

SHORT REPORT

Open Access

The *RAD51* and *DMC1* homoeologous genes of bread wheat: cloning, molecular characterization and expression analysis

Upendra Kumar Devisetty^{1,2}, Katie Mayes¹, Sean Mayes^{1*}

Abstract

Background: Meiotic recombination in eukaryotes requires two homologues of the *E. coli* RecA proteins: Rad51 and Dmc1. Both proteins play important roles in the binding of single stranded DNA, homology search, strand invasion and strand exchange. Meiotic recombination has been well studied in *Arabidopsis*, rice, maize and the orthologues of *RAD51* and *DMC1* have been characterized. However genetic analysis of the *RAD51* and *DMC1* genes in bread wheat has been hampered due to the absence of complete sequence information and because of the existence of multiple copies of each gene in the hexaploid wheat genome.

Findings: In this study we have identified that *TaRAD51* and *TaDMC1* homoeologues are located on group 7 and group 5 chromosomes of hexaploid wheat, respectively. Comparative sequence analysis of cDNA derived from the *TaRAD51* and *TaDMC1* homoeologues revealed limited sequence divergence at both the nucleotide and the amino acid level. Indeed, comparisons between the predicted amino acid sequences of *TaRAD51* and *TaDMC1* and those of other eukaryotes reveal a high degree of evolutionary conservation. Despite the high degree of sequence conservation at the nucleotide level, genome-specific primers for cDNAs of *TaRAD51* and *TaDMC1* were developed to evaluate expression patterns of individual homoeologues during meiosis. QRT-PCR analysis showed that expression of the *TaRAD51* and *TaDMC1* cDNA homoeologues was largely restricted to meiotic tissue, with elevated levels observed during the stages of prophase I when meiotic recombination occurs. All three homoeologues of both strand-exchange proteins (*TaRAD51* and *TaDMC1*) are expressed in wheat.

Conclusions: Bread wheat contains three expressed copies of each of the *TaRAD51* and *TaDMC1* homoeologues. While differences were detected between the three cDNA homoeologues of *TaRAD51* as well as the three homoeologues of *TaDMC1*, it is unlikely that the predicted amino acid substitutions would have an effect on the protein structure, based on our three-dimensional structure prediction analyses. There are differences in the levels of expression of the three homoeologues of *TaRAD51* and *TaDMC1* as determined by QRT-PCR and if these differences are reflected at the protein level, bread wheat may be more dependent upon a particular homoeologue to achieve full fertility than all three equally.

Background

Meiosis is central to all sexually reproducing living organisms and is a highly conserved process. During meiosis a single round of DNA replication is followed by two rounds of cell division resulting in the formation of four daughter cells each having half the chromosome number of the parent. Recombination at meiosis is pivotal not only for

creating genetic diversity through crossing-over but also has a mechanistic role ensuring physical connections in the form of chiasmata that allow the chromosomes to segregate accurately during the first division of meiosis. Bacterial RecA was the first gene to be cloned from *E. coli* that was shown to play an important role in recombination [1,2]. Meiotic recombination in eukaryotic cells generally requires two homologues of the *E. coli* RecA protein - Rad51 and Dmc1 - that were first identified in *Saccharomyces cerevisiae* [3,4]. The Rad51 protein forms a nucleofilament with ATP-dependant strand-exchange

* Correspondence: sean.mayes@nottingham.ac.uk

¹Department of Plant and Crop sciences, School of Biosciences, Sutton Bonington Campus, University of Nottingham, Loughborough LE12 5RD, UK
Full list of author information is available at the end of the article

activity, which has been shown to have roles in meiosis, homologous recombination and DSB repair [5]. *Dmc1* also forms a nucleofilament in cooperation with *Rad51* and is involved in generating hybrid joint molecules specifically in meiosis [3]. The *RAD51* gene exists as a single copy in tomato and *Arabidopsis*, while maize and *Physcomitrella patens* have two copies each [6-11]. The *DMC1* gene sequences have been reported only in a few plant species. While the *Arabidopsis* genome contains one copy of the *DMC1* gene [12], rice has two copies, *OsDMC1A* and *OsDMC1B*, which are located on chromosomes 12 and 11, respectively, most likely produced by chromosomal duplication [13].

Common wheat or bread wheat is an allohexaploid composed of three genomes, each consisting of seven chromosomes (AABBDD; $2n = 6x = 42$). Information on the genes and gene homoeologues of *RAD51* and *DMC1* in bread wheat is limited, primarily due to the added complexity of polyploidy. Given that *RAD51* and *DMC1* gene orthologues are reported to be involved in meiotic recombination in other cereal species and also the model plant *Arabidopsis thaliana*, we have investigated their orthologues in bread wheat. A previous study on meiotic recombination genes in wheat reported the existence of only one copy of *TaDMC1* gene [GenBank accession number EU915561] and two copies of *TaRAD51* gene - *TaRAD51A1* [GenBank accession number EU915557] and *TaRAD51A2* [GenBank accession number EU915558] [14]. We report here for the first time the existence of three homoeologous copies of the *TaRAD51* and *TaDMC1* genes, their genomic location and expression patterns in inflorescence and vegetative tissues in bread wheat. These observations in the Chinese Spring wheat do not completely agree with the results reported earlier [14] and these apparent differences are discussed in this paper.

Methods

Plant materials and DNA isolation

Wild type hexaploid wheat (*Triticum aestivum* c.v Highbury) and Nullisomic-Tetrasomic (NT) wheat stocks (*Triticum aestivum* c.v Chinese Spring) were kindly provided by Prof John Snape and Dr Steve Reader at the John Innes Centre (Norwich, UK). The NT stocks, originally developed by Sears [15], were used for chromosomal assignment of the putative homoeologous genes. Lines Nullisomic for genome A chromosomes were tetrasomics either for B or D chromosomes, lines Nullisomic for genome B chromosomes were tetrasomics either for A or D chromosomes and Nullisomic for genome D chromosomes were tetrasomics either for A or B chromosomes.

Seeds of wheat were first germinated on moistened filter papers in Petri dishes incubated at 20°C in a growth chamber. When coleoptile length had reached ~1 cm,

seedlings were transplanted to 13 cm pots containing John Innes No. 3 compost and grown under standard glasshouse conditions. Total genomic DNA was extracted from young leaf tissue (2 - 4 weeks post emergence) using the GenElute™ Plant Genomic DNA Mini-prep kit (Sigma Aldrich) according to the manufacturer's instructions.

RNA extraction and first strand cDNA synthesis

Total RNA was isolated from the following tissues: root tips (RT, 14 days after germination), mature leaf (L) and immature inflorescence material (I). In addition, meiotically staged anthers were collected for RNA extraction from anther (microspore) tissues. Stages included are: pre-meiosis (PM), leptotene to pachytene (LP), diplotene to anaphase I (DA), telophase I to telophase II (TT), tetrads (T) and immature pollen (IP). Extractions used Trizol reagent (Gibco-BRL) according to the manufacturer's instructions. The isolated RNA was treated with RQ1 RNase-Free DNase (Promega) to remove any residual genomic DNA according to the manufacturers' instructions and first strand cDNA was synthesized using Superscript III Reverse Transcriptase (Invitrogen) using standard methods.

