Research in Plant Sciences, 2013, Vol. 1, No. 2, 24-31
Available online at http://pubs.sciepub.com/plant/1/2/4

DOI:10.12691/plant-1-2-4

Wheat Protein Disulfide Isomerase (PDI) Promoter Sequence Analysis in *Triticum aestivum* cv Chinese Spring and its Wild Relatives

Arun Prabhu Dhanapal[1,2,*], Enrico Porceddu[1,3]

[1]Plant Genetic Resources, Scuola Superiore Sant' Anna, Piazza Martiri della Libertà, Pisa, Italy
[2]Present addressee: Division of Plant Sciences, University of Missouri, Columbia, MO, USA
[3]Dipartimento di Agrobiologia e Agrochimica, Universita della Tuscia, Viterbo, Italy
*Corresponding author: a.dhanapal@sssup.it

Received September 13, 2013; Revised September 26, 2013; Accepted September 29, 2013

Abstract Protein disulphide isomerase (PDI) is an oxidoreductase enzyme abundant in the endoplasmic reticulum (ER). Plant PDIs has been shown to be involved in the folding and deposition of seed storage proteins, which makes this enzyme particularly interesting in wheat, as flour quality is strongly affected by composition and structure of seed storage proteins. Promoter sequences of three homoeologous genes encoding typical PDI, located on chromosome group four of bread wheat, and PDI promoter sequence analysis of *Triticum urartu*, *Aegilops speltoides* and *Aegilops tauschii* had also been reported previously. In this study, we report the isolation, cloning and sequencing of a ~1450 bp region, comprising ~1350 bp of the putative promoter region and 88 bp of the first exon of the typical PDI gene, from *Triticum urartu* (AA), *Aegilops speltoides* (BB) and *Aegilops tauschii* (DD). Sequence analysis indicated close similarity was found within each species and with the corresponding homoeologous PDI sequences of *Triticum aestivum* cv. CS (AABBDD) resulting in an overall high conservation of the sequence in proximal region then distal region of promoter conferring endosperm-specific expression.

Keywords: *protein disulphide isomerase, cloning, promoter, regulatory elements, wheat wild relatives*

Cite This Article: Arun Prabhu Dhanapal, and Enrico Porceddu, "Wheat Protein Disulfide Isomerase (PDI) Promoter Sequence Analysis in Triticum aestivum cv Chinese Spring and its Wild Relatives." *Research in Plant Sciences* 1, no. 2 (2013): 24-31. doi: 10.12691/plant-1-2-4.

1. Introduction

Wheat is adapted to temperate regions of the world and was one of the first crops to be domesticated. Spread over all continents, it is today one of the most important food source for human beings. Bread wheat (*Triticum aestivum*) belongs to the tribe *Triticeae*, and has three genomes A, B and D, each organised in seven homoeologous chromosome groups. The diploid progenitors of the three genomes A, B, and D have been identified in *Triticum uratu*, *Aegilops speltoides* and *Aegilops tauschii*, although the progenitor of B genome is still matter of debate [1]. Throughout their evolutionary history, multiple polyploidization events occurred between species of the *Triticum* and *Aegilops* genera and human manipulation of wild species led to the domestication of different cultivated lineages [2,3,4,5].

Wheat grains are consumed under many different forms, such as flour for leavened, flat and steamed breads, biscuits, cookies, cakes, breakfast cereal, pasta, noodles, couscous and for fermentation to make beer, alcohol, vodka and biofuel. It is composed by starch, proteins and other compounds, which accumulate in significant quantities during its development, in particular grain storage proteins, are responsible for the quality of the end product. The majority of them are prolamins, which account for > 90 % of the total protein content in the wheat grain. Different quality of grain storage proteins produced by the same genotype under different environmental condition has stimulated research on protein folding and assembling.

Secretory protein folding and disulphide bond formation takes place within the Endoplasmic reticulum (ER lumen), but the precise mechanisms involved, and the role of other proteins such as molecular chaperones, are not fully understood [6,7]. PDI plays an important role in assisting protein folding and assembly, catalyzing thiol-disulfide oxidation, reduction and isomeration, this latter occurring directly by intramolecular disulfide rearrangement or through cycles of reduction and oxidation [8,9,10]. During the maturation of the secretory proteins, disulfide bonds cross-linking specific cysteines are added to stabilize a protein or to join covalently different polypeptides. These bonds are crucial for the stability of the final protein structure, thus mispairing of cysteine residues can prevent proteins from attaining their native conformation and lead to misfolding [11].

Typical PDI/Classical PDI is the most prominent member of a family of related proteins (PDI-like) characterised by one, two or three thioredoxin-like active

domains [12]. It has been cloned and sequenced in many plant species, such as alfalfa [13], barley [14], maize [15], castor bean [16] and soybean [17,18,19,20], common and durum wheat [21,22]. A detailed knowledge of the complexity and diversity of genes encoding PDI and PDI-like proteins in *A. thaliana*, wheat and other plant species was described by [23,24,25].

Wheat genes coding for typical PDI in bread wheat have been located in homoeologous chromosome group four [26]. Analyses performed on 23 species of *Triticum* and *Aegilops* [27], indicated that PDI restriction fragments were highly conserved within each species and confirmed that plant PDI is encoded either by one or few copy sequences, respectively in diploid and polyploid species. The nucleotide sequences of the three genes located respectively on genomes A, B, and D (designed as *GPDI-4A, GPDI-4B* and *GPDI-4D)* were 3561bp, 3527bp and 3466 bp long. The comparison of typical PDI gene sequences of wheat, rice and Arabidopsis showed a significant conservation of the exon/intron structure across the three species [28]. More detailed study on the complexity and diversity of genes encoding PDI and PDI-like proteins in wheat and other plant species also showed a significant conservation of the exon/intron structure [23,24,29].

