DOI: 10.1007/s10535-016-0609-3 BIOLOGIA PLANTARUM 60 (3): 505-512, 2016

Molecular characterization and phylogenetic analysis of *Wx* genes from three *Taeniatherum* diploid species

S.F. DAI, J.Q. JIANG, Y.N. JIA, X.F. XUE, D.C. LIU, Y.M. WEI, Y.L. ZHENG, and Z.H. YAN*

Triticeae Research Institute, Sichuan Agricultural University, Chengdu 611130, Sichuan, P.R. China

Abstract

In wheat seeds, starch synthase I or the Waxy protein is an enzyme involved in amylose synthesis. The gene encoding this enzyme is *Wx* and in this study, eight novel *Wx* alleles were identified in three diploid *Taeniatherum* species. The variability of these alleles was evaluated, and their nucleotide sequences were compared with those of homologous alleles from wheat. Two types of *Taeniatherum Wx* alleles were detected in three diploid species *Ta. caput-medusae*, *Ta. asperum*, and *Ta. crinitum*. A phylogenetic analysis indicates that the *Taeniatherum Wx* alleles were more closely related to *Wx* alleles from *Aegilops* species with C, D, M, and U genomes than to *Wx* alleles of other species. These alleles represent a potential genetic resource that may be useful in wheat breeding programs.

Additional key words: amylose, starch, *Triticum aestivum*, Waxy protein.

Introduction

Starch is one of the most important components of wheat endosperm, accounting for 65 - 70 % on a dry mass basis. Starch is composed of two types of glucose polymers, the essentially linear amylose and the highly branched amylopectin. In wheat cultivars, the amount of amylose and amylopectin usually ranges from 22 to 35 % and from 68 to 75 %, respectively (Zeng *et al*. 1997). The amylose/amylopectin ratio is important for determining starch properties such as gelatinization, pasting, and gelation (Zeng *et al*. 1997). Wheat cultivars lacking amylose or with a low amylose content are used as foods and as ingredients in non-food industries (Morell *et al*. 1995).

Starch synthesis is a complex process involving a series of biosynthetic enzymes. Granule-bound starch synthase I (or the Waxy protein) is the key enzyme responsible for amylose synthesis in seeds (Nakamura *et al*. 1995). Bread wheat (*Triticum aestivum* L., 2n = 6x = 42, AABBDD) contains three different *Waxy* (*Wx)* gene loci, namely *Wx-A1*, *Wx-B1*, and *Wx-D1*, which are located on chromosomes 7AS, 4AL (translocated from 7BS), and 7DS, respectively (Chao *et al*. 1989, Ainsworth *et al*. 1993, Yamamori *et al*. 1994). Studies on the composition of Waxy proteins in diploid, tetraploid, and hexaploid wheat have revealed a low polymorphism in tetraploid and hexaploid wheat (Yamamori *et al*. 1995, Rodriguez-Quijano *et al*. 1998, Urbano *et al*. 2002, Caballero *et al*. 2008, Guzmán *et al*. 2009). In hexaploid wheat, eight different *Wx* combinations have been identified. Six of the eight alleles encode a partial Waxy protein that produces wheat with more than a 20 % amylose content. However, amylose content in a triple null mutant may be as low as 1 % (Yamamori *et al*. 1994, Nakamura *et al*. 1995). The absence of naturally occurring triple null alleles at all three *Wx* loci restricts generation of wheat cultivars with no amylose or a very low amylose content. Wheat lines with amylose content between 1 and 20 % have been produced by combining two null *Wx* alleles and one allele with a reduced activity or expression (Yamamori *et al*. 1994, Nakamura *et al*. 1995, Yamamori 2009, Yamamori and Yamamoto 2011).

Novel *Wx* alleles with different expression patterns or those encoding enzymes with altered activity may be used to produce wheat cultivars with a lower amylose content. A rich diversity of *Wx* alleles has been found in some ancient wheat species, and some new *Wx* alleles have been characterized at the molecular level (Yan *et al*. 2000, Yan and Bhave 2001, Guzmán *et al*. 2012b, 2015, Ortega *et al*. 2014b). The *Wx* alleles have been mainly characterized in *Aegilops* species (Ortega *et al*. 2014a).

Submitted 21 October 2015, *last revision* 5 January 2016, *accepted* 18 January 2016.
Abbreviations: pI - isoelectric point; MYA - million years ago; *Wx* - *waxy*.
Acknowledgments: This research was supported by the National Natural Science Foundation of China (U1403185) and the Key Fund Project of the Sichuan Provincial Department of Education (15ZA0021).
* Corresponding author; fax: (+86) 2882650350, e-mail: zhyan104@163.com

The *Aegilops* genus is closely related to hexaploid wheat (Dvořák *et al*. 1998, Maestra and Naranjo 1998). *Aegilops tauschii* (2n = 2x = 14, DD) provides the D genome (Kihara 1944, Dvořák *et al*. 1998), whereas *Ae. speltoides* is most likely the species that contributes the B genome (Maestra and Naranjo 1998).

In addition to their potential use in breeding wheat lines with modified starch properties, complete or partial *Wx* sequences have been used to study the origins and evolutionary relationships of *Triticeae* species (Mason-Gamer 1998, 2001, Yan and Bhave 2001, Ingram and Doyle 2003, Guzmán *et al*. 2012a). There are many relatives of bread wheat within the *Triticeae*. The *Wx* genes of the *Triticum* and *Aegilops* genera have been characterized, but their orthologs in other *Triticeae* genera have not. *Taeniatherum* (TaTa, 2n = 2x = 14) contains three diploid species, *Ta. caput-medusae*, *Ta. crinitum*, and *Ta. asperum* (Frederiksen 1986, Frederiksen and Von Bothmer 1986). The *Taeniatherum* species are distantly related to wheat (Frederiksen and Von Bothmer 1986). To date, there have been no reports on the molecular characterization of *Wx* genes from these wheat relatives. The aim of the present study was to isolate and characterize *Wx* alleles from three diploid *Taeniatherum* species and analyze their phylogenetic relationships with the *Wx* alleles of other *Triticeae* species.