Isolation of full-length cDNA sequences of wheat *RAD51* and *DMC1* genes

Wheat *RAD51* and *DMC1* genes had not been reported at the start of this research, so in order to isolate the wheat *RAD51* cDNA sequence, the rice *OsRAD51A1* cDNA sequence [Gen Bank accession number ABO8O261.2] was retrieved from the GenBank database (<http://www.ncbi.nlm.nih.gov/genbank>) and used to design a total of four primer sets to amplify different sized fragments covering the entire rice cDNA sequence, including 3' and 5' UTR sequences (Table 1). These were then used in RT-PCR using wheat immature inflorescence cDNA as the template. The RT-PCR products were cloned, sequenced and aligned into contigs to obtain the full-length *TaRAD51* cDNA. To isolate the full-length wheat *DMC1* cDNA sequences, primers were designed to the partial *DMC1* cDNA sequence of wheat [GenBank accession number DQ247845] and 3' and 5' RACE was performed. To generate *TaDMC1* 3' ends, mRNA was reverse transcribed using Oligo(dT)₁₈ primer (Table 1). After first strand cDNA synthesis, the 3' region of the *TaDMC1* cDNA was amplified with DGSP1 and the 3' *NotI* anchor primer (Table 1) in the first PCR, and then a nested PCR was done with DGSP2 in combination with the *NotI* anchor primer (Table 1). 5'-RACE to generate *TaDMC1* 5' ends was carried out using a 5'RACE System for Rapid Amplification of cDNA Ends kit (Invitrogen) following the manufacturer's instructions. The PCR products were obtained with AAP or AUAP primers (provided in the kit) and gene

Table 1 List of primers used in the present study

Primer name	Primer sequence (5'-3')	Tm°C	Size (bp)	Use
<i>OsRAD51</i> Exon 1-2-F	CCTCACATCCCAGCATCTC	60	171	Cloning of full-length cDNA of wheat <i>RAD51</i>
<i>OsRAD51</i> Exon 1-2-R	TCACATCAACTGCAGCTATTCCAG			
<i>OsRAD51</i> Exon 8-9-F	GGTAGTGGCCCAAGTGGATG	60	564	
<i>OsRAD51</i> Exon 8-9-R	TGGTGGTCCAATATCACATAGGAG			
<i>OsRAD51</i> Exon 1-7-F	ATGTCGTGGCGGCGG	65	889	
<i>OsRAD51</i> Exon 1-7-R	CTCATCCGCTAATTTCTGAAGGC			
<i>OsRAD51</i> Exon 2-9-F	AAGTCGATAAGATAAT TGAAGCA	60	879	
<i>OsRAD51</i> Exon 2-9-R	CTGAAACCTTGCTTCAGCT			
NotI-(dT) ₁₈	AACTGGAAGAATTCGCGGCCGAGGAA(T) ₁₈			Cloning of full-length cDNA of wheat <i>DMC1</i>
NotI(dT) anchor	AACTGGAAGAATTCGCGG			
<i>TaDMC1</i> -DGSP1	GAAGAGCCTTTCAGGCTTCTG	55	700	
<i>TaDMC1</i> -DGSP2	GGTGAACCTGCAGAGCGCCAG	55	772	
<i>TaDMC1</i> -DGSP3	CTGGCGCTCTGCAAGTTCACC	55	772	
<i>TaDMC1</i> -DGSP4	GGAAGACCCAGTTGGCTCAT	55	461	
<i>TaDMC1</i> -DGSP5	GGTGAACCTGCAGAGCGTCAG	55	207	
<i>TaRAD51</i> Exon 8-F	CCATGATGGTGGAGACAAGG	65	2500	
<i>TaRAD51</i> 3'UTR-R	TGGTGGTCCAATATCACATAGGAG			Amplification of intronic regions in <i>TaRAD51</i> & <i>TaDMC1</i> genes
<i>TaDMC1</i> -5'UTR-F	ATGATCCACATTCCACCCGC	65	2500	
<i>TaDMC1</i> -Exon 6-R	ATGAGCCAACTGGGTCTTCC			
<i>iTaRAD51</i> -7A&7D-F	CGGAAGGATTGGTA AAAAAT	55	450 & 500	
<i>iTaRAD51</i> -7A&7D-R	CACTCAGAAATGACG AAAAAGG			<i>TaRAD51</i> & <i>TaDMC1</i> genome-specific primers based on intronic regions for localization on wheat genome
<i>iTaRAD51</i> -7B-F	ATGGTGGAGACAA GGTGAG	63	500	
<i>iTaRAD51</i> -7B-R	GTACTACTAGCATTAC GGTACAG			
<i>iTaDMC1</i> -5A-F	AGCCTCCGCCCCACTTCCTTC	63	500	
<i>iTaDMC1</i> -5A-R	ACAAACGCAACACGA GCACACG			
<i>iTaDMC1</i> -5B-F	AGCCTTGGCCCCACTTCCTTC	63	500	
<i>iTaDMC1</i> -5B-R	ACGCGCTGCACGCA CCAAA			
<i>iTaDMC1</i> -5D-F	CTCCTCTGACGAG GCGGA	63	389	
<i>iTaDMC1</i> -5D-R	ACGCGCTGCACGC ACCAAA			
<i>TaRAD51</i> -UTR-F	CCTCACATCCCAGCATCTC	63	1528	
<i>TaRAD51</i> -UTR-R	CTCCTATGTGATATTGGACCACCA			<i>TaRAD51</i> cDNA full-length primers including UTR's
<i>eTaRAD51</i> -7A-F	GGGGATACCTCGTGATCAGACT	58	328	
<i>eTaRAD51</i> -R	TGGTGGTCCAATATCACATAGGAG			<i>TaRAD51</i> genome-specific primers based on exons for QRT-PCR analyses
<i>eTaRAD51</i> -7B-F	GATACATCGTGTATCGGACA	58	328	
<i>eTaRAD51</i> -R	TGGTGGTCCAATATCACATAGGAG			
<i>eTaRAD51</i> -7D-F	TGGGGATACATCGTGTATTGGCCT	65	327	
<i>eTaRAD51</i> -R	TGGTGGTCCAATATCACATAGGAG			
<i>TaDMC1</i> -UTR-F	ATGATCCACATTCCACCCGC	63	1268	
<i>TaDMC1</i> -UTR-R	CACTGCAGAAAAGAAATTGGGCAAC			<i>TaDMC1</i> cDNA full-length primers including UTR's
<i>eTaDMC1</i> -5A-F	AGCCTCCGCCCCACTTCCTTC	65	150	
<i>eTaDMC1</i> -R	CTTGTCGATGGACTCGAAGCACTC			<i>TaDMC1</i> genome-specific primers based on exons QRT-PCR analyses
<i>eTaDMC1</i> -5B-F	AGCCTTGGCCCCACTTC CTC	65	150	
<i>eTaDMC1</i> -R	CTTGTCGATGGACTCGAAGCACTC			
<i>eTaDMC1</i> -5D-F	TTCTCCTCCAGCAGCAGCGGAA	65	150	
<i>eTaDMC1</i> -R	CTTGTCGATGGACTCGAAGCACTC			
TUB-F	TCTTCATGGTGGGCTTCGC	55	475	
TUB-R	CGCCTCGGGTGAACCTCATCT			QRT-PCR control primers

specific primers (Table 1). The RT-PCR products were cloned, sequenced and aligned into contigs to obtain the full-length *TaDMC1* cDNA.

