Theor Appl Genet (2015) 128:1617–1629
DOI 10.1007/s00122-015-2536-7

ORIGINAL PAPER

Dynamic evolution of resistance gene analogs in the orthologous genomic regions of powdery mildew resistance gene *MlIW170* in *Triticum dicoccoides* and *Aegilops tauschii*

Yong Liang[1] · De-Yun Zhang[1] · Shuhong Ouyang[1] · Jingzhong Xie[1] · Qiuhong Wu[1] · Zhenzhong Wang[1] · Yu Cui[1] · Ping Lu[1] · Dong Zhang[1] · Zi-Ji Liu[1] · Jie Zhu[1] · Yong-Xing Chen[1] · Yan Zhang[1] · Ming-Cheng Luo[2] · Jan Dvorak[2] · Naxin Huo[3] · Qixin Sun[1] · Yong-Qiang Gu[3] · Zhiyong Liu[1]

Received: 25 February 2015 / Accepted: 2 May 2015 / Published online: 21 May 2015

Abstract

Key message **Rapid evolution of powdery mildew resistance gene *MlIW170* orthologous genomic regions in wheat subgenomes**.

Abstract Wheat is one of the most important staple grain crops in the world and also an excellent model for plant ploidy evolution research with different ploidy levels from diploid to hexaploid. Powdery mildew disease caused by *Blumeria graminis* f.sp. *tritici* can result in significant loss in both grain yield and quality in wheat. In this study, the wheat powdery mildew resistance gene *MlIW170* locus located at the *Triticum dicoccoides* chromosome 2B short arm was further characterized by constructing and sequencing a BAC-based physical map contig covering a 0.3 cM genetic distance region (880 kb) and developing additional markers to delineate the resistance gene within a 0.16 cM genetic interval (372 kb). Comparative analyses of the *T. dicoccoides* 2BS region with the orthologous *Aegilops tauschii* 2DS region showed great gene colinearity, including the structure organization of both types of RGA1/2-like and RPS2-like resistance genes. Comparative analyses with the orthologous regions from *Brachypodium* and rice genomes revealed considerable dynamic evolutionary changes that have re-shaped this *MlIW170* region in the wheat genome, resulting in a high number of non-syntenic genes including resistance-related genes. This result might reflect the rapid evolution in *R*-gene regions. Phylogenetic analysis on these resistance-related gene sequences indicated the duplication of these genes in the *MlIW170* region, occurred before the separation of the wheat B and D genomes.

Communicated by B. Keller.

Y. Liang and D.-Y. Zhang are contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-015-2536-7) contains supplementary material, which is available to authorized users.

✉ Yong-Qiang Gu
yong.gu@ars.usda.gov

✉ Zhiyong Liu
zhiyongliu@cau.edu.cn

[1] State Key Laboratory for Agrobiotechnology/Department of Plant Genetics and Breeding, China Agricultural University, Beijing, China

[2] University of California at Davis, Davis, CA 95616, USA

[3] USDA-ARS West Regional Research Center, Albany, CA 94710, USA

Introduction

Wheat is an excellent model for plant ploidy evolution research because of the presence of its well characterized ancestral species with different ploidy levels from diploid ($2n = 2x = 14$) to hexaploid ($2n = 6x = 42$). Bread wheat is hexaploid with three closely related subgenomes designated A, B, and D which are derived from three diploid species: *Triticum urartu* (AA), an unknown close relative of *Aegilops speltoides* (BB), and *Ae. tauschii* (DD) (Salamini et al. 2002). About 500,000 years ago, the hybridization of A and B genomes resulted in the formation of wild emmer (*T. turgidum* ssp. *dicoccoides*, referred to as *T. dicoccoides* hereafter), which is the ancestor of cultivated emmer (*T. dicoccon*) and *T. durum*. After that the D-genome donor participated in next allopolyploidization event happened 10,000 years ago, forming the modern hexaploid bread wheat (*T. aestivum*, $2n = 42$, AABBDD) as a milestone

in modern human agricultural history (Dubcovsky and Dvorak 2007; Petersen et al. 2006).

The next-generation sequencing technology has made the generation of large sequence data much efficient and cost effective. Genomes of many plant species have been sequenced, including important crops such as rice, sorghum, and maize (Kawahara et al. 2013; Paterson et al. 2009; Schnable et al. 2009). However, the hexaploid wheat genome size is nearly 17-gigabase-pair (Gb) with each subgenome size about 5.5 Gb. Furthermore, the wheat genome is highly repetitive with repeat DNA content approximately 80 % (Brenchley et al. 2012; Eilam et al. 2007). Because of its multiple ploidy level, large genome size, and high repeat DNA content, it still represents a great challenge to generate a high-quality reference genome sequence for bread wheat. To reduce the sequencing and assembly complexity, several strategies have been undertaken in the wheat genome sequencing community. This includes sequencing wheat's diploid ancestor species to leverage the genomics research for crop improvement in bread wheat. For instance, both wheat A-genome progenitor *T. urartu* and D-genome progenitor *Ae. tauschii* have been sequenced using whole-genome shotgun sequencing strategy (Ling et al. 2013; Jia et al. 2013). Another effective strategy adopted by the International Wheat Genome Sequencing Consortium (IWGSC) is to first isolate individual wheat chromosome arm by flow cytometric sorting and then conduct sequencing either using shotgun sequencing, or clone-by-clone sequencing approach, or the combination of both (International Wheat Genome Sequencing Consortium 2014). The wheat chromosome 3B with nearly 1-Gb in size is the first to be sequenced with both clone-by-clone and shotgun sequencing approaches (Choulet et al. 2014).

Bread wheat (*T. aestivum*) is one of the most important staple grain crops worldwide. However, wheat disease causes tremendous yield loss every year and has been one of the major threats that can cause world hunger. Wheat powdery mildew caused by *Blumeria graminis* f. sp. *tritici* (*Bgt*) can both severely reduce grain yield, and change milling and baking quality (Conner et al. 2003; Everts et al. 2001). To combat pathogen attacks, plants have evolved effective defense mechanism. In gene for gene disease resistance pathway, resistance (R) proteins are activated by specific pathogen effectors, resulting in effector-triggered immunity (ETI) (Jones and Dangl 2006). Identification of resistance genes and their use for developing disease-resistant varieties in breeding have been considered as an economical and effective alternative to other measures like pesticides to protect crops from diseases.

One of the major types of plant disease resistance genes encodes nucleotide-binding site leucine-rich repeat proteins (NBS-LRR proteins). The NBS domains contain conserved motifs that bind and hydrolyze ATP and GTP, and the LRR motifs are typically involved in protein–protein interactions (McHale et al. 2006). For plants to adapt to the rapid change of species-specific pathogen spectrum, R genes, as a big gene family, often amplify tandemly or segmentally and tend to be clustered in the genome. Based on the analyses on shotgun genome assemblies, a total of 1219 *Ae. tauschii* genes and 593 *T. urartu* genes encoding R proteins were identified (Jia et al. 2013; Ling et al. 2013). However, little is known about R-gene organization and evolution in a large genomic region, particularly in the context of comparison with other homeologous wheat genomes.

In a previous work, we mapped a powdery mildew resistance gene *MlIW170* originated from *T. dicoccoides* on the short arm of chromosome 2B (Liu et al. 2012). In this paper, we report the following: (1) fine genetic linkage map construction of the *MlIW170* gene using a large segregating population, (2) physical map construction of the *MlIW170* genomic region by chromosome landing and walking, and (3) comparative genomic analyses of the *MlIW170* genomic region on chromosome 2B and the homeologous region of the D genome from *Ae. tauschii* by BAC contig sequencing.

Materials and methods

Plant materials

A powdery mildew-resistant wild emmer (*T. dicoccoides*) accession IW170 (original G-680-M), provided by Drs. T. Fahima and E. Nevo, University of Haifa, Israel, was crossed with a highly susceptible durum line 81086A to construct a large segregating population including 3526 $F_{2:3}$ families for fine genetic linkage map construction.

Chromosome landing and BAC libraries screening

The wild emmer accession TZ-2 (personal collection from Dr. T. Fahima of University of Haifa, Israel) BAC library, representing 3× genome coverage, contains 326,784 BAC clones stored in 851 384-well plates and was arrayed in 175 DNA pools for PCR screening with designed probes. Before the TZ-2 BAC library was available, the *T. turgidum durum* cv. Langdon BAC library (Cenci et al. 2003) was also used for BAC screening. Two flanking markers, SCAR marker *XcauG3* and EST-STS marker *XP2430*, most proximal to *MlIW170*, were selected as start point to screen the BAC libraries for chromosome landing. The positive BACs were sequenced and assembled. New probes at the end of the contigs toward *MlIW170* were designed for next step BAC library screening.