Materials and methods

Plants: Eight accessions from three *Taeniatherum* diploid species, *Ta. caput-medusae* Nevski, *Ta. asperum* Nevski, and *Ta. crinitum* Nevski, were used (Table 1 Suppl). Seven of the eight accessions originated from Turkey, and one (PI 220590) originated from Afghanistan. These accessions were kindly provided by the Agricultural Research Service of the United States Department of Agriculture (http://www.ars-grin.gov/).

DNA isolation and *Wx* cloning: Approximately 2 g of young seedling tissue was frozen in liquid nitrogen and ground to a fine powder. Deoxy-ribonucleic acid was extracted using cetyltrimethylammonium bromide as described previously (Doyle and Doyle 1990). The complete *Wx* sequence is about 2.8 kb long and includes 10 introns and 11 exons. The full-length *Wx* sequence was amplified with 0.6-, 1.2-, and 1.2-kb overlapping fragments, which corresponded to exons 1 to 3, 3 to 6, and 6 to 11, respectively. Primers used for PCR and thermal cycling conditions were the same as those used by Ortega *et al*. (2014a). Primers used to amplify the 0.6-, 1.2-, and 1.2-kb fragments were Wx1Fw/Wx1Rv, Wx2Fw/Wx2Rv, and Wx3Fw/Wx3Rv, respectively.

The amplicons were separated on 0.8 % (m/v) agarose gels and purified using a *DP1602* DNA recovery kit (*Bioteke*, Beijing, China). The PCR products were then inserted into *pMD18-T* vectors (*Takara*, Dalian, China) to generate target plasmids. After transforming plasmids into *Escherichia coli* strain DH10B cells, three positive clones were isolated for sequencing. Sequencing was performed by the *Takara Biotechnical Company* (Dalian, China). Full-length *Wx* sequences were assembled by aligning the three overlapping fragments.

Data analysis: Sequencing data were analyzed using *DNAsp 5.10.01* (Librado and Rozas 2009). This program was also used to calculate nucleotide parameters, including the total number of mutations, the average number of nucleotide differences, and the number of polymorphic sites. We also calculated nucleotide diversity parameters such as theta (θ), the average number of nucleotide differences per site between two sequences (π), the substitution rates of synonymous (Ks) and non-synonymous (Ka) sites, and the Ka/Ks ratio. A neutrality test was performed using Tajima's D statistics (Tajima 1989). Divergence times were evaluated based on the average time of 2.7 million years ago (MYA) for the A and D genomes (Dvořák and Akhunov 2005). Amino acid differences between the deduced *Taeniatherum* Wx (Wx-Ta) proteins and reference proteins were determined by *SIFT Blink* analysis (Sim *et al*. 2012; http://sift.jcvi.org/www/SIFT_BLink_submit.html). Protein molecular masses and isoelectric points (pIs) were calculated using the Compute pI/Mw tool (http://web.expasy.org/compute_pi/).

Phylogenetic analysis: We used *MEGA 5* to construct phylogenetic trees (Tamura *et al*. 2011). A total of 21 representative *Wx* genes from the A, B, C, D, S, M, and U genomes of diploid *Aegilops* species and hexaploid wheat were included in the analyses. The genes included *Wx-A1* from hexaploid wheat, *Triticum urartu*, and *T. monococcum* (GenBank: AB019622, KF612976, and KF612977, respectively), *Wx-B1* from hexaploid wheat (AB019623 and EU719610), *Wx-D1* from hexaploid wheat and *Ae. tauschii* (AB019624 and EU719612, respectively), *Wx-C1* from *Ae. markgrafii* (JX679010 and JX679019), *Wx-M1* from *Ae. comosa* (JX402790 and JX679003), and *Wx-S1* from S-genome diploid species *Ae. longissima* (JX679019 and JX679018), *Ae. searsii* (JX679011 and JX679012), *Ae. speltoides* (JX679013 and JX679014), and *Ae. sharonensis* (JX679016 and JX679017).

A neighbor-joining tree was constructed using the maximum composite likelihood method and the bootstrap value was estimated based on 1 000 replications.

Results

Complete *Wx* genes were constructed using three overlapping PCR fragments from eight accessions of *Ta. caput-medusae*, *Ta. asperum*, and *Ta. crinitum* (Fig. 1).

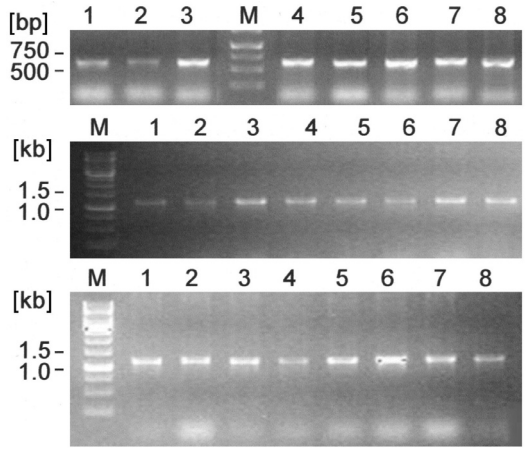

Fig. 1. The complete coding region (2.8 kb) of *Wx* alleles from *Taeniatherum* species amplified by PCR. The coding region was assembled using 0.6- (*A*), 1.2- (*B*), and 1.2-kb (*C*) fragments corresponding to the start of open reading frames to exon 3, exon 2 to exon 6, and exon 6 to the end of the open reading frame, respectively. *Lanes* 1 - 8 represent accessions PI204577, PI220590, PI561091, PI561092, PI561094, PI577708, PI577710, and PI598389, respectively, M - molecular mass markers.

The sizes of the amplified fragments for each accession were as expected (Fig. 1*A*-*C*). The full-length gene sequences were determined by sequencing at least three individual overlapping clones. We obtained eight different *Wx* alleles, which varied in length from 2 777 to 2 808 bp. The alleles were designated *Wx-Ta1* to *Wx-Ta8*, and their sequences were deposited in the National Center for Biotechnology Information (*NCBI*) database under the accession numbers KT878862 to KT878869 (Table 1 Suppl).