Generation of Genome-specific primers for *TaRAD51* and *TaDMC1* using intron sequences

OsRAD51A1 [GenBank accession number ABO8O261] and *OsDMC1A* [GenBank accession number AB079873] cDNA sequences were used to predict the exon/intron boundaries of *TaRAD51* and *TaDMC1*, respectively. Once the boundaries were identified, exon-anchored primers that were expected to amplify products spanning two or three introns were designed (Table 1). Multiple alignment of sequences obtained from the *TaRAD51* and *TaDMC1* cDNA exon-anchored primers on Chinese Spring genomic DNA identified three distinct groups of sequences for each gene with respect to the intronic sequences in bread wheat (data not shown). Putative primers that were expected to amplify a particular homoeologue were designed based on the sequence polymorphisms detected and were coupled with a conserved exon-anchored primer in a PCR reaction containing genomic DNA from Chinese Spring and the appropriate nulli-tetrasomic stocks as templates.

Generation of Genome-specific primers for *TaRAD51* and *TaDMC1* sequences at the exon level

Highbury spikelet cDNA was amplified using primers that generate full-length copies of the *TaRAD51* and *TaDMC1* genes. A major amplification band, approximately 1.5 kb was obtained with both sets of primers. In the absence of differentially sized fragments, genome-specific primers for *TaRAD51* and *TaDMC1* cDNA homoeologues were designed based on sequence polymorphisms identified at the 3' UTRs for *TaRAD51* cDNA homoeologues and at the 5' UTRs for the *TaDMC1* cDNA homoeologues and were paired with a conserved gene primer (Table 1). The genome specificity of the primers was tested using appropriate nulli-tetrasomic stocks as templates.

Comparative amino acid sequence analysis and 3D protein modelling

Comparative amino acid sequence analysis was undertaken using the ClustalW tool of European Bioinformatics Institute (<http://www.ebi.ac.uk/Tools/clustalw2>). The amino acid sequences were processed using the Swiss model workspace (version 8.05, 2009) [16]. The computed outputs were then used to predict the 3D protein structures using Swiss-Pdb Viewer (DeepView; Version 4.0, 2009) [17].

SIFT (Sorting Intolerant From Tolerant) analysis

The potential impact of amino-acid changes on the final protein structure were assessed with the SIFT program

[18]. SIFT is a program which predicts whether an amino acid substitution in the protein affects its protein function or not. SIFT assumes that important amino acids will be conserved in the protein family and so changes at well-conserved positions tend to be predicted as deleterious. SIFT score ranges from 0 to 1. The amino acid substitution is predicted deleterious if the score is ≤ 0.05 and tolerated if the score is > 0.05 . This approach can only give an indication of potential changes to function through mutation.

Phylogenetic analysis

The alignment of nucleotide sequences and subsequent selection of best-fit molecular evolution model was performed using Molecular Evolutionary Genetics Analysis (MEGA) software (version 5.0) [19]. The construction of a phylogenetic tree was performed under maximum likelihood (ML) using RAxML version 7.2.6 [20] on the CIPRES portal [21] with the following parameters: nucleotide search using the GTR + G model, random seed for initial parsimony inference, rapid bootstrapping with automatic bootstopping function (frequency criterion) followed by the final ML search.

Quantitative Real-Time PCR

Harvesting and staging of meiotic anthers for QRT-PCR was performed from immature wheat spikes [22]. Harvested material was immediately frozen and stored at -80°C . Eight independent QRT-PCR amplifications were made for the following tissues: PM, LP, DA, TT, T, IP, L and RT. QRT-PCR was performed in triplicates using wheat Tubulin (Table 1) as an internal standard in a total volume of 12 μl reaction that contained 6 μl of Brilliant[®] SYBR[®] Green (Qiagen), 50 ng cDNA and 1 μl of an equal mix of forward and reverse gene specific primers (Table 1). The reaction conditions for QRT-PCR were as follows: 15 min at 95°C followed by 45 cycles of 30 s at 95°C , 30 s at gene specific annealing temperature and 30 s at 72°C . Primer sequences and annealing temperatures are given in Table 1.

Q-PCR amplification reactions were performed in the LightCycler[®] 480 instrument (Roche Applied Science, UK) using Brilliant[®] SYBR[®] Green (Qiagen) according to the manufacturer's instructions.

Accession numbers

The cDNA sequences of the *TaRAD51* and *TaDMC1* homoeologues were deposited in GenBank at NCBI [GenBank accession numbers FJ594479, FJ594480 and FJ594481 for *TacRAD51-7A*, *TacRAD51-7B* and *TacRAD51-7D* homoeologues, respectively and FJ594476, FJ594477 and FJ594478 for *TacDMC1-5A*, *TacDMC1-5B* and *TacDMC1-5D* homoeologues, respectively].

Results and Discussion

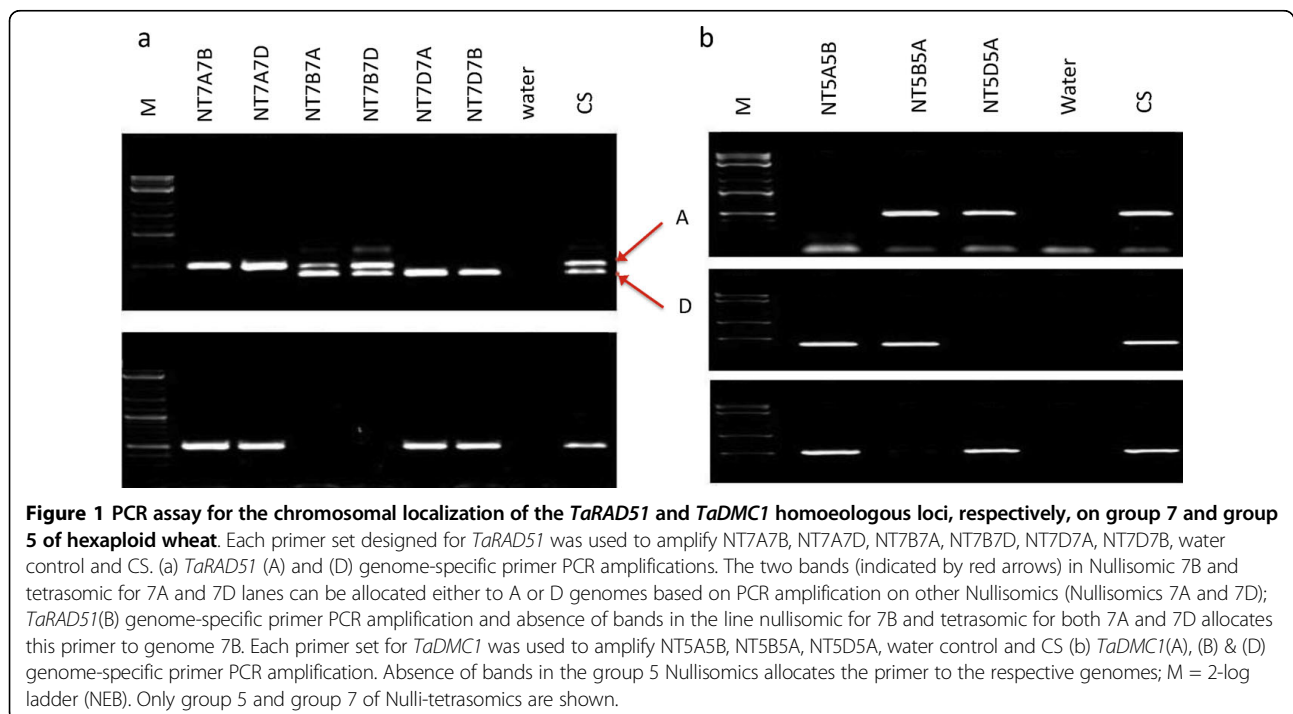
Isolation and genome localization of *RAD51* and *DMC1* gene homoeologous sequences of hexaploid wheat

