Molecular characterization of a novel *waxy* allele (*Wx-A^{u}1a*) from *Triticum urartu*

Carlos Guzmán*, Juan B. Alvarez

Departamento de Genética, Escuela Técnica Superior de Ingeniería Agronómica y de Montes, Edificio Gregor Mendel, Campus de Rabanales, Universidad de Córdoba, ES-14071 Córdoba, Spain.

* Corresponding author: Dr. C. Guzmán Email: ge2gugac@uco.es

Short tittle: *Waxy* gene of *T. urartu*

Keywords: waxy proteins, starch, *T. urartu*, genetic resources.

For GRACE

Abstract

Granule Bound Starch Synthase I, or waxy protein, is the sole enzyme responsible for the accumulation of amylose during the development of starch granules in wheat. The full coding region of the *Wx* gene was sequenced in *T. urartu* ¸ (a wild diploid species) and is related to the A genome of polyploid wheats. The *Wx* gene of *T. urartu (Wx-A^{u}1)* showed a homology of ~ 88.0% with *Wx-A1* from polyploid wheats. A greater homology was found with *Wx-A^{m}1* from the diploid cultivated wheat einkorn. Most of the differences were found in introns although several changes were also detected in exons that led to amino acid changes in the transit peptide and mature protein. These results show the potential of *T. urartu* as a source of new alleles that could be used in the breeding of durum and common wheat in order to synthesize novel starches.

1. Introduction

Granule Bound Starch Synthase I, or waxy protein, is one of the most important determinants of starch synthesis in plants (Baldwin 2001). In common wheat (*Triticum aestivum* L. ssp. *aestivum*; 2n = 6× = 42, AABBDD), three waxy proteins, one for each genome, have been identified and they are controlled by the same gene number located on the chromosomes 7AS, 4AL (translocated from original 7BS) and 7DS, respectively (Yamamori et al. 1994). These proteins have been shown to be the sole enzymes responsible for the accumulation of amylose during the development of starch granules in wheat (Nakamura et al. 1995), and thus have a crucial role in the quality and end use of the product as the amylose status of cereal starches affects important functional properties such as gelatinization, pasting and gelation (Zeng et al. 1997).

To date, different variants of each waxy protein have been reported and some of them have been characterized at the molecular level. Among the alleles described are the *null* alleles that lead to the absence of the protein (Vrinten at al. 1999; Saito et al. 2004; Saito and Nakamura 2005) and others that cause a variation in the waxy protein status (Yamamori 2009; Yamamori and Yamamoto 2011). All of them have been shown to be very important in regulating the amylose content of all wheat cultivars. However the number of alleles that have been detected are small when compared to other endosperm storage proteins involved in wheat quality. This has led to the search for new alleles of these genes in old and ancient wheat varieties (Rodríguez-Quijano et al. 1998; Urbano et al. 2002; Yamamori et al. 1995) together with their wild relatives (Caballero et al. 2008; Guzman et al. 2009, 2011). Among the latter, the diploid wheats could be good candidates

for new Wx-A1 alleles that could be used to enlarge the gene pool of cultivated wheats. One of them is *T. urartu* Thum. *ex* Gandil (2n = 2× = 14, A^uA^u), a wild diploid wheat mainly distributed in the Fertile Crescent region (Johnson 1975; Miller 1987). Numerous studies have suggested that this species generated wild emmer wheat (*T. turgidum* ssp. *dicoccoides* Körn. *ex* Asch. & Graebner em. Thell.; 2n =4× = 28, A^uA^uBB) by a spontaneous cross with an *Aegilops* species (probably *Aegilops speltoides* Tausch.; 2n =2× = 14, SS) with later chromosomal duplication. The domestication of wild emmer wheat produced cultivated emmer wheat (*T. turgidum* ssp. *dicoccum* Schrank ex Schübler em. Thell.; 2n = 4× = 28, A^uA^uBB), the predecessor to durum wheat (*T. turgidum* ssp. *durum* Desf. em. Husnot; 2n = 4× = 28, A^uA^uBB) and of common wheat by one additional cross and chromosomal duplication with *Ae. tauschii* Coss. (2n = 2× = 148, DD). Consequently the A genome of the all polyploid wheats has its origin in this wild diploid wheat (Dvorak et al. 1988).

While *waxy* genes from other diploid wheats, such as *T. monococcum* L. ssp. *monococcum* (2n = 2× = 14, A^mA^m) have been well characterized at the molecular level (Murai et al. 1999; Yan et al. 2000), only two partial sequences of *T. urartu Wx* gene have been reported (Mason-Gamer et al. 1998; Yan and Bhave 2000). Thus, the complete *Wx* coding region and its deduced amino acid sequence are still unknown in this species.

The aims of the current study were to sequence the *Wx* gene in *T. urartu*, which potentially could be an important source for improving starch quality and also provide valuable information on the phylogeny of wild and cultivated wheats.