Sequences for the eight Wx-Ta proteins were predicted based on the corresponding exon sequences. There were 534 (Group I *Wx-Ta* alleles) or 535 (Group II *Wx-Ta* alleles) amino acids in the mature proteins and 70 (Group I *Wx-Ta* alleles) or 71 (Group II *Wx-Ta* alleles) amino acids in the transit peptides (Table 1 Suppl.). The molecular masses of the mature Wx-Ta proteins were 58.60 - 58.69 kDa. They were smaller than the three hexaploid wheat Wx proteins Wx-A1 (58.95 kDa), Wx-B1 (58.80 Da), and Wx-D1 (58.90 kDa). The Wx-Ta proteins were more acidic (pI 5.16 or 5.20) than the Wx-A1, Wx-B1, and Wx-D1 proteins (pI 5.57, 5.44, and 5.60, respectively).

Variations in nucleotide sequences resulted in some amino acid changes. A comparison of the Wx-D1 and Wx-Ta amino acid sequences identified a total of 27 differences with an average of 3.5 different residues per Wx protein (Table 2 Suppl.). Six of the 27 amino acid substitutions were in the transit peptide within the first 70 (group I *Wx-Ta* alleles) or 71 (group II *Wx-Ta* alleles) amino acids. The remaining 21 differences were located randomly throughout the mature protein. The Wx proteins were subjected to *SIFT Blank* analysis to determine how the amino acid sequence differences affected protein function (Sim *et al*. 2012). Two amino acid substitutions with probable adverse effects on protein activity were present in *Wx-Ta1*, *Wx-Ta3*, *Wx-Ta4*, *Wx-Ta6*, *Wx-Ta7*, and *Wx-Ta8*. The Ser to Gly change at a position 330 in *Wx-Ta1* and the Thr to Ile change at a position 444 in *Wx-Ta3*, *Wx-Ta4*, *Wx-Ta6, Wx-Ta7*, and *Wx-Ta8* showed a probability of 0.00, suggesting that both mutations will affect protein activity (Table 2 Suppl., Ortega *et al*. 2014b).

The exons and introns were identified by comparisons with *Wx* alleles from hexaploid wheat (Murai *et al*. 1999) and sequences from the *NCBI* database. The *Wx-Ta* alleles were 88.0 - 95.0 % identical to those in the *NCBI* database with a sequence coverage of 37 - 100 % (data not shown). All of the *Wx-Ta* alleles had 11 exons and 10 introns, which were similar to those of hexaploid wheat (Table 1 Suppl., Murai *et al*. 1999). All of the exons in *Wx-Ta* were the same as those in the hexaploid wheat *Wx* alleles except for exon 1. It contained a 3-bp insertion (in *Wx-Ta3,-Ta4,-Ta6,* and *-Ta8*) / deletion (in *Wx-A1, Wx-B1, Wx-D1, Wx-Ta1,-Ta2,-Ta5,* and *-Ta7*) indels (Fig. 1 Suppl.). All of the introns differed among

Table 1. Polymorphism of DNA and test statistics for eight *Wx*s of *Taeniatherum* species. The *Wx*s in group I included *Wx-Ta1, -Ta2, -Ta5*, and *-Ta7*; the remaining four *Wx*s (*Wx-Ta3, -Ta4, -Ta6, -Ta8*) were included in group II (η - total number of mutations; *k* - average number of nucleotide differences; s - number of polymorphic sites; h - number of haplotypes; θ - Watterson's estimate; π - nucleotide diversity; *D* - Tajima's estimate *D*-test; ns - not significant).

Gene	*n*	Completely nucleotide sequence							Coding region						
		η	k	s	h	$\theta\times10^{-3}$	$\pi\times10^{-3}$	*D*	η	k	s	h	$\theta\times10^{-3}$	$\pi\times10^{-3}$	*D*
Group I	4	57	28.8	56	4	11.1	10.3	-0.757 ns	23	11.7	22	4	6.9	6.4	-0.720 ns
Group II	4	17	9.0	17	4	3.3	3.2	-0.300 ns	11	6.0	11	4	3.3	3.3	0 ns
Overall	8	163	84.1	162	8	22.7	30.4	1.847 ns	58	28.3	57	8	12.3	15.6	1.428 ns

the *Wx-Ta* alleles and between the *Wx-Ta* alleles and hexaploid wheat *Wx* alleles. Based on the exon and intron sequences, the eight *Wx-Ta* alleles were classified into two groups: Group I (*Wx-Ta1*, *Wx-Ta2*, *Wx-Ta5*, and *Wx-Ta7*) and Group II (*Wx-Ta3*, *Wx-Ta4*, *Wx-Ta6*, and *Wx-Ta8*). The Group I allele sequences were very similar except for introns 5 (101 or 102 nucleotides) and 10 (105 or 102 nucleotides). A comparison of the sequences from the two groups revealed significant differences, particularly in exon 1 and introns 1 - 5. All the exons were the same size except for exon 1, which had a 3-bp indel in the Group II sequences. The sizes of the introns varied considerably (Table 1 Suppl.). The group I introns were longer than those of *Wx-A1* and *Wx-B1* but shorter than those of *Wx-D1*. The group II introns were shorter than those of all three *Wx* alleles of hexaploid wheat (Table 1 Suppl.).

Sequence similarity coefficients among the eight *Wx-Ta* alleles were calculated (Table 3 Suppl.). The coefficients based on the complete nucleotide, exon, intron, and amino acid sequences were 94.8 - 99.8 %, 97.3 - 99.8 %, 89.2 - 99.6 %, and 98.3 - 99.8 %, respectively, with an average of 97.0, 98.5, 93.9, and 99.0 %, respectively. These results indicate that sequence diversities of the *Wx* regions were as follows: intron sequences > complete nucleotide sequence > exon sequences (nucleotides > amino acids). Additionally, the within-group sequence similarity coefficients were higher than the between-group coefficients.