The expression analysis of the typical PDI homoeologous genes located on chromosomes 4A, 4B and 4D of bread wheat cv Chinese Spring (CS) [28] showed that the PDI transcripts, although constitutively present at a low-level in all the analyzed tissues, are equally abundant in the developing caryopses, but are differentially expressed in spikelets, roots and leaves [21,28]. The PDI-4A transcription was higher in spikelets that of PDI-4B were higher in roots while the PDI-4 D transcripts were more abundant in leaves. The transcription levels of the three genes were higher in the early stage of seed development (6-14 DAA) and decreased during middle to late stage of (18-34 DAA) grain filling [28]. Within the upstream putative promoter region of the three homoeologous genes cloned from bread wheat cv CS, respectively 1352 bp for PromPDI-4A, 1370 bp for PromPDI-4B and 1292 bp for PromPDI-4D long, several cis-acting elements involved in endosperm specific expression were detected, consistently with the higher PDI expression detected in the kernels.

The variability and evolutionary relationship in a ~700 bp region, comprising ~600 bp of the 5' upstream putative promoter region and 88 bp of the first exon of the typical PDI gene from the diploid species *Triticum uratu* (AA) *Aegilops speltoides* (BB) and *Aegilops tauschii* (DD) has been reported [30,31,32]. This paper reports the cloning and characterization of ~1350 bp of the 5' upstream putative promoter region and 88 bp of the first exon of the typical PDI gene from the diploid species *Triticum uratu* (AA) *Aegilops speltoides* (BB) and *Aegilops tauschii* (DD).

2. Materials and Methods

2.1. Plant Material

A total of three accessions *Triticum uratu* (IG 44831, AA), *Aegilops speltoides* (IG 46812, BB) and *Aegilops tauschii* (AE 1068, DD) were used in this study. Single plant per accessions was grown and used for DNA extraction. Flag leaves were collected at heading stage from plants grown in greenhouse (January- June 2008), immediately frozen in liquid nitrogen and kept at −80 °C until use. About 200 mg of leaf tissue was ground in liquid nitrogen and genomic DNA was extracted using Sigma Gen Elute Plant Genomic DNA Kit (G2N-350, Sigma Aldrich, St. Louis, Mo.).

2.2. Primers Design and PCR Amplification

Primers were designed on the basis of the known homoeologous PDI promoter and gene sequences isolated from bread wheat cv Chinese Spring [28], using DNAMAN 4.15 program. Putative promoters of typical PDI were amplified by using the following primer pairs PDIAPDF1-PDIAPDR1 and PDIAPDF2-PDIAPDR2 for A Genome, PDIAPDF3- PDIAPDR3 and PDIAPDF2-PDIAPDR4 for B genome and PDIAPDF3- PDIAPDR1 and PDIAPDF2- PDIAPDR6 for D genome (Table 1), cloned and sequenced in each accession.

PCR reaction mixture included 10 ng/µl genomic DNA, 0.20 mM dNTPs 5ul, 0.05 units/µl of Taq 0.50 µl (go-taq, Promega), 0.40µM primer 2.0 µl (0.20µM per each), 5X buffer 10 µl and 25.5 µl of ddH2O used per reaction (final volume of 50 µl). PCR condition included initial denaturation step at 95°C for three min, then 32 cycles of (95°C for one min, 59°C for 35 sec, 72°C for one min) and final elongation at 72°C for amplification of the sequences from the A , B and D Genome.

2.3. Cloning and Sequencing

Genome specific primer pair (Table 1) was used to amplify, clone and sequence the DNA extracted from one plant each of the three wild relatives. PCR products of expected size were excised from the gel, purified using the High Pure Purification kit (Roche) according to manufacturer's instructions, and cloned into the pGEM-T easy plasmid vector (Promega). Plasmids were transformed by heat shock into *Escherichia coli* strain $DH5_{\alpha}$. Bacteria were plated onto LB medium containing ampicillin, X-Gal and IPTG, and recombinant plasmids were identified by blue/white screening. For each primer combination two independent PCR reactions were performed and a total of six clones were sequenced. Plasmid DNA for sequencing reaction was prepared from three ml overnight cultures using a plasmid miniprep kit (Qiagen). Sequencing was performed on both strands by the ABI PRISM 377 DNA sequencer (PE Applied Biosystem) using an ABI Prism Dye Terminator sequencing kit (PE Applied Biosystem) and sequenced both with vector and sequence specific primers.

2.4. Data Analysis

Sequences were analysed in Chromas version 2.3 (http://technelysium.com.au/chromas.html) to identify any unresolved bases and subjected to visual inspection. Analysis of three new sequences were done along with the three homoeologous gene and promoter sequences from CS (GPDI-4A, AJ868102; GPDIA-4B, AJ868103; GPDI-4D, AJ868104; PromPDI4A, AJ868108; PromPDI4B, AJ868109; PromPDI4D, AJ868110; [28]). Sequences

were then searched for regulatory elements in PlantCARE (http://sphinx.rug.ac.be:8080/PlantCARE/), a database of plant promoters and (http://www.dna.affrc.go.jp/htdocs/PLACE/; [33]) PLACE database. The six sequences were multiple aligned with Clustal X software version 2.011 [34], IUB as DNA weight matrix with default parameters. A phylogenetic tree was constructed on obtained data by using the neighbour-joining (NJ) method [35], using MEGA version4 with neighbour joining option and 1000 bootstrap [36]. All three cloned sequences have been submitted in EMBL Nucleotide Sequence Database with sequence ID HE588130, HE588131 and HE588132 for *Triticum uratu* (IG 44831), *Aegilops speltoides* (IG 46812) and *Aegilops tauschii* (AE 1068) respectively.

3. Results

3.1. PDI Promoter Sequence

The PDI promoter sequences from one plant each of *Triticum uratu* (IG 44831), *Aegilops speltoides* (IG 46812) and *Aegilops tauschii* (AE 1068) were amplified using the following primer pairs (Table 1) corresponding respectively to sequences comprising putative PDI promoter and part of the first exons of the PDI gene belonging to the A, B and D genome was visualized by gel electrophoresis. Two independent PCR reactions were performed for each of the two primer pair/DNA combinations used and in all cases a specific amplification product of the expected electrophoretic mobility was obtained and cloned. For every each cloned amplification product three clones were randomly chosen and sequenced, for a total of 6 clones for each primer pair/DNA combination. The six DNA sequences deriving from the same primer pair/DNA combination resulted 100% identical.