To identify *Ae. tauschii* BACs contigs (Luo et al. 2013) that are orthologous to the *MlIW170* genomic region,

molecular marker sequences mapped to the *MlIW170* genomic region were used to perform blast analysis against the *Ae. tauschii* SNP marker extend sequence database (http://probes.pw.usda.gov/WheatDMarker/). The orthologous BAC contig was identified and the minimum tilling path (MTP) BACs were then selected for sequencing.

BAC sequencing

Purified BAC DNAs of the positive BAC clones from each round of BAC library screening were pooled together to reduce sequence cost. The pooled BAC DNAs were sheared by nitrogen with 30 psi pressure for 40 s. In general, 1 μg sheared DNA was used to construct the sequencing library using KAPA HTP Library Preparation Kit Illumina platforms following the manufacturer's instructions (Illumina, USA). The prepared libraries were sequenced on the Illumina MiSeq machine using Illumina MiSeq Reagent Kit V3.

Twelve MTP BAC clones from *Ae. tauschii* physical contig 1774 (Luo et al. 2013) were sequenced with Roche 454 platform. In this case, these BACs were divided into two pools. Five microgram of pooled BAC DNA was used to prepare the 454 sequencing library using the GS Titanium rapid library following the manufacturer's instructions (Roche Diagnostics, USA). The 454 sequencing rapid libraries were processed using the GS FLX plus Titanium LV emPCR (Lib-L) and GS FLX plus Titanium sequencing (GS FLX+) kits (Roche Diagnostics, USA) according to the manufacturer's instructions.

Sequence assembly and analysis

The Cross_match program (University of Washington, Seattle, WA, USA) was used to remove the vector and *E. coli* sequences. The SoapDenovo (Luo et al. 2012, http://soap.genomics.org.cn/soapdenovo.html), Roche 454 Newbler, and SeqMan from DNAstar (version 5.05) were applied to assemble the sequence data to contigs and scaffolds. Using the program Cross_match, all the sequence reads that flank the cloning sites are being extracted and marked as EndReads. The BAC end information was very helpful in determination of the order and orientation of assembled contigs or scaffolds. In the end, all of the sequences from overlapped BAC clones were merged into a single fasta file and the gap was filled with N.

Both TriAnnot pipeline from URGI (http://wheat-urgi.versailles.inra.fr/Tools/Triannot-Pipeline) (Leroy et al. 2012) and genome annotation pipeline MAKER-P (Campbell et al. 2014) were used for sequence annotation, where sequences were analyzed with various integrated coding region predictions, homology search analysis, and repetitive DNA analysis programs. To validate the predicted structure of putative genes, BLASTn and BLASTx searches of the National Center for Biotechnology Information (NCBI) nr, EST and transcriptome shotgun assembly database (http://www.ncbi.nlm.nil.gov/BLAST/) were performed and the predicted coding sequences were manually edited if necessary. The functions of predicted genes were described according to the best hit of BLASTx. Multiple alignments of nucleotide coding sequences were performed by ClustalW with default options and then the phylogenetic trees were constructed based on the bootstrap neighbor-joining (NJ) method (Saitou and Nei 1987) with a p-distance (Nei and Kumar 2000) model by MEGA v6.0 (Tamura et al. 2013). Recombination events between the resistance gene-like sequences were detected by Recombination Detection Program (RDP) V4.45 (Martin et al. 2010).

Results

Genetic analysis of wheat powdery mildew resistance *MlIW170* in a larger segregating population

In order to construct a high-resolution genetic linkage map for the powdery mildew resistance gene *MlIW170* region, a large segregating population including 3526 $F_{2:3}$ families were developed. The F_3 families segregated 859 homozygous resistant: 1781 segregating: 884 homozygous susceptible ($x^2_{1:2:1} = 2.29$, $P > 0.05$). Molecular markers linked to *MlIW170* were re-genotyped in the population. *Xcau516* and *XcauG3*, which were co-segregated to *MlIW170* in the previous map (Liu et al. 2012), were mapped in the distal portion with genetic distances of 0.2 and 0.09 cM, respectively. *XP2430* was found to be 0.11 cM proximal to *MlIW170*. These markers were chosen as start point of chromosome landing (Fig. 1).

Chromosome lading, walking, and physical map construction

Two positive BAC clones, 162E15 and 979H8, were identified from *T. turgidum durum* cv. Langdon BAC library using marker *XcauG3* as probe. BAC clones 460J18 and 732C12 were identified from wild emmer wheat accession TZ-2 BAC library using probe *XP2430*. After BAC sequencing, new probes were developed using the single copy sequences located at the BAC ends for next step BAC library screening (Table 1). After five rounds of chromosome walking using the same strategy, twelve and five overlap BAC clones distal and proximal to *MlIW170* were identified, respectively, to construct a preliminary BAC physical map with an unknown size gap between the two BAC contigs (Fig. 1).

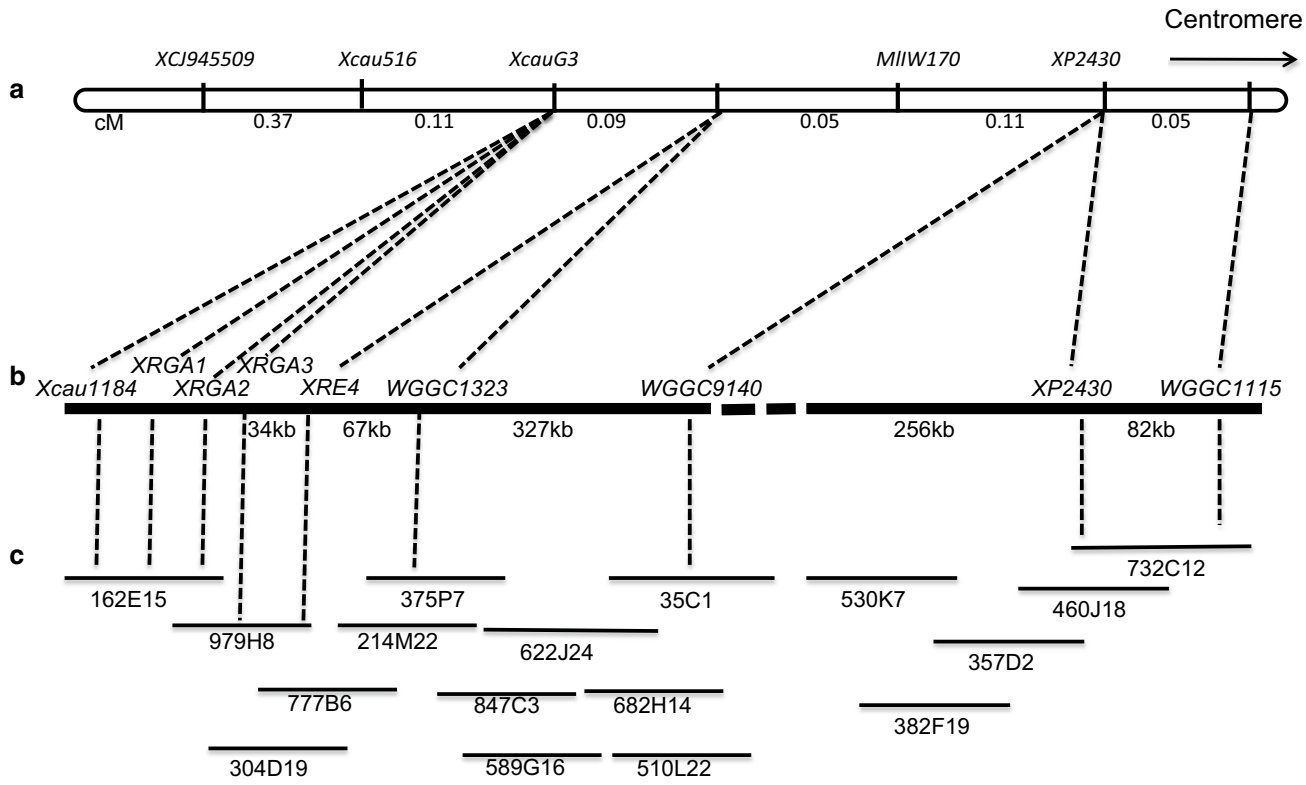

Fig. 1 Genetic linkage map and BAC-based physical map of the powdery mildew resistance locus *MlIW170*. **a** *MlIW170* genetic linkage map on wheat chromosome arm 2BS with genetic distances in cM is shown below. **b** The approximate physical map with the newly developed markers is shown above and the physical distance between markers is shown below. **c** The BAC clone contigs of the distal and proximal regions to *MlIW170* locus. The names of each BAC are shown underneath