We evaluated *Wx-Ta* DNA polymorphisms using the complete sequences (*i.e.,* the exons and introns) and the coding regions (*i.e.,* the exons only) (Table 1). Analysis of the complete sequences of the 8 accessions revealed a high polymorphism (163 mutations at 162 polymorphic sites). The coding regions also exhibited a high polymorphism but had fewer mutations (58) at 57 polymorphic sites. This result indicates that there was more variation in the introns than in the complete sequences.

Table 2. Variations among *Wx* genes from *Taeniatherum* species and estimated divergence times. Ks - *Wx-Ta* group I (*Wx-Ta1, -Ta2, -Ta5,* and *-Ta7*) = 0.2167 ± 0.0041; Ka - *Wx-Ta* group II (*Wx-Ta3, -Ta4, -Ta6,* and *-Ta8*) = 0.2366 ± 0.0009; MYA - Wx-Ta group I/ Wx-Ta group II = 0.37. * - Divergence rate of 0.0533 synonymous substitutions per MYA calculated according to Dvořák and Akhumov (2005), MYA = Ks/0.533 × 10^{-9}.

	Related to *Wx-A1* (AB019622)				Related to *Wx-B1* (AB019623)				Related to *Wx-D1* (AB019624)			
	Ks	Ka	Ka/Ks	MYA*	Ks	Ka	Ka/Ks	MYA*	Ks	Ka	Ka/Ks	MYA*
Wx-Ta1	0.2144	0.0338	0.1576	4.02	0.1544	0.0192	0.1244	2.90	0.1286	0.0132	0.1026	2.41
Wx-Ta2	0.2143	0.0323	0.1507	4.02	0.1544	0.0192	0.1244	2.90	0.1286	0.0117	0.0910	2.41
Wx-Ta3	0.2374	0.0332	0.1398	4.45	0.1463	0.0165	0.1128	2.74	0.1362	0.0125	0.0918	2.56
Wx-Ta4	0.2374	0.0324	0.1365	4.45	0.1464	0.0165	0.1127	2.75	0.1371	0.0118	0.0861	2.57
Wx-Ta5	0.2143	0.0331	0.1545	4.02	0.1544	0.0184	0.1192	2.90	0.1294	0.0125	0.0966	2.43
Wx-Ta6	0.2353	0.0317	0.1347	4.41	0.1455	0.0165	0.1134	2.73	0.1345	0.0110	0.0818	2.52
Wx-Ta7	0.2239	0.0317	0.1416	4.20	0.1595	0.0198	0.1241	2.99	0.1353	0.0125	0.0924	2.54
Wx-Ta8	0.2364	0.0324	0.1371	4.44	0.1472	0.0156	0.1060	2.76	0.1337	0.0118	0.0883	2.51
Wx-Ta				4.25±0.19				2.83±0.09				2.49±0.06

In the multiple sequence alignments, the eight *Wx-Ta* alleles were equally divided into two groups (Fig. 1 Suppl.). Compared with the *Wx* alleles in Group II, the *Wx* alleles in Group I had more polymorphic sites in their complete nucleotide sequences and coding regions (Table 1). The variability of the complete nucleotide sequences was approximately 2.8-fold greater than that of the coding regions for all sequences investigated. The nucleotide sequences of the Group I *Wx-Ta* alleles were more diverse than those of the Group II *Wx-Ta* alleles according to the parameters π and θ (Table 2). The Tajima's *D*-test results are not significant for any case, suggesting that the level of polymorphism fitted a neutral equilibrium model.

A maximum likelihood phylogenetic tree was constructed using the complete sequences of *Wx-Ta* and representative *Wx* alleles from *Triticum* and *Aegilops* species (Fig. 2). All the *Wx* alleles were divided into five clades with high bootstrap values. The *Wx-Ta* alleles formed an independent clade (*Wx-Ta*) that paralleled the *Wx* alleles from the C, U, M, and D genomes of *Ae. markgrafii, Ae. umbellulata*, *Ae. comosa*, *Ae. tauschii*, and the D genome of hexaploid wheat (Clade I). Within the clade containing eight *Wx-Ta* alleles, there were two parallel inner branches, named *Wx-Ta* type I and *Wx-Ta* type II, each containing four *Wx* alleles. These two parallel inner branches were supported by high bootstrap values. All *Wx* alleles from the B genome of hexaploid wheat and the S genome of related diploid species were located in two separate clades (clades II and III). Clade II consisted of the *Wx* alleles from *Ae. searsii*, *Ae. speltoides*, and the B genome of hexaploid wheat. Clade III comprised the *Wx* alleles from *Ae. longissima* and *Ae. sharonensis*. Clade III was located at the basal layer of three parallel clades of *Wx* alleles (clades Wx-Ta, and I–II). The clade Wx-Ta comprised *Wx-Ta* alleles from *Taeniatherum* species; clade I consisted of *Wx* alleles from *Aegilops* species with C, U,

D, and M genomes, and from hexaploid wheat; and clade II consisted of *Wx* alleles from two *Aegilops* species with the S genome (*Ae. searsii* and *Ae. speltoides*) and the B genome of hexaploid wheat. Clade IV included the *Wx* alleles from the A genome species (*i.e.*, hexaploid wheat with the A genome, *T. urata,* and *T. monococcum*) and was located at the outermost layer.

The Ks and Ka substitution rates among the *Wx-Ta*, *Wx-A1*, *Wx-B1*, and *Wx-D1* genes are shown in Table 2. The calculations for these rates were as follows: divergence rate = 5.33×10^{-8} synonymous substitutions/ year; Ks value = 0.144, corresponding to the divergence rate between *Wx-A1* and *Wx-D1* (Guzmán *et al*. 2012); and divergence time = 2.7 MYA for the separation between the A and D genomes (Dvořák and Akhunov 2005). The results show that the divergence times for the *Wx-Ta* alleles and *Wx-A1*, *Wx-B1*, and *Wx-D1* were 4.02 - 4.45, 2.73 - 2.90, and 2.41 - 2.57 MYA, with an average of 4.25, 2.83, and 2.49 MYA, respectively. The divergence times for the group I and group II *Wx-Ta* alleles was about 0.37 MYA.