The length of the cloned region upstream of the translation start codon was 1334 bp for *Triticum uratu* (AA, accession IG 44831), 1332 bp for *Aegilops speltoides tauschii* (DD, accession AE 1068); the three sequences showed 92.38 % identity in the promoter region (Figure 2). The specificity and uniqueness of their respective amplification products was confirmed through cloning all fragments and sequencing.

3.2. Similarity of Sequences with CS

A high degree of conservation was detected in the putative promoter sequences of the entire three wild genomes cloned. Differences were due to both nucleotide substitutions and short insertions/deletions. The promoter sequence of *Triticum uratu* (AA, accession IG 44831) showed 99.40 % identity with one deletion and nucleotide substitution. *Aegilops tauschii* (DD, accession AE 1068) promoter sequence showed 99.04 % identity with few insertion and nucleotide substitution. Whereas promoter sequence of *Aegilops speltoides* (BB, accession IG 46812) showed only 89.59 % identity with large number of deletions and nucleotide substitution.

3.3. Phylogenetic Analysis

The evolutionary relationship between six sequences, of the putative promoter was studied by phylogeny reconstruction. These included one cloned sequence from *Triticum uratu*, *Aegilops speltoides*, *Aegilops* tauschii for a total of 3 new sequences and the three homoeologous sequences from CS previously reported [28]. The phylogenetic tree was constructed using three different methods namely the neighbour-joining (NJ) method, Minimum Evolution (ME) and Maximum Parsimony (MP) methods. As the results of the three methods were similar, only NJ tree is presented here (Figure 1).

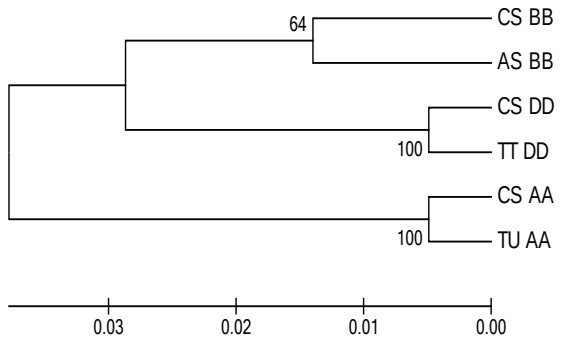

Figure 1. Phylogenetic tree of PDI gene promoter sequence of *Triticum urartu* (TU AA), *Aegilops speltoides* (AS BB) and *Aegilops taushcii* (TT DD) with three homoeologous promoter sequences from CS