Table 1 Probes and positive BAC information for chromosome landing and walking

Probe	Purpose	Forward primer	Reverse primer	Positive BAC clone	Sources
XCAUG3	Landing	5′-GCAGGAGGCGAAACAACT-3′	5′-TAACAACCATTTGTGCCAATG-3′	162E15, 979H8	Langdon
XCAU162	Walking	5′-AGAAAGGTTGGCTGTGGTT-3′	5′-GTCCCTCTGTATCTGTTCCTAT-3′	304D19, 777B6	TZ-2
XP2430	Landing	5′-GTAACATGATGTTGTGGACCTG-3′	5′-GAGGAGATGCAAAACTGAAC-3′	460J18, 732C12	TZ-2
WGGC1336	Walking	5′-ATCCCGTGACTAACTTCT-3′	5′-CTGGGTGGCTGTAGAGGA-3′	214M22, 375P7	TZ-2
WGGC1167	Walking	5′-GCGACGACTGGATTACAT-3′	5′-CCGACTCCCTTCTTACTTAC-3′	357D2	TZ-2
WGGC4380	Walking	5′-CTCCCCCGCTTCTGTAG-TAGTTCT-3′	5′-ATTGTATGTCTGGCTGGGTTT-GAG-3′	847C3, 589G16, 622J24	TZ-2
WGGC5495	Walking	5′-ATACAGGGCCAAGACAATGC-3′	5′-ACATTCCACGTAGGCAGTCC-3′	530K7, 382F19	TZ-2
WGGC6473	Walking	5′-GAGGCAGTACCCCTGCTAGA-3′	5′-TGCCCAAAGTTCTCTTTGCT-3′	35C1, 510L22, 682H14	TZ-2

BAC sequencing and gene annotation of the *MlIW170* region

The 12 and 5 BAC clones distal and proximal to the *MlIW170* locus were sequenced using Illumina MiSeq platform. Assemble of the sequences generated contigs of 590,960 bp (GeneBank accession No. KR021047) and 289,613 bp (GeneBank accession No. KR021048), respectively, in total of 880,573 bp and 21 putative genes with gene density of 23.8 genes/Mb (Table 2). Gene duplications are very common in this region. There are two leucoanthocyanidin reductases (*Td_gene01* and *Td_gene07*), three acetyltransferases (*Td_gene02*, *Td_gene03*, and *Td_gene08*), seven resistance gene-like sequences (*Td_gene04* to *Td_gene06*, *Td_gene14*, and *Td_gene18* to *Td_gene20*), three bidirectional sugar transporters (*Td_gene09*,

Table 2 Gene annotation of the *MlIW170* genomic region in chromosome 2BS of wild emmer (*Triticum dicoccoides*)

Name	Stand	Gene length	Best BLASTx hit to nr database in NCBI	*E* value	Accession number
Td_gene01	−	3779	Leucoanthocyanidin reductase (*Aegilops tauschii*)	2.00E−108	EMT06484.1
Td_gene02	+	1942	Putative acetyltransferase (*Aegilops tauschii*)	4.00E−166	EMT06482.1
Td_gene03	+	1623	Putative acetyltransferase (*Aegilops tauschii*)	5.00E−36	EMT06482.1
Td_gene04	+	4245	Putative disease resistance protein RGA1 (*Aegilops tauschii*)	3.00E−177	EMT09686.1
Td_gene05	+	5553	Putative disease resistance protein RGA1 (*Aegilops tauschii*)	5.00E−121	EMT09686.1
Td_gene06	+	8137	Putative disease resistance protein RGA1 (*Aegilops tauschii*)	5.00E−178	EMT09686.1
Td_gene07	−	1561	Leucoanthocyanidin reductase (*Aegilops tauschii*)	3.00E−123	EMT06484.1
Td_gene08	+	1792	Putative acetyltransferase (*Aegilops tauschii*)	0	EMT06482.1
Td_gene09	+	2994	Bidirectional sugar transporter SWEET6b (*Triticum urartu*)	2.00E−27	EMS55658.1
Td_gene10	–	1259	PREDICTED: sn1-specific diacylglycerol lipase alpha-like (*Setaria italica*)	7.00E−59	XP_004956220.1
Td_gene11	+	2461	Bidirectional sugar transporter SWEET6b (*Triticum urartu*)	2.00E−58	EMS56832.1
Td_gene12	+	2589	Bidirectional sugar transporter SWEET6b (*Triticum urartu*)	2.00E−22	EMS56833.1
Td_gene13	−	1095	Putative WRKY transcription factor 62 (*Triticum urartu*)	1.00E−119	EMS62233.1
Td_gene14	−	4341	Disease resistance protein RPS2 (*Aegilops tauschii*)	2.00E−142	EMT29067.1
Td_gene15	+	9092	Hypothetical protein TRIUR3_03054 (*Triticum urartu*)	9.00E−57	EMS62231.1
Td_gene16	+	4765	Formin-like protein 11 (*Aegilops tauschii*)	3.00E−157	EMT06877.1
Td_gene17	−	1552	Putative WRKY transcription factor 62 (*Aegilops tauschii*)	3.00E−17	EMT15033.1
Td_gene18	−	8185	Disease resistance protein RPS2 (*Aegilops tauschii*)	2.00E−21	EMT29067.1
Td_gene19	−	1434	Disease resistance protein RPS2 (*Aegilops tauschii*)	7.00E−21	EMT29067.1
Td_gene20	−	4026	Disease resistance protein RPS2 (*Aegilops tauschii*)	6.00E−33	EMT29067.1
Td_gene21	−	2955	Formin-like protein 18 (*Triticum urartu*)	1.00E−70	EMS62230.1

Table 3 Markers linked to *MlIW170* developed from the 2BS BAC contig sequence

Markers	Marker type	Forward primer	Reverse primer	Product size (bp)	Dominance
XCAU1184	RJM	TTCAGGTTCTGCGGTCTTCT	GGTACGTCCACCCTGAGATAC	351	Dominant
XRGA1	EST-STS	AGGAGAAGAAGAGTTGGAG	CATCTTTGACTATGTTGTGGC	251	Dominant
XRGA2	EST-STS	TTTGGTCTACCTGCGAGTC	TGAGGCTGGTGAGTTTCTG	510	Dominant
XRGA3	EST-STS	AGGGGTGTAGATGATTCGC	AAGACAGCACGCAATGGA	180	Dominant
XRE4	SSR	AGGCTGGGGAGCTTATATGG	CCTGAACTGCCGAATAGGAA	220/214	Co-dominant
WGGC1323	SSR	TCAAGCATATTTTACACTG	TCCGAAAGAAGAATACCA	217/211	Co-dominant
WGGC9140	SNP(G/T, A/G)	GATATTCCTAGCCTTGTCGTA	GTAGCACCACTGTTGATGA	227	Co-dominant
XP2430	EST-STS	GTAACATGATGTTGTGGACCTG	GAGGAGATGCAAAACTGAAC	700	Dominant
WGGC1115	EST-STS	AAGGGACGGGAGATTGCT	TGCCACCTGTTGATGCTG	800/750	Co-dominant

Td_gene11, and *Td_gene12*), two WRKY transcription factors (*Td_gene13* and *Td_gene17*), and two genes encoding formin-like proteins (*Td_gene16* and *Td_gene21*). Only two genes, *Td_gene10* and *Td_gene15* are single copy.

Marker development and fine genetic linkage map construction of *MlIW170*

The sequence resource generated from BAC clones allowed us to develop more markers to construct a fine genetic map in this region useful for gene cloning and marker-assisted breeding. Based on the sequences generated, EST-STS markers based on the intron flanking sequence, single sequence repeat (SSR) markers, repeat junction markers (RJM) and SNP markers were developed and screened for polymorphism. Finally, nine polymorphic markers were successfully developed and mapped to the *MlIW170* genomic region (Fig. 1; Table 3). *XCAU1184* was a repeat junction marker located in the upstream of *Td_gene01*. Markers *XRGA1*, *XRGA2* and *XRGA3*, *XP2430*, and *WGGC1115* were EST-STS markers derived from *Td_gene04*, *Td_gene05* and *Td_gene06*, *Td-gene18*, and *Td-gene21*, respectively. However, it was found that markers *XRGA1*, *XRGA2,* and *XRGA3*, were co-segregated with *XcauG3* and *Xcau1184* in the large mapping population with 3526 F_2 plants. Two co-segregated SSR