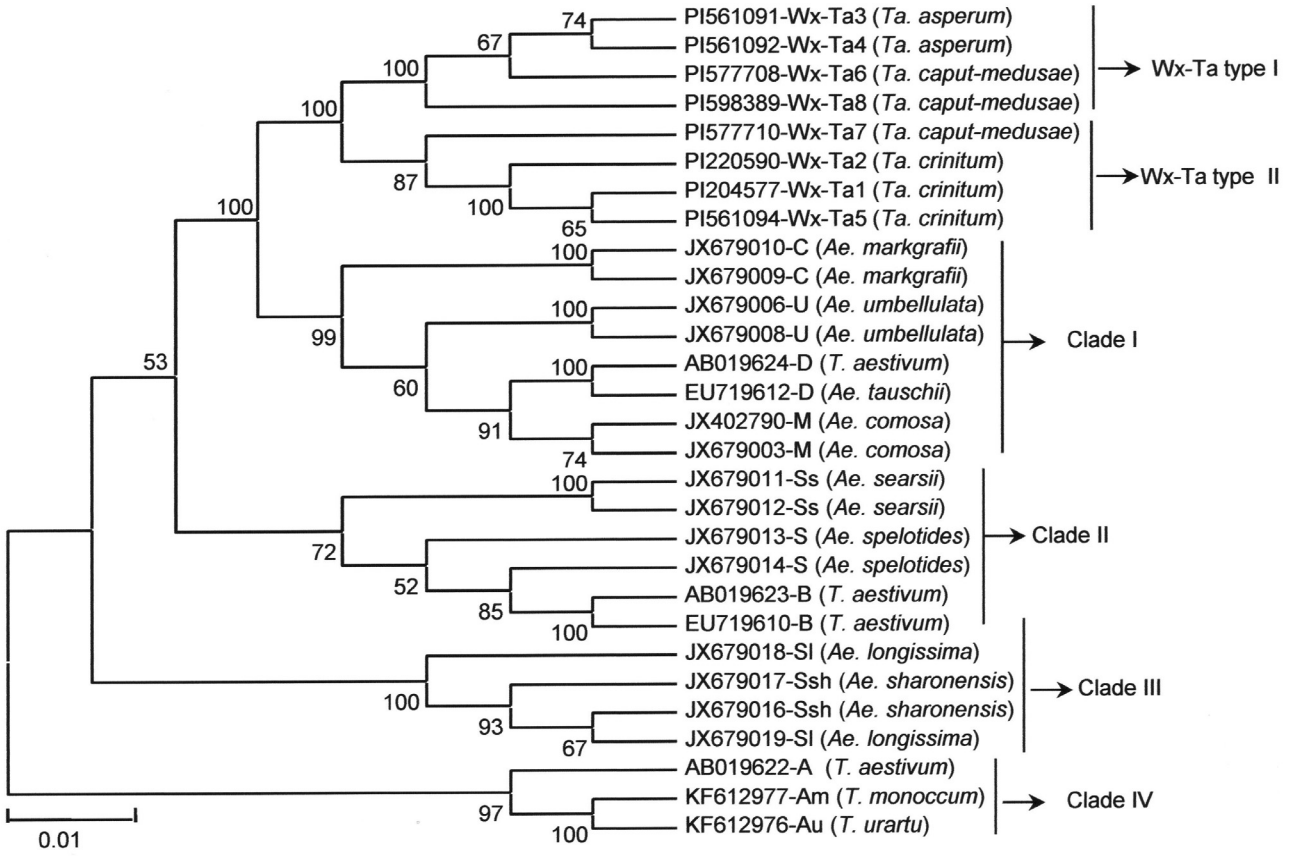

Fig. 2. A phylogenetic tree including *Wx* alleles from *Taeniatherum* and *Aegilops* species, and from hexaploid wheat. The *Wx* alleles from *Aegilops* (including *Ae. markgrafii*, *Ae. comosa*, *Ae. searsii*, *Ae. speltoides*, and *Ae.sharonensis*), *T. aestivum* and A-genome diploid species (*T. monococcum* and *T. urartu*) were reported in Ortega *et al*. (2014a,b) and Murai *et al*. (1999). The other two *Wx* alleles EU719612-D and EU719610-B were obtained from the GenBank database (unpublished data).

Discussion

The *Wx* genes and their encoded Waxy proteins have been extensively studied in wheat and related *Triticum* and *Aegilops* species because of their important effects on starch properties. Researchers have also characterized evolutionary relationships among species based on various *Wx* genes (Mason-Gamer *et al*. 1998, Murai *et al*. 1999, Yan and Bhave 2001, Guzmán *et al*. 2012a,b, Ortega *et al*. 2014a,b). Low polymorphisms in *Wx* genes were detected in tetraploid and hexaploid wheat (Yamamori *et al*. 1995, Rodriguez-Quijano *et al*. 1998, Urbano *et al*. 2002, Guzmán *et al*. 2009, Guzmán and Alvarez 2012). In contrast, the *Wx* genes of some ancient wheat species and their wild relatives show a considerable diversity (Caballero *et al*. 2008, Ortega *et al*. 2014a,b), and are therefore considered to be genetic resources with a potential to broaden diversity of wheat *Wx* genes. *Taeniatherum* species are distantly related wild relatives of wheat (Frederiksen and Von Bothmer 1986).