Table 1. List of primers used for the isolation of PDI promoter analysis

PCR target for Cloning	Primer	Forward	Reverse
	PDIAPDF1	5'-TTTCCAGCTTGAGATGAGGC-3'	
	PDIAPDF2	5'-ACTCCAAATTTGGAACGGG-3'	
AA GENOME	PDIAPDR1		5'-TTGCCGGTTCAGGACTAGG -3'
	PDIAPDR2		5'-GGTGAGCACTGCCTCGGG-3'
	PDIAPDF3	5'- GTTGTTTTCTTCCAAGCTT-3'	
	PDIAPDF2	5'-ACTCCAAATTTGGAACGGG-3'	
BB GENOME	PDIAPDR3		5'-TGGCGGAAATATTTATGTTTG-3'
	PDIAPDR4		5'- GAAGTTGTCGGCGTGCA-3'
	PDIAPDF3	5'- GTTGTTTTCTTCCAAGCTT-3'	
	PDIAPDF2	5'-ACTCCAAATTTGGAACGGG-3'	
DD GENOME	PDIAPDR1		5'-TGGCGGAAATATTTATGTTTG-3'
	PDIAPDR6		5'-GGTGAGCACTGCCTCGGG-3'

```
TU_AA                          TTTCCAGCTTGAGATGAGGCCTATCTCCCGTGCTCAGCTCCTGCAATGCATACAAGAAAGA    61
AS_BB      GTTGTTTTCTTCCAAGCTTTTTCCAGCTTGAGATGAGGCCTATCTCCCGTGCTCAGCTCCTGCAATGCATACAAGAAAGA    80
TT_DD      GTTGTTTTCTTCCAAGC..TTTCCAGCTTGAGATGAGGCTTATCTCCCGTGCTCAA...CTACAATACATACAAG.....    70

TU_AA      CAAAAAGATACATGCAACAAACAAACGTCTCAGCTTTGAGCTGGGTGGGTGGATTACAGGCAGTGCAAACGCTTTTGCAC   141
AS_BB      CAAAAAGATACATGCAACAAACAAACGTCTCAGCTTTGAGCTGGGTGGGTGGATTACAGGCAGTGCAAACGCTTTTGCAC   160
TT_DD      ............................CTTTTTCCAGCTTGGTGG.TGGATTACAGGCAGTGCAAACGCTTTTGCCC   119

TU_AA      TACCGAGTTTGATGTCAGTTGCGCATAACACCAAGATTAGTGTTTCCTTGCACGTAGCATTTGGACCTTGTGCTGCACTT   221
AS_BB      TACCGAGTTTGATGTCAGTTGCGCATAACACCAAGATTAGTGTTTCCTTGCACGTAGCATTTGGACCTTGTGCTGCACTT   240
TT_DD      C.......TTGAT.TCAGTTGCGCATAACACCAAGATTAGTGTTTCCTTGCACGTAAGATTTGGGCCTTGTGTTGCACTT   191

TU_AA      TTCAACCAATTTCGGAATTATTACGTACTGCCGTTAGATGATCTTTTTCATTCGCACCGCATTTTGGCAAAG...GTTGT   298
AS_BB      TTCAACCAATTTCGGAATTATTACGTACTGCCGTTAGATGATCTTTTTCATTCGCACCGCATTTTGGCAAAG...GTTGT   317
TT_DD      TTCAACCATTTTCGGAATTATCACGCACTGCGGTTAGATGATCTTTTTCATTCACACCGCATTTGGACGAAGCAGGTTGT   271

TU_AA      TCGCAAAGGCGTCGC..............GACGATCATTTTGCACCCTCCACTAATGCATTTTTTAACGAAACAGGAAA   363
AS_BB      TCGCAAAGGCGTCGC..............GACGATCATTTTGCACCCTCCACTAATGCATTTTTTAACGAAACAGGAAA   382
TT_DD      TCGCAAATGCGTCGCCGTCAGTTTTGCTGCGACGATTATTTTGCACCGTCCACTAATCCATTTTTTAGCGAAACAGAAAA   351

TU_AA      TGCCATATATGTCCAGTCTGCTAAAAGGACTCGTCTGAAATGCAAAACACCTCAAACAGTAATTATTATTATTTACTCAG   443
AS_BB      TGCCATATATGTCCAGTCTGCTAAAAGGACTCGTCTGAAATGCAAAACACCTCAAACAGTAATTATTATTATTTACTCAG   462
TT_DD      TGCCGCATATGTCCAGTCTGCTAAAAGGACTCATCAGAAATGCAAAACACCTCGAACA.CAAATATTA.........CAG   421

TU_AA      TAGGGGGCAGTGCATATATACATGTGTTGTGAACGTCTGCATTTATATTGCGTTTCTAAAAAAGAAAGTATTTCCCCATG   523
AS_BB      TAGGGGGCAGTGCATATATACATGTGTTGTGAACGTCTGCATTTATATTGCGTTTCTAAAAAAGAAAGTATTTCCCCATG   542
TT_DD      TAATA...ATTATGCATATACGTGT.CTGTGAACGTGTGCATTTGTACTGTGTTTCTAAAAAA..AGGTATTTTTC.ATG   494

TU_AA      CATTTTTTGCGCGACGT.AGATGGTTAATGTGTGGCACGTAGAATCGGCCTACAGTATGAAACGAATCAAAATTACTCCT   602
AS_BB      CATTTTTTGCGCGACGT.AGATGGTTAATGTGTGGCACGTAGAATCGGCCTACAGTATGAAACGAATCAAAATTACTCCT   621
TT_DD      CATTTTTCG.GCGACGTGAGATGGATAATGTGTGGCTCGT.....CACGCTACAGTGTGAAACGAATCAAAATTACGCA.   567

TU_AA      CCTATGCAGAATTATTGTGTTGCACGCGAATCATTAATCACCAAAGATACTCCAAATTTGGAACGGGTTAATTTCTTCAT   682
AS_BB      CCTATGCAGAATTATTGTGTTGCACGCGAATCATTAATCACCAAAGATACTCCAAATTTGGAACGGGTTAATTTCTTCAT   701
TT_DD      .....GAATTATTATTGTGTTCCATGCAAATCATTAATCACCAAAGATACTCCAAATTTGGAACGGGTTAATTTCTTCAT   642

TU_AA      ACATTTATATGGAGTGAATGAATGAAGGATCTCCTCAGATAAGGATGGAATTAATAATCTTCCAAACATAAATATTTCCG   762
AS_BB      ACATTTATAAGCAGCGAATGAATGAATGATCTCCTGAGATA.GGATGGAATTAATAATCTTCCAAACATAAATATTTCCG   780
TT_DD      ACATTTATACGCAGCGAATGAATGAATGATCTCCTCGGCTACGGATGGAATTAATAATCTTCCAAACATAAATATTTCCG   722

TU_AA      CCATATCT.CCTAGTCCTGAACCGGCAACAGTGCCCGGTGCGCGTCGGCTGAGCAGCCTCGGCTAATCAGGTAGCAGTGG   841
AS_BB      CC.......................GAAACAGTGCCCGGTGCGCGTCGGCTGAGTAGCCTCGGCTAATCAGGAGGCAGTGG   838
TT_DD      CCATATCTTCCTAGTCCTGANNCGGCAACAGTGCCCGGTGCGCGTCGGCTGAGTAGCCTCGGCTAATCAGGAGGCAGTAG   802

TU_AA      ACATGTCACTGACACGACGTGGCGAGTTCCGGTTGGTCTGACGCCATACGGGTACGTGTTTGCAGCCCACGTGTATCGTA   921
AS_BB      ACATGTCACTGACACGACGTGGCGAGTCCCGATTGGTCTGGCCCCATACGGGTACGTGTTTGCAGCCCACGTGTACCGTA   918
TT_DD      ACGTGTCACTCACATGACGTGGCGCCCTGCGATTCGCCTGACCCCATCAGGGTACGTTTTTGCAGCCCACGTGTACCGTA   882

TU_AA      CACTGGCGGCCGCACGAGCGTGACGTGGCGTCCCGGATCCGCTAACCCCTCCCGTTTCGACTCACGCGGCATATTCCAGT  1001
AS_BB      CACTGGCGGC.GCACGAGCGTGACGTGGCGTCCCGGATCCGCTAACCCCTCCCTTTTCGAGTCACGCGGCATGTTCCAGT   997
TT_DD      CAGTGGCGGC.GCACGAGCGTGACGTGGCGTCCCGGATCCGCTAACCCCTCCCTTTTCGAGTCACGCGGCATATTCCAGT   961

TU_AA      TAATTGGATCAACACCCTCGGTAGGGACCTCATAGTAATTCACTAGCCTAAC.TTTAGCGGCCTTAGCACCTTAAATTGG  1080
AS_BB      TAATTGGATCGACACCCTCGGTAGGGACCTCATTGTAATTTATCAGCCTAACATTTAGCGGCCTTAACACCTCAGATTGG  1077
TT_DD      TAATTGGATCAACACCCTCGGTAGGGACCTCATTGTAATTTACTAGCCTAACTTTAAGCGGCCTTAACACCTCAAATTGG  1041

TU_AA      GATTCTGATAGTTACGGGGACCTGAGTGTGAAAAGTCCTAGTGATCATCCCGTGATCTCTGTGGGCCGCGCCCACAGAGC  1160
AS_BB      GATTCGGATAGCAACAGGGACCTGAGTGCGAAAAGTCCGAGTAATCCTCCCGTGATCTCTGTGGGCCGCGCCCACAGCGC  1157
TT_DD      GATTCGGATAGCAACAGGGACCTGAGTGCAAAACATCCCAGTAATCATCCCGTGATCTCTGTGGGCCGCGCCCACGGAGC  1121

TU_AA      CCTGGCCGGCGTGGGACCCGTGGACATGGTCGAATATGCCGTGGACCGCGTCCACGTAGGAGACTCTAAACCCC.ACCTC  1239
AS_BB      CCTGGCCGTCGTGGGACCCGTGGACGCGGTCCAGTCCACCGTGGACCGCGTCCACGTTGGAGACTCTAAACCCCCACCTC  1237
TT_DD      CCTGGCCGTCGTGGGACCCGTGGACACGGTCCAGTCCACCGTGGACCGCGTCCACGTTGGAGACTCTAAACCCC.ACCTC  1200

TU_AA      TCGGCCTCTCCGCTCCTATTAAAACCCGTGGAGCACCGCGGCGTGGTCGTGTCCGCCCCGAGAAATCTCATCACTGCTCC  1319
AS_BB      TCGGCCTCTCCGCTCCTATTAAAACCCGTGGAGCACCGCGGCGTGGTCGTGTCCGCCCCGAGAAATCTCATCACTGCTCC  1317
TT_DD      TCGGCCTCTCCGCTCCTATTAAAACCCGTGGAGCACCGCGGCGTGGTCGTGTCCGCCCCGAGAAATCTCATCACTGCTCC  1280

TU_AA      CCAGTCCCTTCCGCC                                                                  1334
AS_BB      CCAGTTCCTTCCGCC                                                                  1332
TT_DD      CCAGTCCCTTCCGCC                                                                  1295
```

