton MI, Nicholson P (2012) The 'Green Revolution' dwarfing genes play a role in disease resistance in *Triticum aestivum* and *Hordeum vulgare*. J Exp Bot 63:1271–1283

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nature Biotech 18:233–234

Smith SM, Yuan Y, Doust AN, Bennetzen JL (2012) Haplotype analysis and linkage disequilibrium at five loci in *Eragrostis tef*. Genes Genomes Genetics 2:407–419

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

Srinivasachary, Gosman N, Steed A, Simmonds J, Leverington-Waite M, Wang Y, Snape J, Nicholson P (2008) Susceptibility to Fusarium head blight is associated with the *Rht-D1b* semi-dwarfing allele in wheat. Theor Appl Gen 116:1145–1153

Srinivasachary, Gosman N, Steed A, Hollins TW, Bayles R, Jennings P, Nicholson P (2009) Semi-dwarfing *Rht-B1* and *Rht-D1* loci of wheat differ significantly in their influence on resistance to Fusarium head blight. Theor Appl Genet 118:695–702

Swigonova Z, Lai J, Ma J, Ramakrishna W, Llaca V, Bennetzen JL, Messing J (2004) Close split of sorghum and maize genome progenitors. Genome Res 14:1916–1923

Uchida N, Townsley B, Chung K-H, Sinha N (2007) Regulation of *SHOOT MERISTEMLESS* genes via an upstream-conserved noncoding sequence coordinates leaf development. Proc Natl Acad Sci USA 104:15953–15958

Van Ooijen JW (2006) Software for the calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen

Xu Z, Wang H (2007) LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res 35(Web Server issue):W265–W268

Yan W, Li HB, Cai SB, Ma HX, Rebetzke GJ, Liu CJ (2011) Effects of plant height on type I and type II resistance to Fusarium head blight in wheat. Plant Path 60:506–512