In this study, we amplified and sequenced eight *Wx* genes from three diploid species, *Ta. caput-medusae*, *Ta. asperum*, and *Ta. crinitum*. Sequencing analyses show that the eight *Wx* alleles had different nucleotide lengths ranging from 2 777 bp to 2 808 bp. The *Wx* alleles showed differences in length in *Ta. caput-medusae* (*Wx-Ta6* and *Wx-Ta8* were 2 777 bp and *Wx-Ta7* was 2 806 bp) but not in *Ta. asperum* (*Wx-Ta3* and *WxTa4* were 2 777 bp) or *Ta. crinitum* (*Wx-Ta1*, *Wx-Ta2*, and *Wx-Ta5* were 2 808 bp). The *Wx-Ta* alleles were significantly different from the *Wx* alleles of hexaploid wheat in terms of complete gene length and intron length. Additionally, the *Wx-Ta* alleles showed some nucleotide substitutions and indels compared with the *Wx* alleles of hexaploid wheat.

Based on gene length, structure, and phylogenetic analyses (Table 1 Suppl. and Fig. 2), the eight *Wx-Ta* genes formed two groups, each group forming a separate sub-clade (*Wx-Ta* type I and *Wx-Ta* type II), with high bootstrap values. The *Wx* alleles from *Ta. asperum* and *Ta. crinitum* were in two separate sub-clades, whereas the *Ta. caput-medusae Wx* alleles were present in both sub-clades. Because the limited number of accessions used in this study had very similar origins, it is unclear whether the *Ta. asperum* and *Ta. crinitum* alleles would still be distributed in separate sub-branches if more accessions were included. Therefore, it is most likely that the *Wx-Ta* alleles form two groups independent of species. The phylogenetic analyses of *Wx-B1* from polyploid species and *Wx* alleles from the *Sitopsis* section of *Aegilops* (Ortega *et al*. 2014a) produced similar results. The *Wx-B1* gene formed two groups: *Wx-B1* type I and *Wx-B1* type II (Guzmán *et al*. 2012b). The *Wx-B1* type I consisted of *Wx* alleles from emmer (*T. turgidum* ssp. *dicoccum*), common wheat, and spelt (*T. turgidum* ssp. *dicoccum*), accessions originating from Asia and Europe. The *Wx-B1* type II group included only *Wx* alleles from Spanish spelt and the durum wheat cv. Langdon. The *Wx* alleles from the *Sitopsis* section of different S genomes also formed two independent branches. One contained *Wx* alleles from *Ae. searsii* and *Ae. speltoides*, whereas the other included *Wx* alleles from *Ae. sharonensis* and *Ae. longissima* (Ortega *et al*. 2014a). These results suggest that the *Wx* genes within these genomes diverged. For example, the *Wx-B1* type I alleles separated from the *Wx-B1* type II alleles approximately 0.6 MYA (Guzmán *et al*. 2012b), at the same time as the divergence of the *Wx-S1* allele from *Ae. speltoides* (Ortega *et al*. 2014a).

The nucleotide sequence diversity of *Wx* alleles in *Taeniatherum* species was much higher than that of *Wx* alleles in *T. urartu* and Einkorn wheat (Ortega *et al*. 2014a) but lower than that of *Wx* alleles in *Aegilops* species (Liu *et al*. 2009; Ortega *et al*. 2014b). *Taeniatherum* species are strictly self-pollinators, in which reproductive isolation occurs extensively among different species or different accessions of the same species (Frederiksen and Von Bothmer 1989). This restriction results in separate *Taeniatherum* gene pools. *T. urartu* is more closely related to *T. aestivum* than to *T. monococcum* and *Taeniatherum* species. Additionally, the agricultural practices used to cultivate these species may have reduced genetic variability through genetic drift processes. *Aegilops* species harbor a number of genomes (Kilian *et al*. 2011), which might lead to a greater diversity than in single-genome *Taeniatherum* species.

Previous studies have shown that *Wx* exons are consistently the same size, but there is a greater variability in intron size (Li *et al*. 2012, Ortega *et al*. 2014). Previous studies have shown that there are a number of indels in the *Wx* introns of several *Aegilops* species (Ortega *et al*. 2014a) and hexaploid wheat (Murai *et al*. 1999, Yan *et al*. 2000). In this study, the *Wx-Ta* intron size was conserved within the same group but differed between the groups I and II. Additionally, there were differences in length between the *Wx-Ta* introns and those of the *Wx* alleles of hexaploid wheat (Table 3 Suppl.) and *Aegilops* species (data no shown). However, the *Wx-Ta* alleles showed close relationships with the *Wx* alleles from the D genome of hexaploid wheat and the C, U, M, and D genomes of *Aegilops* species (Fig. 2, Ortega *et al*. 2014a). This suggests that the origin of the *Wx-Ta* alleles differs from those of the *Wx* alleles of hexaploid wheat and *Aegilops* species. Furthermore, the *Wx-Ta* alleles diverged from the *Wx-A1, Wx-B1*, and *Wx-D1* alleles of hexaploid wheat at different times. The *Wx-Ta* alleles first diverged from *Wx-A1* approximately 4.25 MYA, then diverged from *Wx-B1* about 2.83 MYA, and finally separated from *Wx-D1* about 2.49 MYA.

In the phylogenetic analyses, the Wx-Ta proteins were more similar to Wx-D1 than to Wx-A1 and Wx-B1 of the hexaploid wheat cv. Chinese Spring. The comparison of the Wx-Ta and Wx-D1 proteins revealed 27 amino acid substitutions in Wx-Ta. Although some of the substitutions were located in the transit peptides, most were located in the mature protein. Ortega *et al*. (2014a) predicted that two substitutions (Arg to Met at a position 250 and Thr to Ile at a position 538) in the *Ae. searsii* Wx protein would have deleterious effects on protein activity. In this study, we also identified two substitutions that are likely to affect adversely protein activity. The effects of the Ser to Gly and Thr to Ile substitutions on protein function should be investigated in greater detail in future studies.