Figure 2. Multiple sequence alignment of PDI promoter sequences from *Triticum uratu* (TU_AA), *Aegilops speltoides* (AS_BB) and *Aegilops tauschii* (TT_DD)

Table 2. Main regulatory motifs found in the 5' Upstream of typical PDI gene sequence in *Triticum aestivum, Triticum urartu, Aegilops speltoides* and *Aegilops taushcii* in both positive and negative strands from ATG

Motif	Function	Sequence	A Genome				B Genome				D Genome			
			Triticum aestivum (4A)		Triticum urartu (AA)		Triticum aestivum (4B)		Aegilops speltoides (BB)		Triticum aestivum (4D)		Aegilops tauschii (DD)	
			+ Strand	- Strand	+ Strand	- Strand	+ Strand	- Strand	+ Strand	- Strand	+ Strand	- Strand	+ Strand	- Strand
Prolamin box	Cis-acting element associated with GCN4 in prolamin genes	TGAAAAGT CGAAAAGT CGAAAAGT TGAAAAGT TGAAAAGT TGAAAAGT	-226 -1117		-226		-227 -1135		-227		-226	-1105	None	
TATA box		TATTAAA TATTAAA TATTAAA	-79		-79		-79		-79		-79		-79	
Skn-1	Cis-acting regulatory element for endosperm expression	ACGAC ACGAC ATGAC		-480	-480	-50		-481	-481	-50, -168		-479	-480	-50,-167
GCN4	Cis-regulatory element involved in seed specific expression	CGACTCA TGAGTCA ACTGTGC CGAGTCA CATGTCA CGTGTCA CGAGCCA	-356 -492	-484 -780	-492 -356	-484	-358 -493	-485 -775	-358 -493	-479 -358	-357 -491	-483 -769	-358 -492	-484
			+ Strand (number)	- Strand (number)	+ Strand (number)	- Strand (number)	+ Strand (number)	- Strand (number)	+ Strand (number)	- Strand (number)	+ Strand (number)	- Strand (number)	+ Strand (number)	- Strand (number)
AACA	Cis-acting element conserved in rice glutelin genes and involved in	AACC	1		8	5	1		11	7	1		10	6
CAAT box	Cis-acting element common in promoter and enhancer regions		12	12	4	6	14	17	3	8	12	15	6	8

3.4. Conserved Regulatory Motifs

The search for regulatory motifs in the promoter sequences of *Triticum uratu* (IG 44831), *Aegilops speltoides* (IG 46812) and *Aegilops tauschii* (AE 1068) upstream the coding region in the database of plant promoters (PlantCARE) and PLACE database [http://www.dna.affrc.go.jp/htdocs/PLACE/; [33]] detected, a TATA box located at −79 nt from the start codon, and a number of different cis-acting regulatory elements (Table 2) including several motifs (AACA, prolamin box, GCN4, Skn-1) involved in the regulation of endosperm specific genes [41,42,43,44].

Several CAAT-boxes were present in both (+) strand and (-) strands of A, B and D genome sequences controlling endosperm and tissue specific expression. AACA motif was present in (+) strand of all genomes, eight in A genome, eleven in B genome and 10 in D genome, but in (-) strand, only five in A, seven in B and six in D genome were identified. A prolamine-box was discovered on (+) strands at position -226 bp in *Triticum uratu* and at position -227 bp in *Aegilops speltoides.* No prolamine-box was detected in the target region of *Aegilops tauschii*. The prolamine-box was present in an identical position in the corresponding promoter sequences of the PDI gene in CS located on homoeologous chromosomes 4A and 4B of CS as described in previous study [28]. Skn-1 like element were found at position -480 bp, -481bp and -480 bp on (+) strand respectively in the A, B and D chromosomes. In the homoeologous sequences of CS, [28] identified them at position -480bp, -481bp and -479bp respectively in the sequence of the putative PDI promoter from genome 4A, 4B and 4D. In addition to the above-mentioned regulatory elements (AACA, GCN-4, prolamin box, Skn-1,) involved in the regulation of endosperm specific genes, other regulatory motifs identified in the promoter region was CAAT-box responsible for tissue specific gene expression. Search for regulatory motifs have identified several cis-acting regulatory elements and other motifs conferring endosperm specific and tissue specific expression. All promoter elements identified in partial promoter region of three wild species reported previously was reconfirmed in our present study [32].