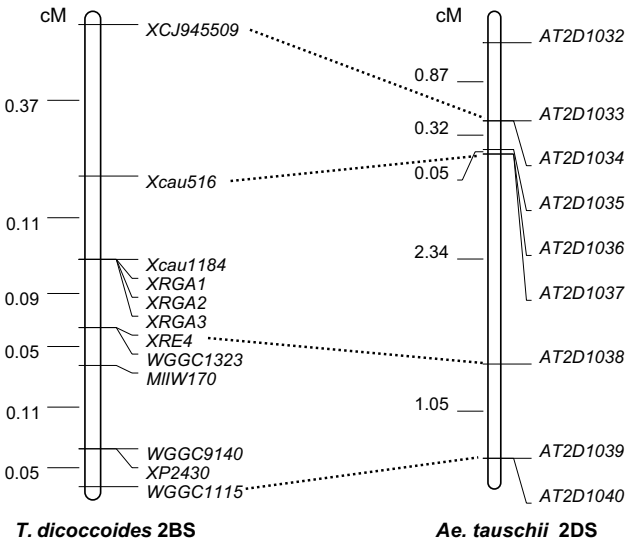

Fig. 2 Fine genetic linkage map of *MlIW170* and its orthologous genetic region of *Ae. tauschii*. Genetic distances in cM are shown on the *left* with genetic markers shown on the *right*. *Dotted lines* indicate the orthologous markers between the two subgenomes

markers, *XRE4* [contains $(TC)_{30}$ between *Td_gene08* and *Td_gene09*] and *WGGC1323* [contains $(CA)_{16}$ between *Td_gene09* and *Td_gene10*] were 67 kb apart (Fig. 1). *WGGC9140* was a SNP marker designed according to a T/G transversion and an A/G transition SNPs in *Td_gene14* where two hybrid peaks in the sequencing track file could be detected in the heterozygote lines (Supplementary Figure 1). It appeared that *WGGC9140* and *XP2430* were co-segregated in the genetic map with at least 256 kb physical distance (Fig. 1). Finally, the powdery mildew resistance gene *MlIW170* could be delimited in a 0.16 cM genetic interval between markers *WGGC1323* and *WGGC9140* spanning 327 kb containing at least five putative genes (*Td_gene09* to *Td_gene13*) (Fig. 1; Table 2).

Identification of *MlIW170* orthologous subgenome region from *Ae. tauschii*

In order to compare the orthologous sub-genomic region difference of *MlIW170* in chromosome 2BS of *T. dicoccoides* and 2DS of *Ae. tauschii*, the sequences of EST-STS markers mapped to the *MlIW170* region were used to blast against the *Ae. tauschii* SNP marker extend contig database (http://probes.pw.usda.gov/WheatDMarker/phpblast/blast.php). *XCJ945509*, *Xcau516*, *XRE4*, and *WGGC1115* were found to be orthologous to *AT2D1034*, *AT2D1036*, *AT2D1038*, and *AT2D1040* anchored in BAC contig 1774 in the *Ae. tauschii* chromosome 2D physical map (Fig. 2; Luo et al. 2013). The 0.20 cM genetic interval of *MlIW170* region between *Xcau516* and *XRE4* at 2BS of *T. dicoccoides* is corresponding to a 2.34 cM genetic interval between *AT2D1036* and *AT2D1038* at 2DS of *Ae. tauschii*. The 0.21 cM genetic region of *MlIW170* between *XRE4* and *WGGC1115* at 2BS of *T. dicoccoides* is corresponding to a 1.05 cM genetic interval between *AT2D1038* and *AT2D1040* at 2DS of *Ae. tauschii*. The *Ae. tauschii* overlapping minimum tiling path (MTP) BACs of this region on 2DS contig 1774 were then selected for sequencing using Roche 454 platform. Finally, an 852,217 bp single scaffold (GeneBank accession No. KR021046) was assembled after manual editing and integration of BAC end sequence information, covering the genomic region from *AT2D1036* to *AT2D1040* of BAC contig1774. Using the same annotation strategy, a total of 26 putative genes (*At_gene01* to *At_gene26*) were identified with gene density of 30.5 gene/Mb (Table 4). Similar situation of gene duplications were found at this 2DS region. There are four phosphoglycerate mutase-like genes (*At_gene01* to *At_gene04*), twelve resistance gene-like sequences (*At_gene05* to *At_gene10*, At_gene13, At_gene14, *At_gene19* to *At_gene22*), three bidirectional sugar transporters (*At_gene15* to *At_gene17*). The other seven genes (*At_gene11*, *At_gene12*, *At_gene18*, *At_gene23* to *At_gene26*) are single copy genes in this region.

Comparative analyses of the orthologous *MlIW170* regions between *T. dicoccoides* 2BS and *Ae. tauschii* 2DS

Comparison of the *MlIW170* orthologous genomic regions on 2BS of *T. dicoccoides* and 2DS of *Ae. tauschii* revealed good micro-colinearity between the two sub-genomic regions. All of the predicted genes from *T. dicoccoides* chromosome 2BS have their orthologs on the orthologous genomic region of *Ae. tauschii* chromosome 2DS, except *Td_gene10*, a sn1-specific diacylglycerol lipase-like gene. The first four genes (*At_gene01* to *At_gene04*) were predicated as phosphoglycerate mutase-like genes on 2DS. The orthologous genomic region of these genes was not covered by the current 2BS physical map. The two leucoanthocyanidin reductase (*Td_gene01* and *Td_gene07*) and three acetyltransferase (*Td_gene02*, *Td_gene03*, and *Td_gene08*) on 2BS are corresponding to single copy genes *At_gene11* and *At_gene12* on 2DS, respectively. In similar situations, the two WRKY transcription factors (*Td_gene13* and *Td_gene17*) and two formin-like proteins (*Td_gene16* and *Td_gene21*) on 2BS are orthologous to single copy genes *At_gene18* and *At_gene24* on 2DS, respectively.

Dynamic duplications and divergence of disease resistance-like sequences were found in the orthologous genomic regions between 2BS and 2DS. Seven and twelve disease resistance-like genes were predicted on 2BS and 2DS, respectively. Phylogenetic analyses revealed that the 19 disease resistance gene analogs found in the 2BS and 2DS regions are classified as three subgroups,

Table 4 Gene annotation of the *MlIW170* orthologous genomic region in chromosome 2DS of *Ae. tauschii*

Name	Stand	Gene length	Best BLASTx hit to nr database in NCBI	*E* value	Accession number
At_gene01	−	5084	Phosphoglycerate mutase-like protein 1 (*Brachypodium distachyon*)	1.00E−82	XP_003580912.1
At_gene02	−	5862	Phosphoglycerate mutase-like protein 1 (*Brachypodium distachyon*)	1.00E−121	XP_003580912.1
At_gene03	−	3257	Phosphoglycerate mutase-like protein 1 (*Brachypodium distachyon*)	2.00E−142	XP_003580912.1
At_gene04	−	4922	Phosphoglycerate mutase-like protein 1 (*Brachypodium distachyon*)	5.00E−136	XP_003580912.1
At_gene05	+	4268	Disease resistance protein RPM1 (*Aegilops tauschii*)	6.00E−10	EMT24451.1
At_gene06	+	5543	Disease resistance protein RGA2 (*Aegilops tauschii*)	0.00E+00	EMT06480.1
At_gene07	+	5299	Disease resistance protein RPM1 (*Triticum urartu*)	0.00E+00	EMS67965.1
At_gene08	−	2673	Disease resistance protein RGA2 (*Aegilops tauschii*)	0.00E+00	EMT06480.1
At_gene09	+	10284	Putative disease resistance protein RGA1 (*Aegilops tauschii*)	0.00E+00	EMT09686.1
At_gene10	−	7115	Putative disease resistance protein RGA1 (*Aegilops tauschii*)	1.00E−62	EMT09686.1
At_gene11	−	1910	Leucoanthocyanidin reductase (*Aegilops tauschii*)	6.00E−74	EMT06484.1
At_gene12	+	4064	Putative acetyltransferase (*Aegilops tauschii*)	0.00E+00	EMT06482.1
At_gene13	−	4519	Disease resistance protein RGA2 (*Aegilops tauschii*)	3.00E−98	EMT06480.1
At_gene14	+	3695	Disease resistance protein RPM1 (*Triticum urartu*)	1.00E−36	EMS59336.1
At_gene15	+	2869	Bidirectional sugar transporter SWEET6b (*Triticum urartu*)	5.00E−37	EMS55658.1
At_gene16	+	2972	Bidirectional sugar transporter SWEET6b (*Triticum urartu*)	6.00E−59	EMS56832.1
At_gene17	+	2486	Bidirectional sugar transporter SWEET6b (*Triticum urartu*)	2.00E−19	EMS56833.1
At_gene18	−	1107	Putative WRKY transcription factor 62 (*Aegilops tauschii*)	2.00E−36	EMT15033.1
At_gene19	−	7179	Disease resistance protein RPS2 (*Aegilops tauschii*)	2.00E−76	EMT29067.1
At_gene20	−	3156	Disease resistance protein RPS2 (*Aegilops tauschii*)	4.00E−50	EMT29067.1
At_gene21	−	6970	Disease resistance protein RPS2 (*Aegilops tauschii*)	0.00E+00	EMT29067.1
At_gene22	−	4202	Disease resistance protein RPS2 (*Aegilops tauschii*)	0.00E+00	EMT29067.1
At_gene23	−	1243	Hypothetical protein TRIUR3_03054 (*Triticum urartu*)	5.00E−25	EMS62231.1
At_gene24	−	6459	PREDICTED: formin-like protein 18 (*Zea mays*)	7.00E−47	XP_008669741.1
At_gene25	+	4024	Mortality factor 4-like protein 1 (*Aegilops tauschii*)	1.00E−09	EMT06876.1
At_gene26	−	3562	PREDICTED: cytochrome P450 93A2-like (*Zea mays*)	1.00E−55	XP_008663013.1