'Highbury' wheat spikelet cDNA amplifications resulted in the isolation of the full-length *RAD51* and *DMC1* cDNA sequences from hexaploid wheat (see methods). In order to localize the isolated *RAD51* and *DMC1* cDNA sequences on hexaploid wheat, genome specific primers were developed in intronic regions and the genome specificity was confirmed using nulli-tetrasomic analysis. The analysis revealed that the *TaRAD51* and *TaDMC1* sequences were located on group 7 and group 5 chromosomes, respectively, with a copy in each of the A, B and D genomes (Figure 1). Exon specific genome specific primers were developed based on sequence polymorphisms identified (Figure 2) and the genome specificity was confirmed using group 7 and group 5 nulli-tetrasomics of wheat (Figure 3). Later full-length cDNA sequences of the *TaRAD51* and *TaDMC1* homoeologues were isolated using these primers and the sequences were submitted to NCBI database. The locations of *TaRAD51* homoeologues cloned here contrast with a previous report [14] in bread wheat that *TaRAD51* exists as two paralogues - *TaRAD51A1* and *TaRAD51A2* on group 7.

TaRAD51 and *TaDMC1* cDNA homoeologues are highly conserved at the amino acid level

Sequence alignment of the *TaRAD51* cDNA homoeologues at nucleotide and amino acid level revealed 97% and 99% identity respectively. Two of the *TaRAD51* cDNA

homoeologues (*TaRAD51-7A* and *-7B*) have a continuous ORF of 1032 bp, capable of encoding a protein of 343 amino acids and the *TaRAD51-7D* ORF has a deletion of three nucleotides at the N-terminal end, thus putatively encoding a predicted protein of 342 amino acids residues. Further analysis of the three translated *TaRAD51* cDNA homoeologues revealed that this leads to the expected deletion of one amino acid at position 17 corresponding to an 'E' residue (Glutamic acid) and single amino acid substitutions are also expected between the three homoeologues at positions 4 (A, compared with B&D), 31 (A&B, compared with D) and 115 (A&D, compared with B) (Figure 4a). SIFT predictions suggested that the amino acid substitutions are not expected to affect protein 3D structure for the three cDNA homoeologues of *TaRAD51* (Figure 4b). We have also found that the previously reported [14] *TaRAD51A1* is identical to the *TaRAD51-7D* homoeologue isolated here and that the reported *TaRAD51A2* sequence appears to be a truncated version of the *TaRAD51-7A* homoeologue albeit with a few amino acid differences at the truncated end (listed in Additional file 1). However the paralogous nature of the reported *TaRAD51A1/A2* [14] was not observed in this study. This supports the idea that there is only one copy of *TaRAD51* per haploid wheat genome, at least for the wheat genotype studied here. The predicted 3D homoeologue overlays superimposed onto each other *TaRad51* revealed there is a high level of predicted conservation for secondary and tertiary structures (Figure 4c). The only noticeable structural dissimilarity observed between the three cDNA



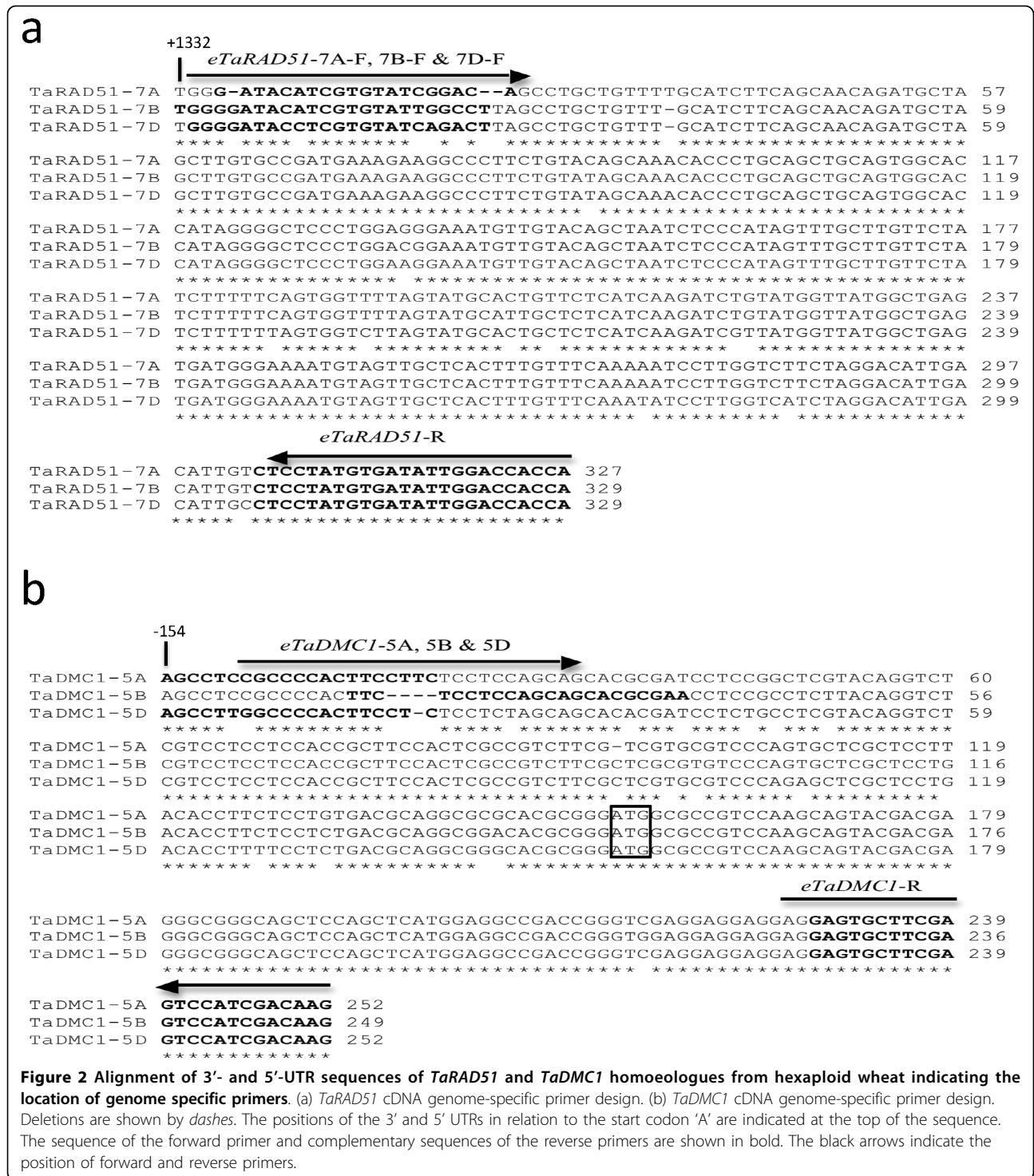


Figure 2 Alignment of 3'- and 5'-UTR sequences of *TaRAD51* and *TaDMC1* homoeologues from hexaploid wheat indicating the location of genome specific primers. (a) *TaRAD51* cDNA genome-specific primer design. (b) *TaDMC1* cDNA genome-specific primer design. Deletions are shown by dashes. The positions of the 3' and 5' UTRs in relation to the start codon 'A' are indicated at the top of the sequence. The sequence of the forward primer and complementary sequences of the reverse primers are shown in bold. The black arrows indicate the position of forward and reverse primers.