4. Discussion

The promoter sequence of the three homoeologous genes in hexaploid wheat *Triticum aestivum* cv Chinese Spring showed overall identity of 89% in our previous study; high degree of conservation was found in the 700 nt proximal sequence, with identity exceeding 93%, when compared to distal region with 80% identity [28]. In our present study overall identity was 92 % among three sequences from wild realtives; high degree of conservation was found in 650 nt proximal sequence, with identity exceeding 94 %, when compared to distal region with 89 % identity.

All sequences were grouped into two major clusters, one containing the sequences derived from A genome, and one containing two subclusters each formed respectively by the B and D genome sequences. The sequence from *Aegilops speltoides* (BB) closely clustered together with the sequence located on the chromosome 4B of CS. The homoeologous B genome of hexaploid wheat is closely related to the genome of *Aegilops speltoides* that is proposed to be its wild progenitor [37,38]. The sequence from the *Aegilops tauschii* (DD) accessions closely clustered with sequence located on the 4D chromosome of CS. Several studies underline that *Aegilops tauschii*, also known as wild goat grass from the Middle East, crossed with modern versions of emmer wheat to produce bread wheat [38,39,40]. As expected also sequence of *Triticum uratu* (AA) closely clustered with the sequence from CS located on chromosome 4A.

Our present study had shown strong evolutionary relationship from three new sequences *of Triticum uratu*, *Aegilops speltoides* and *Aegilops* tauschii cloned in this study with previously identified three homoeologous sequences from CS. Similar result was also obtained from sequences of partial promoter region from *Triticum uratu*, *Aegilops speltoides* and *Aegilops* tauschii reported previously [32,49,50].

Former studies in wheat had been restricted to the characterization of genes encoding the typical PDI and their promoter, and the cloning and characterization of complete set of genes encoding PDI and PDI like proteins in bread wheat (*Triticum aestivum* cv Chinese Spring) and the comparison of their sequence, structure and expression with homologous genes from other plant species, which is of special interest for its involvement in determining bread making qualities and technological properties of flour. The interest of extending the study to promoters of wild species is due to high conservation of gene, due to their relevant metabolic functions, as well as to the interesting expression pattern found in previous work. Present study confirmed the conservation of important motifs conferring the endosperm specific expression with evolutionary relationship between wild relatives and bread wheat (*Triticum aestivum* cv Chinese Spring). An intensive analysis carried out on the transcription levels of nine PDI and PDI-like genes in a set of 29 samples of wheat, including tissues, developmental stages and temperature stresses, showed their constitutive although very variable expression. Highly diversified expression rates and expression patterns were evidenced not only by genes belonging to different phylogenetic groups, but also in close paralog genes. This variable expression pattern may be due to differing regulatory elements and their numbers. The very high expression of the gene TaPDIL1-1, encoding the typical PDI, in the developing caryopses, which is consistent with its hypothesised role in the folding, aggregation and deposition of seed storage proteins was confirmed. This function of the typical PDI has been demonstrated experimentally in maize, rice and soybean [23,24,45,46,47,48,51,52].

Future studies should involve characterization of proximal and distal end of the Typical PDI promoter in diploid and tetraploid progenitor from diverse geographic origin of world to determine to determine the similarity and identity. Functional analysis of Chinese spring PDI promoter (full and partial sequence) driving the GUS gene and selective silencing of the PDI and PDI-like genes in wheat plants. The characterization of the regulatory motifs through the expression studies of the progressive deletions of their promoters, as well as the expression analysis of the PDI gene from accession of wild and cultivated wheat

is currently under way. This analysis will help in elucidating the function of some regulatory elements controlling the spatial and temporal specific expression of the PDI and PDI –like genes.

5. Conclusions

The very high expression of the gene TaPDIL1-1, encoding the typical PDI, in the developing caryopses, which is consistent with its hypothesised role in the folding, aggregation and deposition of seed storage proteins was confirmed. This function of the typical PDI has been demonstrated experimentally in maize, rice and soybean. Future studies should involve characterization of proximal and distal end of the Typical PDI promoter in diploid and tetraploid progenitor from diverse geographic origin of world to determine to determine the similarity and identity. Functional analysis of Chinese spring PDI promoter (full and partial sequence) driving the GUS gene and selective silencing of the PDI and PDI-like genes in wheat plants. The characterization of the regulatory motifs through the expression studies of the progressive deletions of their promoters, as well as the expression analysis of the PDI gene from accession of wild and cultivated wheat is currently under way. This analysis will help in elucidating the function of some regulatory elements controlling the spatial and temporal specific expression of the PDI and PDI –like genes.

Acknowledgement

I thank Shalini Narasimhan for helping me in editing grammatical errors and finalizing the manuscript. My Ph.D programme scholarship entitled “International Doctoral Scholarship for Agrobiodiversity” from Scuola Superiore Sant’ Anna Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy from February 2007 to February 2010 is highly acknowledged.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

[1] Gill, B.S. et al., 2004. “A workshop report on wheat genome sequencing: international genome research on wheat consortium.” *Genetics*, 168: 1087-1096.

[2] Kihara, H., 1919. “Uber cytolosgische Studien bei einigen Getreidearten. I. Spezies- Bastarde des Weizens und Weizenmggen-Bastarde.” *Botanical Magazine (Japanese)*, 33: 17-38.

[3] Kihara, H., 1924. “Cytologische und genetische Studien bei wichtigen Getreidearten mit besonderer Rucksicht auf das Verhalten der Chromosomen und die Sterilität in den Bastarden.” *Memoirs of the College of Science, Kyoto Imperial University*, 8: 1-200.

[4] Morris, R. and Sears, E.R., 1967. “The cytogenetics of wheat and its relatives. In: Wheat and Wheat Improvement (Quisenberry KS and Reitz LP, Eds.).” *American Society of Agronomy*, Madison WI: 19-87.

[5] Sax, K., 1922. “Sterility in wheat hybrids. II. Chromosome behavior in partially sterile hybrids” *Genetics*, 7(6): 49-68.