RGA1/2-like, RPS2-like, and RPM-like (Fig. 4). The RGA1/2-like resistance genes include five genes identified from 2DS (*At_gene06*, *At_gene08*, *At_gene09*, *At_gene10*, and *At_gene13*) and three genes from 2BS (*Td_gene04*, *Td_gene05*, and *Td_gene06*). The RPS2-like sub-family contains four orthologous gene pairs between 2BS and 2DS (Fig. 4). The three RPM1-like resistance genes *At_gene05*, *At_gene07*, and *At_gene14* were only found on 2DS but not on 2BS.

Syntenic relationship of the *MlIW170* genomic region with *Brachypodium* and rice

Our previous comparative map demonstrated the existence of macro-colinearity of the wheat *MlIW170* genetic region with the corresponding genomic regions in *Brachypodium* chromosome 5 and rice chromosome 4 (Liu et al. 2012). Here we compared these regions from these genomes at a large stretch of sequence level to understand the micro-colinearity relationship. Compared to six and ten predicted genes in the rice and Brachypodium orthologous genomic regions, 21 and 26 putative genes were identified in the sequenced genomic regions of *T. dicoccoides* 2BS and *Ae. tauschii* 2DS, respectively (Fig. 3). The *Brachypodium* gene *Bd5g02400* and rice gene *Os04g01230* are orthologs of four duplicated genes (*At_gene01* to *At_gene04*) in *Ae. tauschii* 2DS (Fig. 3). Therefore, the orthologs of these genes in 2BS were likely not covered in the current available physical map (Figs. 1, 3). No homologs of the *Brachypodium* genes *Bd5g02440* and *Bd5g02447* and their orthologs in rice were found in the corresponding *T. dicoccoides* and *Ae. tauschii* genomic regions. The *Brachypodium* genes *Bd5g02460*, *Bd5g02470*, and *Bd5g02480* and their orthologousrice genes have orthologous gene pairs (*At_gene24* to *At_gene26*) in *Ae. tauschii* 2DS. However, a segmental inversion was found in *Ae. tauschii* and *T. dicoccoides* (Fig. 3).

Orthologous gene pairs were identified for all of the six predicted rice genes in the corresponding *Brachypodium* genomic region. However, four predicted *Brachypodium* genes (*Bd5g02410*, *Bd5g02420*, *Bg5g02425*, and *Bd5g02430*) have no orthologs in the corresponding rice

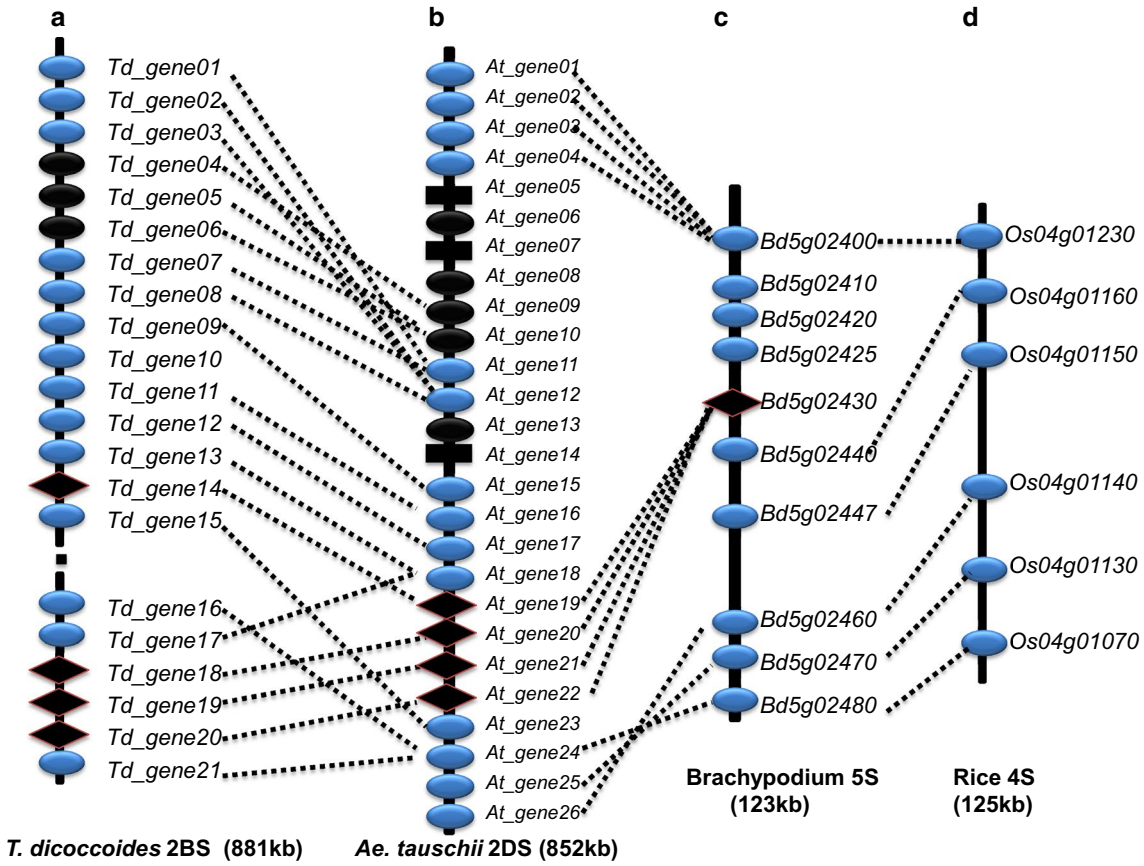

Fig. 3 Comparative genomics mapping of powdery mildew resistance gene *MlIW170*. **a** Sequence map of the *MlIW170* genomic region on *T. dicoccoides* chromosome 2BS with predicted genes shown on the *right*. **b** Sequence map of *MlIW170* orthologous genomic region on *Ae. tauschii* chromosome 2DS with predicted genes shown on the *right*. **c** The *MlIW170* orthologous genomic region on *Brachypodium* chromosome 5S. **d** The *MlIW170* orthologous genomic region on rice chromosome 4S. The resistance gene-like sequences are shown in *dark black color* and the RGA1/2-like, RPM-like, and RPS2-like groups are shown in different shapes of *oval*, *rectangle*, and *rhombus*, respectively

genomic region. Both *Bd5g02420* and *Bd5g02425* encoded retrotransposon proteins and might not represent functional genes in plant. *Bd5g02430* is a RPS2 disease resistance-like protein and *Bd5g02410* is a 552 bp gene fragment which shared 66 % identities at nucleic level and 47 % identities at protein level to *Bd5g02430*. *Bd5g02430* is colinear to four duplicated RPS2-like proteins in *Ae. tauschii* 2DS as well as in *T. dicoccoides* 2BS. However, one of the RPS2-like proteins *Td_gene14* is separated to other three RPS2-like tandem duplication in 2BS of *T. dicoccoides* due to segmental inversion and rearrangement (Fig. 3). Another three and eight resistance gene-like sequences, as well as the remaining genes identified in *T. dicoccoides* 2BS and *Ae. tauschii* 2DS, respectively, have no homolog in the corresponding *Brachypodium* and rice genomic regions, indicating fast evolution of plant disease resistance genes and genomic divergence across the grass genomes (Fig. 3; Tables 2, 4).