homoeologues is in peptide loops seen between α -helix 13 and α -helix 14 (indicated by the white arrow in Figure 4c)
 Conceptual translation and subsequent sequence alignment of *TaDMC1* showed 97% nucleotide sequence identity and 98% identity at the amino acid level between the three translated homoeologous copies. All

cDNA homoeologues of the *TaDMC1* have a continuous ORF of 1035 bp which encodes a predicted protein of 344 amino acids. Further analysis of the three translated *TaDMC1* cDNA homoeologues revealed that the only differences identified between the three cDNA homoeologues were due to nucleotide differences

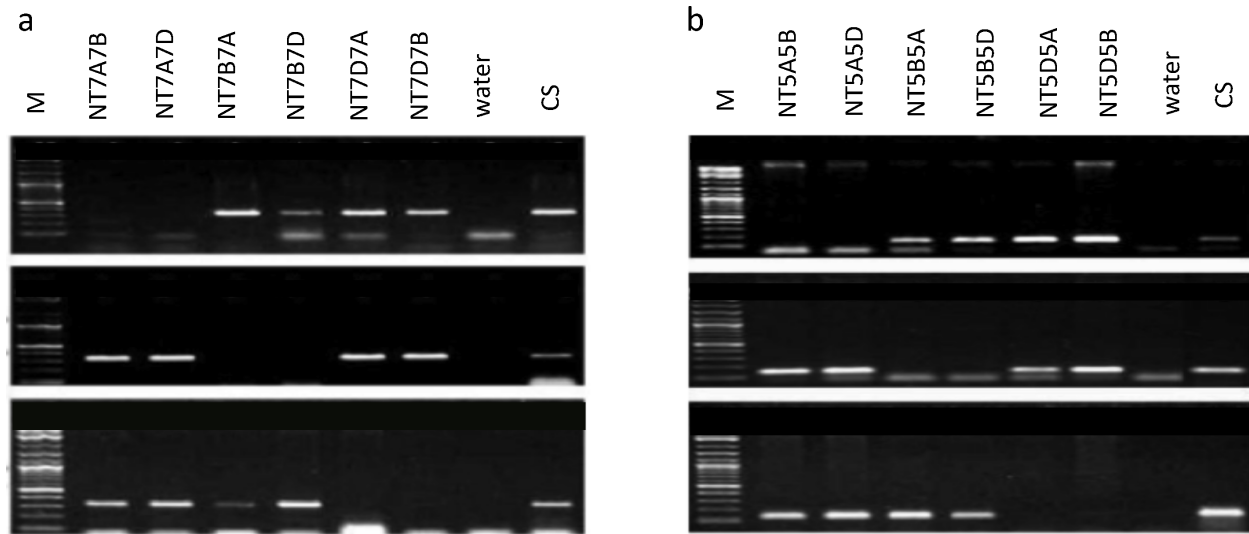


Figure 3 PCR assay for genome specificity of *TaRAD51* and *TaDMC1* cDNA homoeologues in bread wheat. Each primer set for *TaRAD51* was used to amplify NT7A7B, NT7A7D, NT7B7A, NT7B7D, NT7D7A, NT7D7B, water control and CS (a) *TaRAD51*(A), (B) & (D) genome-specific primer PCR amplification. Each primer set for *TaDMC1* was used to amplify NT5A5B, NT5A5D, NT5B5A, NT5B5D, NT5D5A, NT5D5B, water control and CS.(b) *TaDMC1*(A), (B) & (D) genome-specific primer PCR amplification. Absence of bands in the group 7 & 5 Nullisomics allocates the primer to the respective genomes; M = 2-log ladder (NEB).

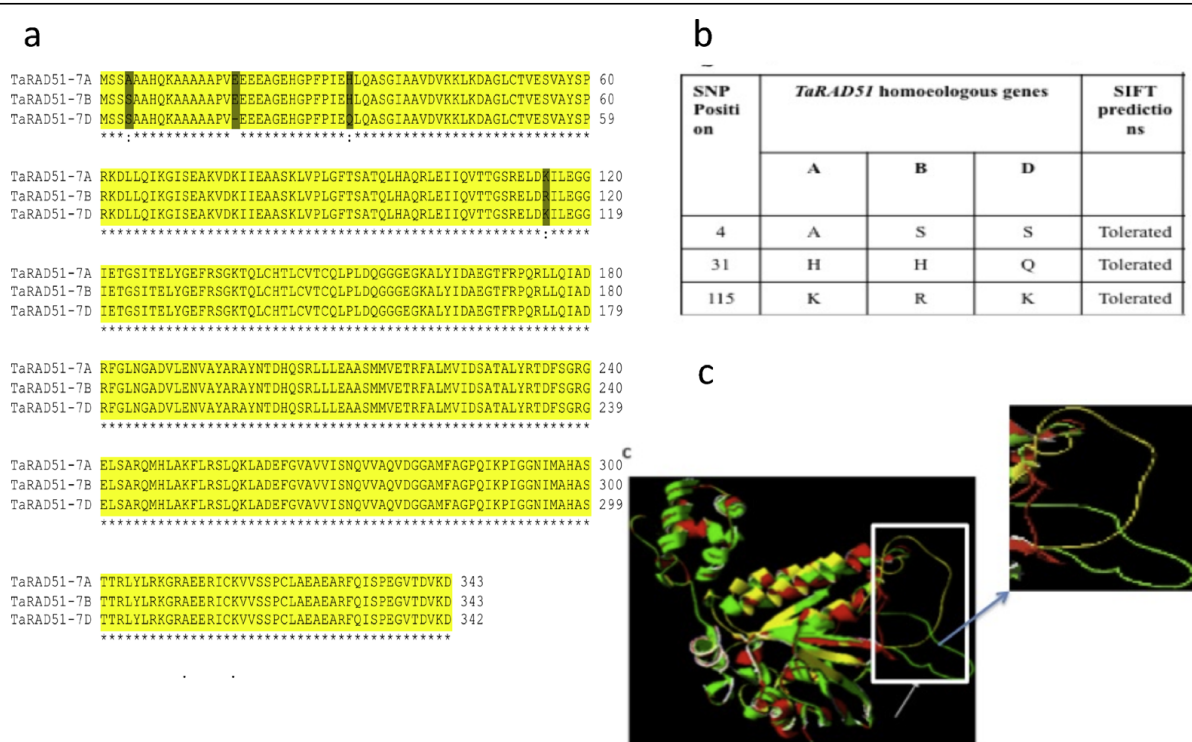


Figure 4 The deduced amino acids alignments of *TaRAD51* homoeologous proteins and their 3D models. (a) Multiple alignments of the three *TaRAD51* sequences identified in the three bread wheat genomes (A, B and D). Conserved amino acids are indicated by black with a yellow background. The amino acid differences between the three cDNA homoeologous proteins are indicated by black with a grey background. (b) SIFT predictions for the amino acid substitutions for the three cDNA homoeologues of *TaRAD51*. (c) The 3D structure of TaRad51-7A is represented in red, TaRad51-7B in green and TaRad51-7D in yellow. Blue arrow shows the magnified image of the side chains of three TaRad51 homoeologue proteins, which indicates the only structural dissimilarity.