[6] Shewry, P.R., Napier, J.A. and Tatham, A.S., 1995. “Seed storage proteins: structure and biosynthesis.” *Plant Cell*, 7(7): 945-956.

[7] Wang, C. et al., 2012. “Human protein-disulfide isomerase is a redox-regulated chaperone activated by oxidation of domain a'.” *Journal of Biological Chemistry*, 287(2): 1139-1149.

[8] Cho, E.J. et al., 2011. “Protein disulfide isomerase-2 of Arabidopsis mediates protein folding and localizes to both the secretory pathway and nucleus, where it interacts with maternal effect embryo arrest factor.” *Molecules and Cells*, 32(5): 459-475.

[9] Onda, Y. et al., 2011. “Distinct roles of protein disulfide isomerase and P5 sulfhydryl oxidoreductases in multiple pathways for oxidation of structurally diverse storage proteins in rice.” *Plant Cell*, 23(1): 210-223.

[10] Schwaller, M., Wilkinson, B. and Gilbert, H.F., 2003. “Reduction-reoxidation cycles contribute to catalysis of disulfide isomerisation by protein-disulfide isomerase.” *Journal of Biological Chemistry*, 278(9): 7154-7159.

[11] Tu, B.P. and Weissman, J.S., 2004. “Oxidative protein folding in eukaryotes:mechanisms and consequences.” *Journal of Cell Biology*, 64(3): 341-346.

[12] Ferrari, D.M. and Soling, H.D., 1999. “The protein disulphide-isomerase family: unravelling a string of folds.” *Biochemistry Journal*, 339(1): 1-10.

[13] Shorrosh, B.S. and Dixon, R.A., 1991. “Molecular cloning of a putative plant endomembrane protein resembling vertebrate protein disulfide-isomerase and a phosphatidylinositol-specific phospholinase.” *Proceedings of National Academy of the Sciences*, 88(23): 10941-10945.

[14] Chen, F. and Hayes, P.M., 1994. “Nucleotide sequence and developmental expression of duplicated genes encoding protein disulfide isomerase in barley (Hordeum vulgare L.).” *Plant Physiology*, 106(4): 1705- 1706.

[15] Li, C.P. and Larkins, B.A., 1996. “Expression of protein disulfide isomerase is elevated in the endosperm of the maize floury-2 mutant.” *Plant Molecular Biology*, 30(5): 873-882.

[16] Coughlan, S.J., Hastings, C. and Winfrey, R.J., 1996. “Molecular characterization of plant endoplasmic reticulum: Identification of protein disulfide-isomerase as the major reticuloplasmin.” *European Journal of Biochemistry*, 235(1-2): 215-224.

[17] Iwasaki, K. et al., 2009. “Molecular cloning and characterization of soybean protein disulfide isomerase family proteins with non classic active center motifs.” *FEBS Journal*, 276(15): 4130-4141.

[18] Kamauchi, S. et al., 2008. “Molecular cloning and characterization of two soybean protein disulfide isomerases as molecular chaperones for seed storage proteins.” *FEBS Journal*, 275(10): 2644-2658.

[19] Wadahama, H., Kamauchi, S., Ishimoto, M., Kawada, T. and Urade, R., 2007. “Protein disulfide isomerase family proteins involved in soybean protein biogenesis.” *FEBS Journal* 274(3): 687-703.

[20] Wadahama, H. et al., 2008. “A novel plant protein disulfide isomerase family homologous to animal P5: molecular cloning and characterization as a functional protein for folding of soybean seed storage proteins.” *FEBS Journal*, 275(3): 399-410.

[21] Ciaffi, M., Paolacci, A.R., Dominici, L., Tanzarella, O.A. and Porceddu, E., 2001. “Molecular characterization of gene sequences coding for protein disulfide isomerase (PDI) in durum wheat (*Triticum turgidum* ssp. durum).” *Gene* 265(1-2): 147-156.

[22] Shimoni, Y., Segal, G., Zhu, X. and Galili, G., 1995a. “Nucleotide sequence of a wheat cDNA encoding protein disulfide isomerase.” *Plant Physiology*, 107(1): 281.

[23] d'Aloisio, E. et al., 2010. “The Protein Disulfide Isomerase gene family in bread wheat (*T. aestivum* L.).” *BMC Plant Biology*, 10: 101.

[24] d’Aloisio, E. et al., 2011. “Protein disulfide isomerase family in bread wheat (*Triticum aestivum* L): genomic structure, synteny conservation and phylogenetic analysis.” *Plant Genetic Resources*, 9: 347-351.

[25] Houston, N.L. et al., 2005. “Phylogenetic Analyses Identify 10 Classes of the Protein Disulfide Isomerase Family in Plants, Including Single-Domain Protein Disulfide Isomerase-Related Proteins.” *Plant Physiology*, 137(2): 762-778.

[26] Ciaffi, M., Dominici, L., Tanzarella, O.A. and Porceddu, E., 1999. “Chromosomal assignment of gene sequences coding for protein disulphide isomerase (PDI) in wheat.” *Theoritical and Applied Genetics*, 98(3-4): 405-410.

[27] Ciaffi, M., Dominici, L., Umana, E., Tanzarella, O.A. and Porceddu, E., 2000. “Restriction fragment length polymorphism

(RFLP) for protein disulfide isomerase (PDI) gene sequences in *Triticum* and *Aegilops* species." *Theoritical and Applied Genetics*, 98(1-2): 405-410.

[28] Ciaffi, M., Paolacci, A.R., d'Aloisio, E., Tanzarella, O.A. and Porceddu, E., 2006. "Cloning and characterization of wheat PDI (Protein disulfide isomerase) homoeologous genes and promoter sequences." *Gene*, 366(2): 209-218.

[29] Han, X. et al., 2012. "The failure to express a protein disulphide isomerase-like protein results in a floury endosperm and an endoplasmic reticulum stress response in rice." *Journal of Experimental Botany*, 63(1): 121-130.

[30] Dhanapal, A.P and Porceddu, E., 2013. "Funtional and Evolutionary Genomics of Protein Disulphide Isomerase (PDI) Gene Family in Wheat and its Wild relatives: A Review." *Annual Review and Research in Biology*, 3(4): 935-958.