Discussion

The fine genetic and physical maps of powdery mildew resistance gene *MlIW170*

MlIW170 is a powdery mildew resistance loci identified previously in a mapping population from a cross between resistance wild emmer accession IW170 and susceptible durum wheat line 81086A (Liu et al. 2012). A high-resolution genetic linkage map and physical map cover the resistant gene locus, which are the major key steps toward the final cloning of the resistance gene. After several rounds of chromosome landing and walking, a 880 kb BAC contig was constructed to cover the *MlIW170* locus in the 2BS chromosome of *T. dicoccoides*. Application of BAC pooling and next-generation sequencing technology significantly improved the efficiency of sequencing and chromosome walking process at reduced cost. Nine additional

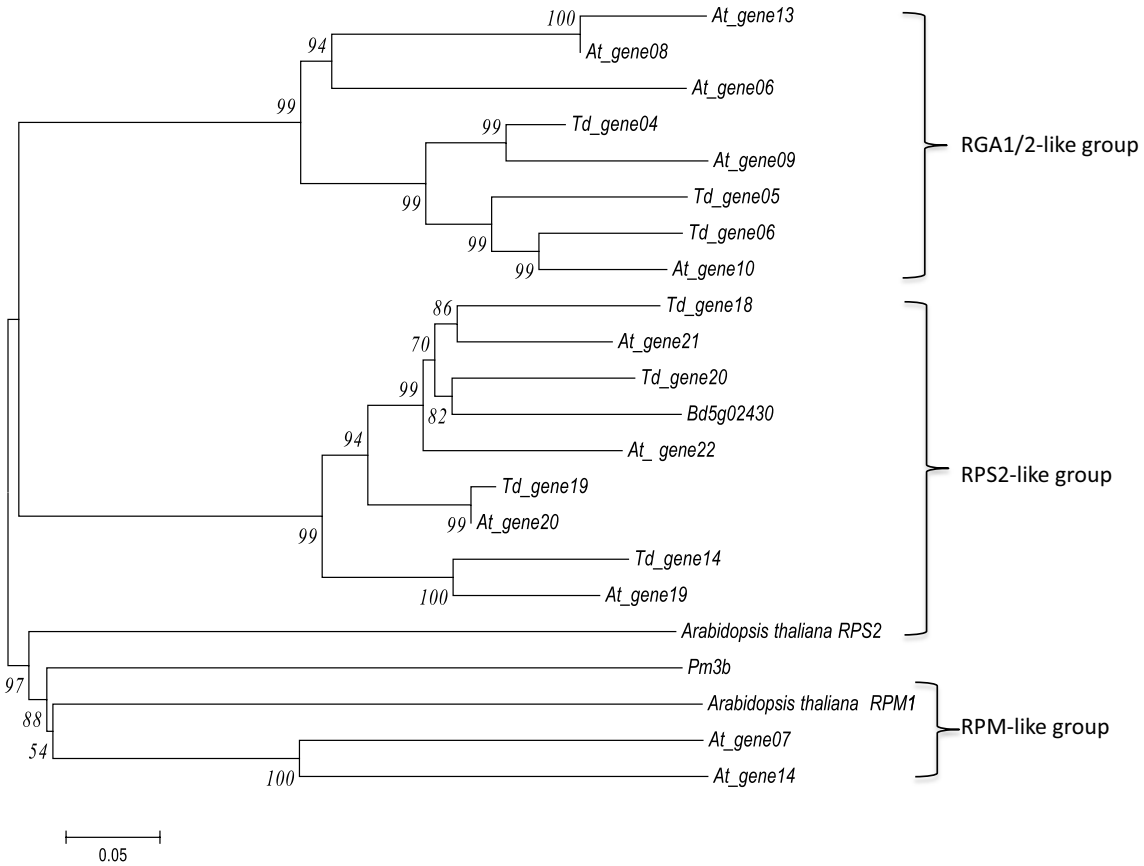

Fig. 4 Phylogenetic tree of the resistance gene-like sequences identified from the 2BS and 2DS genomic regions. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. The wheat powdery mildew resistance gene *Pm3b* (AY325736), *Arabidopsis thaliana* disease resistance genes *RPS2* (NM_118742.2), and *RPM1* (NM_111584.2) were used as outgroup in the evolution analysis. *At_gene05* and *Bd5g02410* are not included in the phylogenetic analyses

markers tightly linked to *MlIW170* were developed from the BAC sequences to construct a fine genetic linkage map of the resistance locus. Finally, the powdery mildew resistance locus is delineated in a 0.16 cM genetic interval spanning 327 kb containing at least five putative genes (Fig. 1; Table 2).

Gene contents and recombination rates comparisons between the 2BS and 2DS homoeologous genomic regions

Sequencing of the 17 overlapping BACs in the *MlIW170* region revealed high number of duplicated genes with high gene density of 23.8 genes/Mb. High gene density of 30.5 gene/Mb was also observed for the orthologous genomic region on 2DS of *Ae. tauschii*. This high gene density might be correlated with their distal bin locations on the chromosome. Similar result was also reported in wheat chromosome 3B with gene density of 19–27.9 genes/Mb in the distal region that is much higher than the average gene density of 9 ± 5 genes/Mb in the whole chromosome (Choulet et al. 2014).

The genetic maps and the sequenced contigs in the *MlIW170* region on chromosome 2BS allowed us to examine and compare the physical to genetic distance between the two homoeologous chromosome regions. A 0.3 cM genetic region between markers *Xcau1184* and *WGGC1115* is corresponding to an 881 kb physical distance on chromosome 2BS of *T. dicoccoides*. Orthologous genomic region was identified in *Ae. tauschii* as a 3.39 cM genetic region between markers *AT2D1036* and *AT2D1040* that corresponds to a 852 kb physical distance. The recombination rates are 0.35 and 3.98 cM/Mb for the 2BS and 2DS genomic regions, respectively. A tenfold higher recombination rate is observed for the chromosome 2DS genomic region than that of the chromosome 2BS genomic region even through they have similar gene contents. It is not clear if the lower recombination rate in the 2B region is

associated with recombination suppression or other mechanisms. However, recombination suppression in resistance loci was observed in wheat *Pm1*/*Lr20*/*Sr15* genetic region on chromosome 7AL (Neu et al. 2002). Distribution of meiotic recombination rates usually showed extreme variations in different genomic regions (Akhunov et al. 2003; Saintenac et al. 2009). The study of wheat chromosome 3B indicated that the average recombination rate along the chromosome ranged from 0 to 2.3 cM/Mb with an average of 0.16 cM/Mb. The average recombinant rate for the *Ae.tauschii* 3D chromosome is 0.32 cM/Mb (Choulet et al. 2014; Luo et al. 2013). The difference of average recombination rates in 3B and 3D could be partially explained by the large physical size of the chromosome 3B in *T. aestivum* cv. Chinese Spring with their similar genetic distances.

Fast evolution and divergence of disease resistance gene-like sequences between the 2BS and 2DS sub-genomic regions

Analyzing the genomic difference of the genomic region flanking wheat resistance gene *MlIW170* on 2BS of *T. dicoccoides* and its homoeologous region on 2DS of *Ae. tauschii* could shed light on the evolution and divergence of disease resistance genes between the two wheat subgenomes. Gene annotations of the two homoeologous genomic regions indicated that a cluster of disease resistance gene analogs exist in both subgenomes. Three types of resistance gene analogs, RGA1/2-like, RPS2-like, and RPM-like, were identified based on blast and sequence alignment analyses. Neighbor-Joining trees were constructed to analyze the evolutionary relationship among the homologous resistance gene analogs. Despite the close relatedness, divergence that distinguishes the two subgenomes is also clearly observed.

In the RGA1/2-like group, *Td_gene04*, *Td_gene05*, and *Td_gene06* from *T. dicoccoides* are three tandem RGA1-like genes, which showed highly similarity with *Ae. tauschii* disease resistance protein RGA1 (Accession No. EMT09686.1). *Td_gene04* and *Td_gene06* are most closely related to *At_gene09* and *At_gene10* from *Ae. tauschii*, suggesting these RGA1-like genes are already present in the progenitor genome before the divergence of the B and D subgenomes. Another three RGA-like genes, *At_gene06*, *At_gene08*, and *At_gene13*, from *Ae. tauschii* showed higher similarity with *Ae. tauschii* disease resistance protein RGA2 (Accession No. EMT06480.1), but not with RGA1 in sequence level. Homolog of the RGA2-like sequences were not found in rice and *Brachypodium* genomes, indicating that they are likely presented only in *Trticeae* and occurred after the separation of *Brachypodium* and wheat lineages. It is likely that these three RGA2-like genes were paralogs and originated from local gene duplications (Fig. 4).

In the RPS2-like group, four copies of RPS2-like genes which showed highly similarity to *Ae. tauschii* RPS2 (Accession No. EMT29067.1) were found in both 2BS and 2DS genomic regions to form four gene pairs (Figs. 3, 4). A four RPS2-like genes tandem repeat cluster (*At_gene19* to *At_gene22*) was identified in the 2DS genomic region, while the RPS2-like tandem repeat cluster in the 2BS genomic region was separated by an insertion containing at least three putative genes between *Td_gene14* and *Td_gene18*. *Td_gene20* is most closely related to *Brachypodium* gene *Bd5g02430*, indicating they are orthologs. It seems that these RPS2-like genes are derived from ancestral *Pooideae* species after divergence from *Oryzoideae*. Local duplications of the RPS2-like sequence could occur after the separation of *Brachypodium* and *Triticeae*. A subgenome rearrangement was presented in the 2BS genomic region after the divergence of B and D subgenomes at 4–5 million years ago (MYA) (Marcussen et al. 2014), resulting the RPS2-like gene *Td_gene14* to be separated from other three RPS2-like genes (Figs. 3, 4).