Several phylogenetic analyses have been conducted using *Wx* gene sequences (Mason-Gamer 1998, 2001, Yan *et al*. 2000, Ingram and Doyle 2003, Ortega *et al*. 2014a). For example, phylogenetic relationships among the *Aegilops* and *Triticum* species were investigated using *Wx* gene sequences (Yan *et al*. 2000, Guzmán *et al*. 2012a, Ortega *et al*. 2014a) and their 5′-untranslated regions (Li *et al*. 2012). The results of the phylogenetic analysis in this study are consistent with those of previous phylogenetic analyses based on the sequences of low molecular mass glutenin subunits (Wang *et al*. 2011)

and repeated nucleotide sequences (Dvořák and Zhang 1992) although there were some differences. In this study, phylogenetic relationships among *Taeniatherum*, *Aegilops*, and *Triticum* species were assessed using the *Wx* gene sequences. The results suggest that *Taeniatherum* species are more closely related to *Aegilops* species with C, D, M, and U genomes than to *Aegilops* species with S genomes in the *Sitopsis* section or to *Triticum* species with A and B genomes. Our phylogenetic results are also consistent with previous findings on species relationships based on comparisons of conserved domain sequences of high molecular mass glutenin subunits although there were some minor differences (Dai *et al*. 2013).

In summary, eight novel *Wx* genes were isolated and characterized from three diploid *Taeniatherum* species. The intron sizes in the eight new *Wx* genes were significantly different from those in the *Wx* alleles of *Aegilops* and *Triticum* species, but the exon sizes were conserved. The predicted Wx-Ta proteins were smaller and more acidic than the Wx proteins from hexaploid wheat. The new *Wx* genes showed some variations in nucleotide sequences, and some of these changes resulted in amino acid substitutions in the transit peptides and mature proteins. It is likely that at least two of these changes will modify protein function leading to synthesis of a novel kinds of starch. Expression and processing these novel *Wx* alleles should be studied in a modern wheat background *via* hybridization or genetic transformation. These alleles represent a new genetic resource that may be relevant to wheat breeding. The *Wx-Ta* alleles were divided into two groups based on gene length, domain structure, and phylogenetic characterristics. The *Wx-Ta* alleles differed from the *Wx* alleles of *Aegilops* and *Triticum* species and showed the closest relationships to those of *Aegilops* species with C, D, U, and M genomes. These results suggest that the origins of *Wx-Ta* alleles differ from those of the *Wx* alleles in *Aegilops* and *Triticum* species.

References

Ainsworth, C., Clark, J., Balsdon, J.: Expression, organisation and structure of the genes encoding the waxy protein (granule-bound starch synthase) in wheat. - Plant. mol. Biol. **22**: 67-82, 1993.

Caballero, L., Bancel, E., Debiton, C., Branlard, G.: Granule bound starch synthase (GBSS) diversity of ancient wheat and related species. - Plant. Breed. **127**: 548-553, 2008.

Chao, S., Sharp, P.J., Worland, A.J., Warham, E.J., Koebner, R.M.D., Gale, M.D.: RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. - Theor. appl. Genet. **78**: 495-504, 1989.

Dai, S.F., Pu, Z.J., Liu, D.C., Wei, Y.M., Zheng, Y.L., Hu, X.K., Yan, Z.H.: Characterization of novel HMW-GS in two diploid species of *Eremopyrum*. - Gene **519**: 55-59, 2013.

Doyle, J.J., Doyle, J.L.: Isolation of plant DNA from fresh tissue. - Focus **12**: 13-15, 1990.

Dvořák, J., Akhunov, E.D.: Tempos of gene locus deletions and duplications and their relationship to recombination rate during diploid and polyploid evolution in the *Aegilops-Triticum* alliance. - Genetics **171**: 323-332, 2005.

Dvořák, J., Luo, M.C., Yang, Z.L., Zhang, H.B.: The structure of the *Aegilops tauschii* gene pool and the evolution of hexaploid wheat. - Theor. appl. Genet. **97**: 657-670, 1998.

Dvořák, J., Zhang, H.B.: Reconstruction of the phylogeny of the genus *Triticum* from variation in repeated nucleotide sequences. - Theor. appl. Genet. **84**: 419-429, 1992.

Frederiksen, S.: Revision of *Taeniatherum* (*Poaceae*). - Nordic J. Bot. **6**: 389-397, 1986.

Frederiksen, S., Von Bothmer, R.: Intergeneric hybridization between *Taeniatherum* and different genera of *Triticeae*, *Poaceae*. - Nordic J. Bot. **9**: 229-240, 1989.

Frederiksen, S., Von Bothmer, R.: Relationships in *Taeniatherum* (*Poaceae*). - Can. J. Bot. **64**: 2343-2347, 1986.

Guzmán, C., Alvarez, J.B.: Molecular characterization of a novel waxy allele (Wx-Au1a) from *Triticum urartu* Thum. *ex* Gandil. - Genet. Resour. Crop Evol. **59**: 971-979, 2012.

Guzmán, C., Caballero, L., Alvarez, J.B.: Variation in Spanish cultivated einkorn wheat (*Triticum monococcum* L. ssp. *monococcum*) as determined by morphological traits and waxy proteins. - Genet. Resour. Crop Evol. **56**: 601-604, 2009.

Guzmán, C., Caballero, L., Martín, L.M., Alvarez, J.B.: Waxy genes from spelt wheat: new alleles for modern wheat breeding and new phylogenetic inferences about the origin of this species. - Ann. Bot. **110**: 1161-1171, 2012a.

Guzmán, C., Caballero, L., Yamamori, M., Alvarez, J.B.: Molecular characterization of a new waxy allele with partial expression in spelt wheat. - Planta **235**: 1331-1339, 2012b.

Guzmán, C., Ortega, R., Yamamori, M., Pena, R.J., Alvarez, J.B.: Molecular characterization of two null Waxy alleles in Mexican Bread wheat landraces. - J. Cereal Sci. **62**: 8-14, 2015.

Ingram, A.L., Doyle, J.J.: The origin and evolution of *Eragrostis tef* (*Poaceae*) and related polyploids: evidence from nuclear waxy and plastid rps16. - Amer. J. Bot. **90**: 116-122, 2003.

Kihara, H.: Discovery of the DD analyser, one of the ancestors of *Triticum vulgare*. - Agr. Hort. **19**: 13-14, 1944.