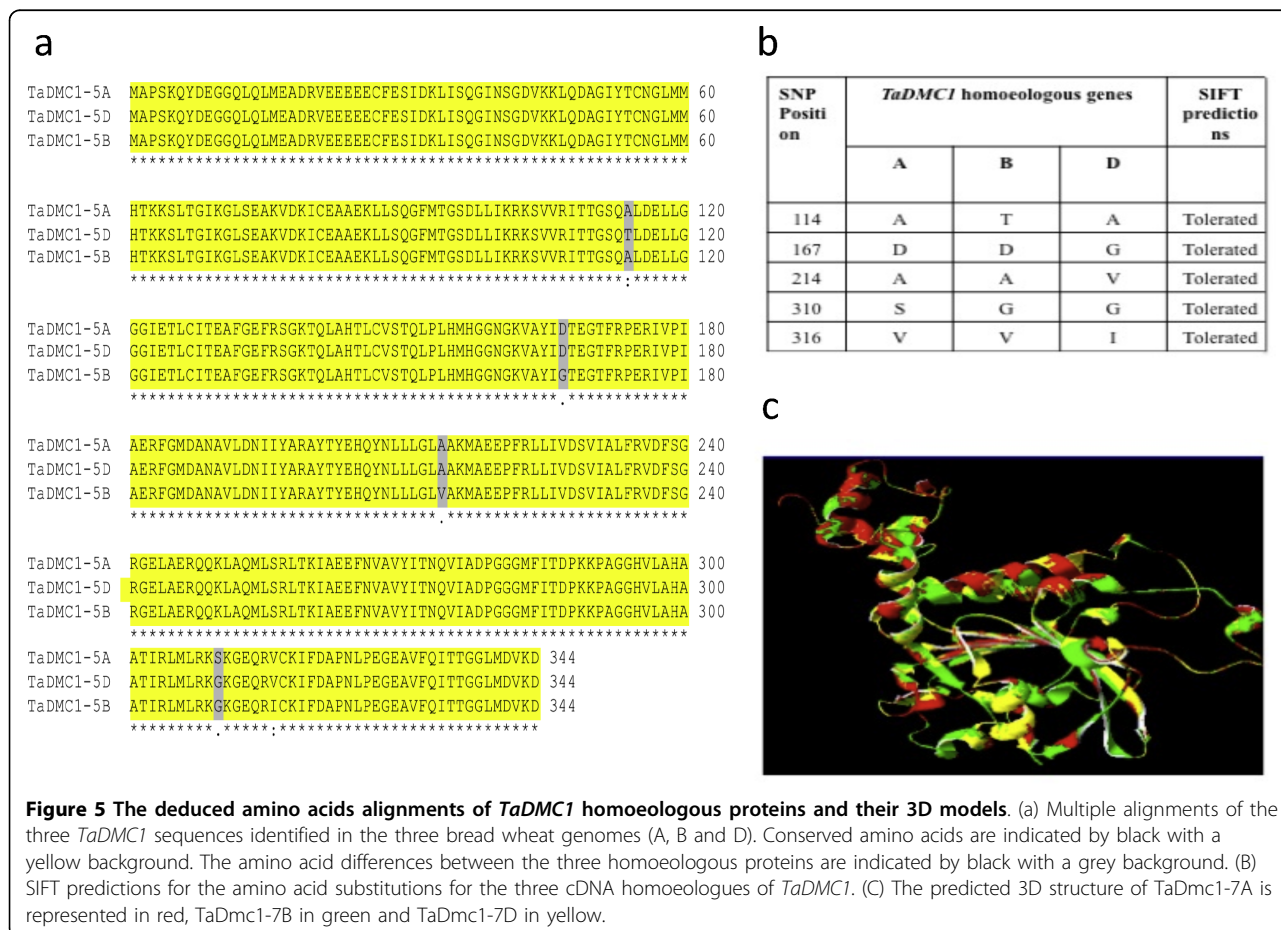
predicted to produce single amino acid substitutions at positions 114 (A&D, compared with B), 166 (A&B compared with D), 214 (A&B compared with D), 310 (A, compared with B&D) and 316 (A&B, compared with D); no insertions or deletions were detected (Figure 5a). SIFT predictions for the three translated homoeologues of *TaDMC1* suggested that the amino acid substitutions are not expected to affect protein secondary and tertiary structure (Figure 5b). Multiple alignment of the amino acid sequence of the three translated homoeologues of *TaDMC1* with the previously reported *TaDMC1* cDNA sequence [14] revealed that the reported *TaDMC1* is the D genome homoeologue (data not shown). The predicted 3D overlays for the translated *TaDMC1* homoeologues superimposed onto each other revealed very high levels of conservation for secondary and tertiary structures (Figure 5c).

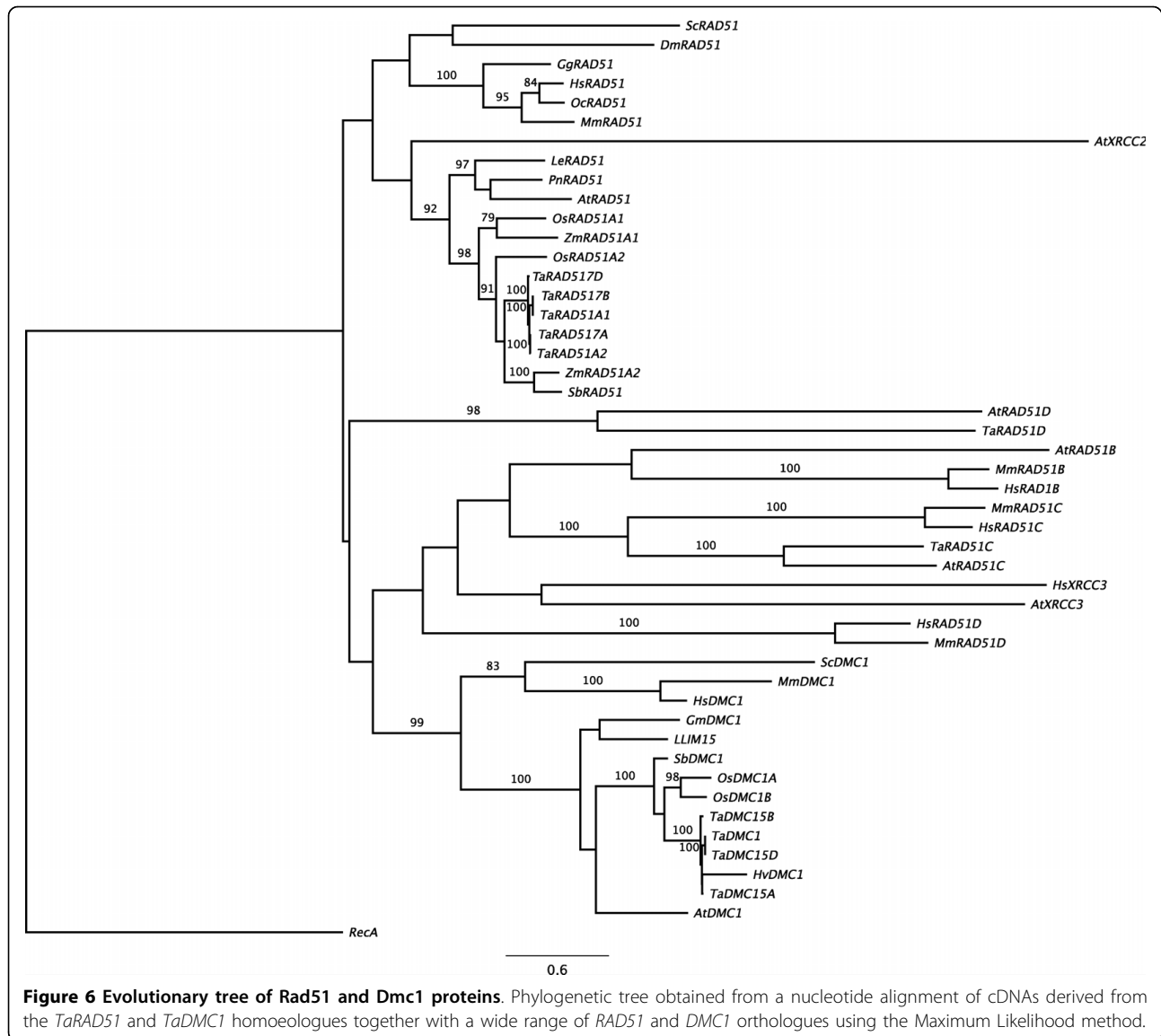
Phylogenetic analysis of *TaRAD51* and *TaDMC1* cDNA homoeologues reveals evolutionary conservation