At_gene05, *At_gene07*, and *At_gene14* in the 2DS genomic region are three RPM-like genes based on the BLASTx results. The predicted *At_gene05* is a short gene fragment and therefore was not included in the evolutionary analyses. *At_gene07* and *At_gene14* are in the same clade and more related to the cloned wheat powdery mildew resistance gene *Pm3b* (Accession No. AY325736) and *Arabidopsis thaliana* disease resistance gene *RPM1* (Accession No. NM111584.2) (Fig. 4). Homolog of the RPM-like sequences was not found in the homoeologous chromosome region on 2BS, suggesting that the RPM-like genes might have an independent evolution process. The results in this study provide further evidence of rapid and dynamic evolution of resistance gene sequences in plant genomes (Bergelson et al. 2001).

Sequence fragments recombination between the disease resistance gene-like sequences was detected using Recombination Detection Program (RDP V4.45; Martin et al. 2010). No recombination event was detected in the RPM and RGA2 subgroups (Supplementary Figure 2). Within the RGA1 subgroup, a 510 bp minor parent fragment, 89 % similarity to *Td_gene06*, was found in recombinant sequence *Td_gene05*. And two minor parent fragments, 1979 bp with 75.9 % similarity to *Td_gene06*, and 567 bp with 88.5 % similarity to *Td_gene05*, were detected in recombinant sequence *At_gene09*. Within the RPS2 gene family, three minor parent fragments, 364 bp with 94.2 % similarity to *At_gene21*, 395 bp with 74.4 % similarity to *At_gene21*, and 627 bp with 85.8 % similarity to *Td_gene20*, respectively, were found in recombinant sequence *At_gene22*. Furthermore, recombination events were also

detected in *Td_gene18* and *At_gene21* (Supplementary Figure 2). No recombination events were observed between subgroups. These results indicated that the recombination events happened within the subgroups before the divergence of B and D genomes.

Comparative analyses of the *MlIW170* orthologous genomic regions within grass family

Besides the large genome size and polyploid nature, the wheat genome also showed much dynamic in regard to evolution events including inversions, deletions, duplications, and other rearrangements that can cause the disruption of gene colinearity among related grass species (wheat, rice, and *Brachypodium*). When performing whole-genome level comparisons between rice, *Brachypodium*, sorghum, and *Ae. tauschii*, the highest structural arrangements (e.g., sequence inversion) were observed in *Ae. tauschii* and the lowest was found in rice (Massa et al. 2011).

The BAC contig sequences in the orthologous genomic regions of powdery mildew resistance gene *MlIW170* on 2BS of *T. dicoccoides* and 2DS of *Ae. tauschii* allowed us to perform a detailed comparison of gene content and gene order among different grass genomes. *Brachypodium* and rice showed highly colinearity in the corresponding genomic regions except the insertion of a RPS2-like gene *Bd5g02430*, a RPS2-like gene fragment *Bd5g02425*, and two putative TE elements *Bd5g02410* and *Bd5g02420* in *Brachypodium* (Fig. 3). In the 2BS and 2DS genomic regions, there are at least 17 putative genes that are not present from the orthologous genomic regions of rice and *Brachypodium* (Fig. 3). Besides the resistance gene analogs, there are other non-syntenic genes with different functions (Table 2).

Another feature for the 2BS and 2DS genomic regions is the presence of high numbers of duplicated genes. Compared with single copy genes *Bd5g02400* and *Bd5g02430* in *Brachypodium*, four homologous genes were identified in the 2DS genomic region for each of them, respectively (Fig. 3), indicating local gene duplications in the *Ae. tauschii* genome. It seems that gene duplications, amplifications, and rearrangements are some of the common features in this region during the evolutionary process of *Triticum* species, resulting in higher number of non-syntenic genes in the 2B and 2D subgenomes. Comparison of the 2BS and 2DS homeologous sub-genomic regions indicated a good colinearity despite the presence of insertion, duplication, and rearrangement events that can differentiate the two subgenomes. Most of predicted genes from the 2BS genomic region can find their orthologs in the corresponding 2DS genomic region. The identification and sequencing of the MTP BAC contig from *Ae. tauschii* greatly facilitated the marker development and fine genetic map construction, as well as the chromosome walking efforts for physical map construction of *MlIW170* in the 2BS genomic region of *T. dicoccoides*.

Several researches indicated that wheat disease resistance genes are located in resistance gene analog (RGA) cluster such as *Pm3b* (Yahiaoui et al. 2004), *Sr33* (Periyannan et al. 2013) and *Lr1* (Cloutier et al. 2007). In barley, the powdery mildew gene *MLA* family was located in adjacent region of multiple classes of genes associated with plant resistance responses (Wei et al. 2002). The wheat stem rust resistance gene *Sr2* region was associated with a 867 kb gene cluster containing ten copies of Germin-like proteins (GLP) (Mago et al. 2014). It seems that powdery mildew resistance gene *MlIW170*, likely to be tightly linked or allelic to another powdery mildew resistance gene *Pm26* (Rong et al. 2000; Liu et al. 2012), also fallen into a resistance gene analog cluster. Seven resistance gene analogs are predicated from the 2BS region BAC contigs flanking *MlIW170* in our study, and the relationship between these RGA and *MlIW170* should be characterized in the future.

Analysis of the 2BS and 2DS orthologous genomic regions revealed the presence of RGA1/2-like, RPS2-like, and RPM-like genes. The RPS2-like genes are likely to be ancestral genes present in both *Triticum* and *Brachypodium*, while the RGA1/2-like and RPM-like genes are newly evolved genes unique to the *Triticum* genome. Phylogenetic analyses of the RGAs identified from 2BS and 2DS genomic regions revealed that the most closely related genes are orthologous gene pairs from homeologous subgenomes, but not the paralogs in the same subgenome, suggesting that these gene duplications likely occurred before the separation of the two subgenomes. Taking together with multiple duplicated copies of disease resistance-like genes and a considerable number of non-syntenic genes, the powdery mildew resistance locus *MlIW170* genomic region has undergone dynamic evolutionary changes in the *T. dicoccoides* genome. Cloning and functional characterization of *MlIW170* in the future will further enhance our understanding of evolutionary and functional divergence of the plant disease resistance genes.

Author contribution statement YL, DYZ, SO, YG, and ZL designed the experiments; YL, DYZ, SO, JX, QW, ZW, YC, DZ, PL, ZJL, JZ, YXC, YZ, and NH performed the experiments; MCL, JD, and QS provided materials; YL, YG, and ZL wrote the paper.

Acknowledgments This work was financially supported by Ministry of Science and Technology of China (2011AA100104, 2013CB127705), Ministry of Agriculture of China (2011ZX08009-003-001), National Natural Science Foundation of China (31030056), the Program of Introducing Talents of Discipline to Universities

(111-02-3), and Monsanto's Beachell-Borlaug International Scholars Program.

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

References

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

Bergelson J, Kreitman M, Stahl EA, Tian DC (2001) Evolutionary dynamics of plant R-genes. Science 292:2281–2285

Brenchley R, Spannagl M, Pfeifer M, Barker GLA, D'Amore R, Allen AM, McKenzie N, Kramer M, Kerhornou A, Bolser D, Kay S, Waite D, Trick M, Bancroft I, Gu Y, Huo N, Luo MC, Sehgal S, Gill B, Kianian S, Anderson O, Kersey P, Dvorak J, McCombie WR, Hall A, Mayer KFX, Edwards KJ, Bevan MW, Hall N (2012) Analysis of the bread wheat genome using whole-genome shotgun sequencing. Nature 491:705–710

Campbell MS, Law M, Holt C, Stein JC, Moghe GD, Hufnagel DE, Lei J, Achawanantakun R, Jiao D, Lawrence CJ (2014) MAKER-P: a tool kit for the rapid creation, management, and quality control of plant genome annotations. Plant Physiol 164:513–524

Cenci A, Chantret N, Kong X, Gu Y, Anderson OD et al (2003) Construction and characterization of a half million clone BAC library of durum wheat (*Triticum turgidum* ssp. *durum*). Theor Appl Genet 107:931–939