Kilian, B., Mammen, K., Millet, E., Sharma, R., Graner, A., Salamini, F., Hammer, K., Özkan, H.: *Aegilops*. - In: Kole, C. (ed.): Wild Crop Relatives: Genomic and Breeding Resources: Cereals. Pp. 1-76, Springer, Berlin - Heidelberg 2011.

Li, W., Gao, Z., Xiao, W., Wei, Y.M., Liu, Y.X., Chen, G.Y., Pu, Z.E., Chen, H.P., Zheng, Y.L.: Molecular diversity of restriction enzyme sites, Indels and upstream open reading frames (uORFs) of 5′ untransalted regions (UTRs) of *Waxy* genes in *Triticum* L. and *Aegilops* L. species. - Genet. Resour. Crop Evol. **59**: 1625-1647, 2012.

Librado, P., Rozas, J.: *DnaSP v5*: a software for comprehensive

analysis of DNA polymorphism data. - Bioinformatics **25**: 1451-1452, 2009.

Liu, Y.X., Li, W., Wei, Y.M., Chen, G.Y., Zheng, Y.L.: Molecular characterization of the *Waxy* gene in einkorn wheat. - J. Plant Sci. **4**: 114-121, 2009.

Maestra, B, Naranjo, T.: Homoeologous relationships of *Aegilops speltoides* chromosomes to bread wheat. - Theor. appl. Genet. **97**: 181-186, 1998.

Mason-Gamer, R.J.: Origin of North American *Elymus* (*Poaceae*: *Triticeae*) allotetraploids based on granule bound starch synthase gene sequences. - Syst. Bot. **26**: 757-768, 2001.

Mason-Gamer, R.J., Weil, C.F., Kellogg, E.A.: Granule-bound starch synthase: structure, function, and phylogenetic utility. - Mol. biol. Evol. **15**: 1658-1673, 1998.

Morell, M.K., Rahman, S., Alrahams, S.L., Appeles, R.: The biochemistry and molecular biology of starch synthesis in cereals. - Aust. J. Plant Physiol. **22**: 647-660, 1995.

Murai, J., Taira, T., Ohta, D.: Isolation and characterization of the three *Waxy* genes encoding the granule-bound starch synthase in hexaploid wheat. - Gene **234**: 71-79, 1999.

Nakamura, T., Yamamori, M., Hirano, H., Hidaka, S.: Identification of three Wx proteins in wheat (*Triticum aestivum* L.). - Biochem. Genet. **31**: 75-86, 1993.

Nakamura, T., Yamamori, M., Hirano, H., Hidaka, S.: Production of waxy (amylose-free) wheats. - Mol. gen. Genet. **248**: 253-259, 1995.

Ortega, R., Alvarez, J.B., Guzmán, C.: Characterization of the *Wx* gene in diploid *Aegilops* species and its potential use in wheat breeding. - Genet. Resour. Crop Evol. **61**: 369-382, 2014a.

Ortega, R., Guzmán, C., Alvarez, J.B.: *Wx* gene in diploid wheat: molecular characterization of five novel alleles from enikorn (*Triticum monococcum* L. ssp. *monococcum*) and *T. urartu*. - Mol. Breed. **34**: 1137-1146, 2014b.

Rodriguez-Quijano, M., Nieto-Taladriz, M.T., Carrillo, J.M. Polymorphism of waxy proteins in Iberian hexaploid wheats. - Plant. Breed. **117**: 341-344, 1998.

Sim, N.L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., Ng, P.C.: SIFT web server: predicting effects of amino acid substitutions on proteins. - Nucl. Acids Res. **40**: W452-W457, 2012.

Tajima, F.: Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. - Genetics **123**: 585-595, 1989.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. *MEGA5*: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. - Mol. Biol. Evol. **28**: 2731-2739, 2011.

Urbano, M., Margiotta, B., Colaprico, G., Lafiandra, D.: Waxy proteins in diploid, tetraploid and hexaploid wheats. - Plant. Breed. **121**: 465-469, 2002.

Wang, S., Li, X., Wang, K., Wang, X., Li, S., Zhang, Y., Guo, G., Zeller, F.J., Hsam, S.L.K., Yan, Y., Gustafson, P.: Phylogenetic analysis of C, M, N, and U genomes and their relationships with *Triticum* and other related genomes as revealed by *LMW-GS* genes at *Glu-3* loci. - Genome **54**: 273-284, 2011.

Yamamori, M.: Amylose content and starch properties generated by five variant Wx alleles for granule-bound starch synthase in common wheat (*Triticum aestivum* L.). - Euphytica **165**: 607-614, 2009.

Yamamori, M., Nakamura, T., Endo, T.R., Nagamine, T.: Waxy protein deficiency and chromosomal location of coding genes in common wheat. - Theor. appl. Genet. **89**: 179-184, 1994.

Yamamori, M., Nakamura, T., Nagamine, T.: Ploymorphism of two waxy proteins in the emmer group of tetraploid wheat, *Triticum dococcoides*, *T. dicoccum* and *T. durum*. - Plant. Breed. **114**: 215-218, 1995.

Yamamori, M., Yamamoto, K.: Effects of two novel Wx-A1 alleles of common wheat (*Triticum aestivum* L.) on amylose and starch properties. - J. Cereal Sci. **54**: 229-235, 2011.

Yan, L., Bhave, M., Fairclough, R., Konik, C., Rahman, S., Appels, R.: The genes encoding granule-bound starch synthases at the waxy loci of the A, B, and D progenitors of common wheat. - Genome **43**: 264-272, 2000.

Yan, L., Bhave, M.: Characterization of waxy proteins and waxy genes of *Triticum timopheevii* and *T. zhukovskyi* and implications for evolution of wheat. - Genome **44**: 582-588, 2001.

Zeng, M., Morris, C.F., Batey, I.L., Wrigley, C.W.: Sources of variation for starch gelatinization, pasting, and gelation properties in wheat. - Cereal Chem. **74**: 63-71, 1997.