[31] Dhanapal, A.P., 2012. "Evolutionary Relationship of Wheat Protein Disulphide Isomerase (PDI) Gene Promoter Sequence Based on Phylogenetic Analysis." *American Journal of Plant Sciences* 3(3): 373-380.

[32] Dhanapal, A.P., Ciaffi, M., Porceddu, E. and d Aloisio, E., 2011. "Protein disulphide isomerase promoter sequence analysis of *Triticum urartu, Aegilops speltoides* and *Aegilops tauschii." Plant Genetic Resources*, 9: 338-341.

[33] Higo, K., Ugawa, Y., Iwamoto, M. and Higo, H., 1999. "PLACE: a database of plant cis-acting regulatory DNA elements." *Nucleic Acids Research*, 26(1): 358-359.

[34] Larkin, M.A. et al., 2007. "Clustal W and Clustal X version 2.0." *Bioinformatics*, 23(21): 2947-2948.

[35] Saitou, N. and Nei, M., 1987. "The neighbor-joining method: A new method for reconstructing phylogenetic trees." *Molecular Biology and Evolution* 4(4): 406-425.

[36] Tamura, K., Dudley, J., Nei, M. and Kumar, S., 2007. "MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0." *Molecular Biology and Evolution* 24(8): 1596-1599.

[37] Dvorak, J., Luo, M.C. and Yang, Z.L., 1998a "Genetic evidence on the origin of T. aestivum L." *In: The Origins of Agriculture and the Domestication of Crop Plants in the Near East*, edited by A. DAMANIA. ICARDA, Aleppo, Syria: 235-251.

[38] Gitte, P., Seberg, O., Yde, M. and 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)." *Molecular Phylogenetics and Evolution*, 39(1): 70-82.

[39] Kimber, G. and Feldman, M., 1987. "Wild wheat, an introduction." *Columbia (MO): College of Agriculture University of Missouri Special Report* 353: 146.

[40] Simonite, T., 2006. "Ancient genetic tricks shape up wheat Turning back the evolutionary clock offers better crops for dry regions." *Nature*.

[41] Albani, D. et al., 1997. "The wheat transcriptional activator SPA: a seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes." *Plant Cell* 9(2): 171-184.

[42] Chandrasekharan, M.B., Bishop, K.J. and Hall, T.C., 2003. Module-specific regulation of the beta-phaseolin promoter during embryogenesis. Plant Journal, 33(5): 853-866.

[43] Wu, C., Washida, H., Onodera, Y., Harada, K. and Takaiwa, F., 2000. "Quantitative nature of the Prolamin-box, ACGT and AACA motifs in a rice glutelin gene promoter: minimal cis-element requirements for endosperm-specific gene expression." *Plant Journal*, 23(3): 415-421.

[44] Wu, C.Y., Suzuki, A., Washida, H. and Takaiwa, F., 1998. "The GCN4 motif in a rice glutelin gene is essential for endosperm-specific gene expression and is activated by Opaque-2 in transgenic rice plants. *Plant Journal*, 14(6): 673-683.

[45] d'Aloisio, E., Dhanapal, A.P., Paolacci, A.R., Tanzarella. O.A., Ciaffi, M., Porceddu, E., "The PDI (Protein Disulfide Isomerase) gene family in wheat." *Proceedings of the 11th International Wheat Genetics Symposium*,24-28 August 2008b, Brisbane QLD, Australia.

[46] d'Aloisio, E., Dhanapal, A.P., Paolacci, A.R., Tanzarella. O.A., Ciaffi, M., Porceddu, E., "Protein Disulphide Isomerase Gene family in Bread Wheat – Structural Characterization and Expression analysis." *Proceedings of the International Workshop on Plant Genetic Resources For Food And Agriculture- General aspects and research opportunities*, 2009b;5-6 November 2009b, Rome, Italy.

[47] d'Aloisio, E., Dhanapal, A.P., Paolacci, A.R., Tanzarella. O.A., Ciaffi, M., Porceddu, E., "Genomic Structure and Synteny Conservation Towards Rice Genome of Genes Encoding Proteins of The Protein Disulfide Isomerase Family In Wheat." *Proceedings of the 53rd Annual Congress Societa' Italiana Di Genetica Agraria*, 2009a;16-19 Septemeber, Torino.

[48] d'Aloisio, E., Dhanapal, A.P., Paolacci, A.R., Tanzarella. O.A., Porceddu, E., Ciaffi, M., "Typical PDI and its homologous in Wheat." *Proceedings of the International Durum Wheat Symposium*, June 30- July 3 2008a Bologna, Italy.

[49] Dhanapal, A.P., d'Aloisio, E., Porceddu, E., Ciaffi, M., "Genetic Variability in the Promoter Regions of Protein Disulfide Isomerase In Durum Wheat and Its Wild Relatives." *Proceedings of the 53rd Annual Congress Societa' Italiana Di Genetica Agraria*, 16-19 Septemeber 2009a Torino Italy.

[50] Dhanapal, A.P., d'Aloisio, E., Ciaffi, M., Porceddu, E., "Conserved regulatory motifs identified in Protein Disulfide Isomerase Promoter sequence analysis in wheat wild relatives." *Proceedings of the International Workshop on Plant Genetic Resources For Food And Agriculture-General aspects and research opportunities.* 5-6 November 2009b, Rome, Italy.

[51] Paolacci, A.R., Ciaffi, M., Dhanapal, A.P., Tanzarella. O.A., Porceddu, E., d'Aloisio, E., "Protein disulfide isomerase family in bread wheat (*Triticum aestivum* L): protein structure and expression analysis" *Plant Genetic Resources.* 2011; 9(2):342-346.

[52] Ciaffi, M., Paolacci, A.R., d'Aloisio, E, Dhanapal, A.P., Tanzarella, O.A., Porceddu, E., "Protein disulfide isomerase gene family in wheat:Genomic structure, phylogenetic and expression analyses". *Proceedings of the 54th Italian Society of Agricultural Genetics Annual Congress Matera*, 27-30 September, 2010, Italy.