Choulet F, Alberti A, Theil S, Glover N, Barbe V, Daron J, Pingault L, Sourdille P, Couloux A, Paux E, Leroy P, Mangenot S, Guilhot N, Le Gouis J, Balfourier F, Alaux M, Jamilloux V, Poulain J, Durand C, Bellec A, Gaspin C, Safar J, Dolezel J, Rogers J, Vandepoele K, Aury JM, Mayer K, Berges H, Quesneville H, Wincker P, Feuillet C (2014) Structural and functional partitioning of bread wheat chromosome 3B. Science 345:1249721

Cloutier S, McCallum BD, Loutre C, Banks TW, Wicker T, Feuillet C, Keller B, Jordan MC (2007) Leaf rust resistance gene *Lr1*, isolated from bread wheat (*Triticum aestivum* L.) is a member of the large *psr567* gene family. Plant Mol Biol 65:93–106

Conner RL, Kuzyk AD, Su H (2003) Impact of powdery mildew on the yield of soft white spring wheat cultivars. Can J Plant Sci 83:725–728

Dubcovsky J, Dvorak J (2007) Genome plasticity a key factor in the success of polyploid wheat under domestication. Science 318:393

Eilam T, Anikster Y, Millet E, Manisterski J, Sagi-Assif O, Feldman M (2007) Genome size and genome evolution in diploid Triticeae species. Genome 50:1029–1037

Everts KL, Leath S, Finney PL (2001) Impact of powdery mildew and leaf rust on milling and baking quality of soft red winter wheat. Plant Dis 85:423–429

International Wheat Genome Sequencing Consortium (2014) A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science 345:1251788

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 KF, 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; International Wheat Genome Sequencing C, 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

Jones JDG, Dangl JL (2006) The plant immune system. Nature 444:323–329

Kawahara Y, de la Bastide M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, Schwartz DC, Tanaka T, Wu JZ, Zhou SG, Childs KL, Davidson RM, Lin HN, Quesada-Ocampo L, Vaillancourt B, Sakai H, Lee SS, Kim J, Numa H, Itoh T, Buell CR, Matsumoto T (2013) Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. Rice 6:4

Leroy P, Guilhot N, Sakai H, Bernard A, Choulet F, Theil S, Reboux S, Amano N, Flutre T, Pelegrin C, Ohyanagi H, Seidel M, Giacomoni F, Reichstadt M, Alaux M, Gicquello E, Legeai F, Cerutti L, Numa H, Tanaka T, Mayer K, Itoh T, Quesneville H, Feuillet C (2012) TriAnnot: a versatile and high performance pipeline for the automated annotation of plant genomes. Front Plant Sci 3:5

Ling HQ, Zhao S, Liu D, Wang J, Sun H, Zhang C, Fan H, Li D, Dong L, Tao Y, Gao C, Wu H, Li Y, Cui Y, Guo X, Zheng S, Wang B, Yu K, Liang Q, Yang W, Lou X, Chen J, Feng M, Jian J, Zhang X, Luo G, Jiang Y, Liu J, Wang Z, Sha Y, Zhang B, Wu H, Tang D, Shen Q, Xue P, Zou S, Wang X, Liu X, Wang F, Yang Y, An X, Dong Z, Zhang K, Zhang X, Luo MC, Dvorak J, Tong Y, Wang J, Yang H, Li Z, Wang D, Zhang A, Wang J (2013) Draft genome of the wheat A-genome progenitor *Triticum urartu*. Nature 496:87–90

Liu Z, Zhu J, Cui Y, Liang Y, Wu H, Song W, Liu Q, Yang T, Sun Q, Liu Z (2012) Identification and comparative mapping of a powdery mildew resistance gene derived from wild emmer (*Triticum turgidum* var. *dicoccoides*) on chromosome 2BS. Theor Appl Genet 124:1041–1049

Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam TW, Wang J (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. GigaScience 1:18

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 PE, Pasternak S, Stein JC, Ware D, Kramer M, McCombie WR, Kianian SF, Martis MM, Mayer KF, Sehgal SK, Li W, Gill BS, Bevan MW, Simkova H, Dolezel J, Weining S, Lazo GR, 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

Mago R, Tabe L, Vautrin S, Simkova H, Kubalakova M, Upadhyaya N, Berges H, Kong X, Breen J, Dolezel J, Appels R, Ellis JG, Spielmeyer W (2014) Major haplotype divergence including multiple Germin-like protein genes, at the wheat *Sr2* adult plant stem rust resistance locus. BMC Plant Biol 14:30

Marcussen T, Sandve SR, Heier L, Spannagl M, Pfeifer M; International Wheat Genome Sequencing C, Jakobsen KS, Wulff BB, Steuernagel B, Mayer KF, Olsen OA (2014) Ancient hybridizations among the ancestral genomes of bread wheat. Science 345:1250092

Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P (2010) RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics 26:2462–2463

Massa AN, Wanjugi H, Deal KR, O'Brien K, You FM, Maiti R, Chan AP, Gu YQ, Luo MC, Anderson OD, Rabinowicz PD, Dvorak J, Devos KM (2011) Gene space dynamics during the evolution of

Aegilops tauschii, *Brachypodium distachyon*, *Oryza sativa*, and *Sorghum bicolor* Genomes. Mol Biol Evol 28:2537–2547

McHale L, Tan XP, Koehl P, Michelmore RW (2006) Plant NBS-LRR proteins: adaptable guards. Genome Biol 7:212

Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, Oxford

Neu C, Stein N, Keller B (2002) Genetic mapping of the *Lr20-Pm1* resistance locus reveals suppressed recombination on chromosome arm 7AL in hexaploid wheat. Genome 45:737–744

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, Schmutz J, Spannagl M, Tang HB, Wang XY, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang Y, Zhang LF, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboob-ur-Rahman Ware D, Westhoff P, Mayer KFX, Messing J, Rokhsar DS (2009) The *Sorghum bicolor* genome and the diversification of grasses. Nature 457:551–556

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

Petersen G, Seberg O, Yde M, Berthelsen K (2006) Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). Mol Phylogenet Evol 39:70–82

Rong J, Millet E, Manisterski J, Feldman M (2000) A new powdery mildew resistance gene: introgression from wild emmer into common wheat and RFLP-based mapping. Euphytica 115:121–126

Saintenac C, Falque M, Martin OC, Paux E, Feuillet C, Sourdille P (2009) Detailed recombination studies along chromosome 3B provide new insights on crossover distribution in wheat (*Triticum aestivum* L.). Genetics 181:393–403

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

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

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, Du F, Kim K, Abbott RM, Cotton M, Levy A, Marchetto P, Ochoa K, Jackson SM, Gillam B, Chen W, Yan L, Higginbotham J, Cardenas M, Waligorski J, Applebaum E, Phelps L, Falcone J, Kanchi K, Thane T, Scimone A, Thane N, Henke J, Wang T, Ruppert J, Shah N, Rotter K, Hodges J, Ingenthron E, Cordes M, Kohlberg S, Sgro J, Delgado B, Mead K, Chinwalla A, Leonard S, Crouse K, Collura K, Kudrna D, Currie J, He R, Angelova A, Rajasekar S, Mueller T, Lomeli R, Scara G, Ko A, Delaney K, Wissotski M, Lopez G, Campos D, Braidotti M, Ashley E, Golser W, Kim H, Lee S, Lin J, Dujmic Z, Kim W, Talag J, Zuccolo A, Fan C, Sebastian A, Kramer M, Spiegel L, Nascimento L, Zutavern T, Miller B, Ambroise C, Muller S, Spooner W, Narechania A, Ren L, Wei S, Kumari S, Faga B, Levy MJ, McMahan L, Van Buren P, Vaughn MW, Ying K, Yeh CT, Emrich SJ, Jia Y, Kalyanaraman A, Hsia AP, Barbazuk WB, Baucom RS, Brutnell TP, Carpita NC, Chaparro C, Chia JM, Deragon JM, Estill JC, Fu Y, Jeddeloh JA, Han Y, Lee H, Li P, Lisch DR, Liu S, Liu Z, Nagel DH, McCann MC, SanMiguel P, Myers AM, Nettleton D, Nguyen J, Penning BW, Ponnala L, Schneider KL, Schwartz DC, Sharma A, Soderlund C, Springer NM, Sun Q, Wang H, Waterman M, Westerman R, Wolfgruber TK, Yang L, Yu Y, Zhang L, Zhou S, Zhu Q, Bennetzen JL, Dawe RK, Jiang J, Jiang N, Presting GG, Wessler SR, Aluru S, Martienssen RA, Clifton SW, McCombie WR, Wing RA, Wilson RK (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326:1112–1115

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729

Wei FS, Wong RA, Wise RP (2002) Genome dynamics and evolution of the *Mla* (powdery mildew) resistance locus in barley. Plant Cell 14:1903–1917

Yahiaoui N, Srichumpa P, Dudler R, Keller B (2004) Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene *Pm3b* from hexaploid wheat. Plant J 37:528–538