To further analyze the level of conservation and evolutionary relationships of the Rad51 and Dmc1 predicted proteins across a diverse range of species and to

examine the predicted identity of the *TaRAD51* and *TaDMC1* cDNA homoeologues, previously annotated database entries (listed in Additional file 2) were used to construct a phylogenetic tree. The tree identified three main branches, with Rad51 and Dmc1 members clustered into two separate branches with Rad51B, Rad51C, Rad51D, Xrcc2 and Xrcc3 clustered into a third (Figure 6). This is consistent with the reports that Rad51 and Dmc1 are paralogues and likely descendants of the ancestral RecA gene [23]. As expected, the three cDNA homoeologues of *TaRAD51* and *TaDMC1* clustered together and fall into their respective branches. The strongest similarity was found between the nucleotide sequences of wheat and rice *RAD51* and *DMC1*. This provides strong evidence that the cDNA homoeologues isolated in this study are the true orthologues of the rice *RAD51* and *DMC1* genes.

This study also found that there was only one copy of *TaRAD51* per haploid genome of wheat in contrast to two copies of the *RAD51* gene reported in rice (*OsRAD51A1/A2*; rice chromosomes 11 and 12) and maize (*ZmRAD51A1/A2*; maize chromosomes 7 and 8) [8,13]. The presence of two *RAD51* copies in rice and



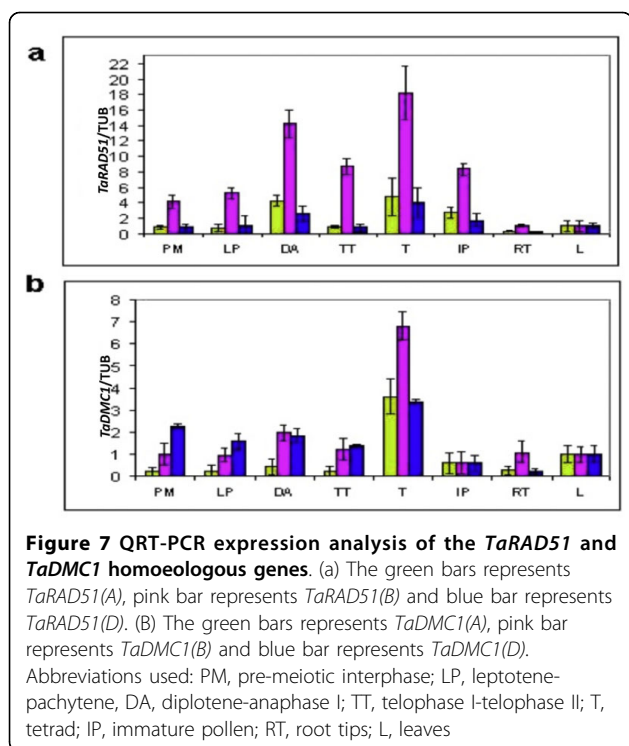


maize and only one copy of *RAD51* per genome in wheat may argue for gene loss in the ancestors of wheat. This conclusion is supported by the phylogenetic tree (Figure 6). The genomes of rice and maize contain both products of a predicted earlier gene duplication before divergence of the two species i.e *RAD51A1* and *RAD51A2*, while all three of the *RAD51* genes in wheat appear to be homologues of the *RAD51A2* gene, suggesting that the *RAD51A1* lineage has been lost in bread wheat. There are also two versions of the *DMC1* gene in rice and maize and only one in the haploid wheat genome. This may have a similar origin, but the greater conservation of *DMC1* in these species makes an assessment more difficult and the presence of two copies may simply reflect a local duplication of R11 and R12 in rice and the fact that maize is an ancient tetraploid. In both

case, it is worth noting that *DMC1* homologues are more highly conserved than the *RAD51* homologues, arguing for a higher level of selective pressure preventing mutation or a more recent duplication event.

All homoeologues of *TaRAD51* and *TaDMC1* genes were highly expressed during prophase I of meiosis

The QRT-PCR analysis of *TaRAD51* cDNA homoeologues indicated that all the three homoeologues were expressed in both vegetative and meiotic tissues (Figure 7a). Comparatively higher expression levels were observed in meiotic stages than vegetative tissues (Figure 7a, stages PM to TT). There is a sharp drop in expression of the *TaRAD51* cDNA homoeologue transcripts at the end of Meiosis I but then expression increases again at the Tetrad stage. There was however



differences in expression level among the three cDNA homoeologues, particularly during meiotic stages. This may suggest that the individual homoeologues contribute to differing extents in meiotic recombination. The expression level of *TaRAD51-7B* was significantly higher than *TaRAD51-7A* and *TaRAD51-7D* in all meiotic stages suggesting it could be the version with the greatest role in meiosis.

QRT-PCR analysis of *TaDMC1* cDNA homoeologues indicated that all three homoeologues were expressed in both vegetative and meiotic tissues (Figure 7b). However higher expression levels were found in meiotic stages (Figure 7b, stages PM to TT) compared to non-meiotic tissues and we observed an increase in expression of all three *TaDMC1* cDNA homoeologues during the Tetrad stage. There were differences in the expression levels among all three cDNA homoeologues of *TaDMC1* in the observed tissues, which suggests that they could contribute to differing extents in meiotic recombination. Both *TaDMC1-5B* and *TaDMC1-5D* homoeologues were highly expressed in all meiotic stages, although not the Tetrad stage, compared to the transcript level of *TaDMC1-5A* suggesting they could be the versions with the greatest role in meiosis.

Overall, we have reported the cloning and characterisation of the hexaploid wheat versions of the major strand-exchange proteins *DMC1* and *RAD51*. Primer tests have been developed which are specific to individual homoeologues and these have been used to characterise the levels

of genome-specific expression in a range of tissues through Q-PCR. The same primers could be used to screen mutation and deletion stocks of wheat to detect loss-of-function mutations, allowing *in planta* characterisation of the effect of these mutations on meiosis.

Additional material

Additional file 1: Alignment and comparison of the deduced amino acids of *TaRAD51* cDNA homoeologues with already reported *TaRAD51A1* & *TaRAD51A2* paralogues. Conserved amino acids are indicated by black with a yellow background. The aa similarities between *TaRAD51A1* and *TaRAD51-7D* is indicated by black with green background and aa similarities between *TaRAD51A2* and *TaRAD51-7A* is indicated by black with grey background. Deletions are shown by dashes

Additional file 2: Accession numbers of the nucleotide sequences used in phylogenetic analysis.

Abbreviations

AA: Amino Acids; DSB: Double Strand Breaks; NT: Nulli-Tetrasomics; RACE: Rapid Amplification of cDNA Ends; CS: Chinese Spring; RT-PCR: Reverse Transcription Polymerase Chain Reaction.

Acknowledgements

The authors gratefully acknowledge Prof John Snape for Highbury wheat lines, Dr Steve Reader for Nullisomic-Tetrasomic wheat stocks, Dr Hongying Li for helping with Q-PCR experiments and Dr Jason Able for providing valuable suggestions for meiotic staging of wheat anthers and for useful discussion during the project. UKD was supported by a University of Nottingham Fees PhD Scholarship. The work was partly supported by a New Lecturer grant to SM.

Author details

¹Department of Plant and Crop sciences, School of Biosciences, Sutton Bonington Campus, University of Nottingham, Loughborough LE12 5RD, UK. ²Department of Plant Biology, 1 Shields Ave, University of California Davis, CA 95616, USA.

Authors' contributions

UKD conducted the actual research, analyzed the data pertaining to experiments and drafted the manuscript. KM designed the primers for the isolation of *TaRAD51* and *TaDMC1* genes. SM guided the research and proofread this manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 25 May 2010 Accepted: 29 September 2010

Published: 29 September 2010

References

- Clark AJ, Margulies AD: Isolation and characterization of recombination-deficient mutants of *Escherichia coli* K12. *Proc Natl Acad Sci USA* 1965, 53:451-459.
- Lusetti , Cox MM: The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu Rev Biochem* 2002, 71:71-100.
- Bishop DK, Park D, Xu L, Kleckner N: *DMC1*: a meiosis-specific yeast homolog of *E. coli recA* required for recombination, synaptonemal complex formation, and cell cycle progression. *Cell* 1992, 69:439-456.
- Shinohara A, Ogawa H, Ogawa T: Rad51 protein involved in repair and recombination in *S. cerevisiae* is a RecA-like protein. *Cell* 1992, 69:457-470.
- Symington LS: Role of *RAD52* epistasis group genes in homologous recombination and double-strand break repair. *Microbiol Mol Biol Rev* 2002, 66:630-670.

6. Stassen NY, Logsdon JM, Vora GV, Offenberg HH, Palmer JD, Zolan ME: Isolation and characterization of *rad51* orthologs from *Coprinus cinereus* and *Lycopersicon esculentum*, and phylogenetic analysis of eukaryotic RecA homologs. *Curr Genet* 1997, **31**:144-157.
7. Doutriaux MP, Couteau F, Bergounioux C, White C: Isolation and characterisation of the *RAD51* and *DMC1* homologs from *Arabidopsis thaliana*. *Mol Gen Genet* 1998, **257**:283-291.
8. Franklin AE, McElver J, Sunjevaric I, Rothstein R, Bowen B, Cande WZ: Three-dimension microscopy of the *RAD51* recombination protein during meiotic prophase. *The Plant Cell* 1999, **11**:809-824.
9. Ayora S, Piruat JI, Luna R, Reiss B, Russo VEA, Aguilera A, Alonso JC: Characterization of two highly similar Rad51 homologs of *Physcomitrella patens*. *J Mol Biol* 2002, **316**:35-49.
10. Markmann-Mulisch U, Hadi MZ, Koepchen K, Alonso JC, Russo VEA, Schell J, Reiss B: The organization of *Physcomitrella patens RAD51* genes is unique among eukaryotic organisms. *Proc Natl Acad Sci USA* 2002, **99**:2959-2964.
11. Li W, Changbin C, Ullrich MM, Ljudmilla T, Elmon S, Ma H, Reiss R: The *Arabidopsis AtRAD51* gene is dispensable for vegetative development but required for meiosis. *Proc Natl Acad Sci USA* 2004, **101**:10596-10601.
12. Klimyuk VI, Jones JD: *AtDMC1*, the *Arabidopsis* homologue of the yeast *DMC1* gene: characterization, transposon-induced allelic variation and meiosis-associated expression. *Plant J* 1997, **11**:1-14.
13. Kathiresan A, Khush GS, Bennett J: Two rice *DMC1* genes are differentially expressed during meiosis and during haploid and diploid mitosis. *Sex Plant Reprod* 2002, **14**:257-267.
14. Khoo KHP, Jolly HR, Able JA: The *RAD51* gene family in bread wheat is highly conserved across eukaryotes, with *RAD51A1* upregulated during early meiosis. *Funct Plant Biol* 2008, **35**:1267-1277.
15. Sears ER: Nullisomic-tetrasomic combinations in hexaploid wheat. In *Chromosome manipulation and plant Genetics*. Edited by: Riley R, Lewis KR. Oliver 1966:29-45.
16. Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T: Protein structure homology modelling using SWISS-MODEL Workspace. *Nat Protoc* 2009, **4**:1.
17. Kaplan Warren, Littlejohn TG: Swiss-PDB Viewer (Deep View). *Brief Bioinform* 2001, **2**:195-197.
18. Pauline CN, Henikoff S: SIFT: predicting amino acid changes that affect protein function. *Nucl Acid Res* 2003, **31**:3812-3814.
19. Tamura K, Dudley J, Nei M, Kumar S: MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol and Evol* 2007, **24**:1596-1599.
20. Stamatakis A, Ludwig T, Meier H: Raxml-iii: A fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* 2005, **21**:456-463.
21. Miller MA, Holder MT, Vos R, Midford PE, Liebowitz T, Chan L, Hoover P, Warnow T: The CIPRES Portals. [http://www.phylo.org/sub_sections/portal].
22. Crismani W, Baumann U, Sutton T, Shirley N, Webster T, Spangenberg G, Langridge P, Able J: Microarray expression analysis of meiosis and microsporogenesis in hexaploid bread wheat. *BMC Genomics* 2006, **7**.
23. Lin Z, Kong H, Nei M, Ma H: Origins and evolution of the *recA/RAD51* gene family: Evidence for ancient gene duplication and endosymbiotic gene transfer. *Proc Natl Acad Sci USA* 2006, **103**:10328-10333.

doi:10.1186/1756-0500-3-245

Cite this article as: Devisetty et al.: The *RAD51* and *DMC1* homoeologous genes of bread wheat: cloning, molecular characterization and expression analysis. *BMC Research Notes* 2010 **3**:245.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