2. Material and Methods

2.1 Plant material

One Iraq accession (MG 26992) of *T. urartu* obtained from Instituto del Germoplasma (Bari, Italy) was used. This accession was previously characterized for morphological traits by Castagna et al. (1997) confirming its identity; and was used to the development of introgression lines between *T. urartu* and durum wheat (Alvarez et al. 2009).

2.2 DNA extraction and PCR amplification

For DNA extraction, approximately 100 mg of young leaf tissue was excised, immediately frozen in liquid nitrogen and stored at -80°C. DNA was isolated using the CTAB method (Stacey and Isaac, 1994).

Three pairs of primers were designed to amplify the waxy gene in three fragments: Waxy1 Fw: 5′-TTGCTGCAGGTAGCCACACC-3′, and Waxy1 Rv: 5′-CCGCGCTTGTAGCAGTGGAA-3′; Waxy2 Fw: 5′-ATGGTCATCTCCCCGCGCTA-3′, and Waxy2 Rv: 5′-GTTGACGGCGAGGAACTTGT-3′; and Waxy3 Fw: 5′-GGCATCGTCAACGGCATGGA-3′, and Waxy3 Rv: 5′-ATGGACGTCAGCGAGTGGGA-3′. Each 15-μl reaction included 50 ng DNA, 1.5 mM $MgCl_2$, 0.2 μM of each primer, 0.2 mM dNTPs, 1.5 μl 10x PCR Buffer and 0.75U DNA polymerase (Promega, Madison, WI, USA). The PCR conditions included an initial denaturation step of 3 min at 94°C followed by 35 cycles as follows: for Waxy1 Fw/Waxy1 Rv, 40 s at 94°C, 45s at 62°C then 1 min at 72°C; for Waxy2 Fw/Waxy 2 Rv, 40 s at 94°C, 45 s at 62°C then 1 min 45 s at 72°C; and for Waxy3 Fw/Waxy 3 Rv, 40 s at 94°C, 45 s at 62°C then 1 min 30s at 72°C. After the 35 cycles all reactions included a final extension of 5 min at 72°C.

2.3 Sequencing analysis of PCR products

Amplification products were fractionated in vertical PAGE gels at 8% (w/v, C: 1.28%) and the bands were visualized by ethidium bromide staining. After, PCR products were excised from polyacrylamide gel and cloned into pGEM T-easy vector (Promega) for sequencing. Inserts were sequenced from at least three different clones using an ABI Prism 310 Genetic Analyzer (Applied Biosystems, Carlsban, CA, USA).

The sequence reported in the current study was compared to the sequences of the Wx-A1 gene available in the GenBank (NCBI) of common wheat cv. Chinese Spring (AB019622), durum wheat cv. Langdon (AB029063), wild emmer (AB029061), emmer (HM751941) and einkorn (AF110373), using Geneious Pro ver. 5.0.3 (Biomatters Ltd.) software. A neighbour-joining cluster with all sequences analysed was generated using the Maximun Composite likelihood method for the DNA sequences (Tamura et al. 2004) and the Poisson correction method for aminoacid sequences (Zuckerkandl and Pauling 1965). In both cases, one bootstrap consensus from 1000 replicates was used (Felsenstein 1985). The analyses were conducted in MEGA5 software (Tamura et al. 2011).

3. Results and Discussion

The entire coding sequence for the *Wx* gene of *T. urartu* (*Wx-A^{u}1*) was 2783 bp long, divided into eleven exons and ten introns. This structure and size is similar to the *Wx* genes present in other wheats (2781 bp in common wheat). The complete sequence is available in GenBank (JN857937).

Due to the A^u genome potentially being an ancestor of the A genome found in polyploid wheats (Dvorak et al. 1988), the *Wx-A^{u}1* sequence was compared with different

Wx-A1 alleles detected in several species of polyploid wheat, together with the ones present in cultivated diploid wheat (einkorn). This alignment and comparison is shown in Figure 1. The initiation codon, ATG and the termination codon TGA for translation, as well as the splice junctions of each intron of *Wx-A^{u}1*, were in homologous positions to those in these other *Wx-A1* genes. In all cases the exons were the same size (Table 1), with the exception of exon 1 that presented one additional codon in both diploid wheats (einkorn and *T. urartu*). In contrast, nine out of ten introns of the *Wx-A^{u}1* gene were different to those of the *Wx-A1* gene of polyploid wheat, where only intron 4 was the same size. These differences were smaller between the diploid wheats, which only showed different sizes for introns 1, 4 and 8.

The comparison between the six nucleotide sequences showed that the homology among the *Wx-A1* genes from polyploidy wheats and *Wx-A^{u}1* was around 88.0%, whereas this homology was clearly higher between *Wx-A^{m}1* and *Wx-A^{u}1* (97.4%). Most of the differences were found in introns, although several SNPs were also detected in exons. These changes led to marked differences in the deduced sequences of the respective proteins (Figure 2). In total, 27 amino acids changed in *Wx-A^{u}1* compared to *Wx-A1a*. The comparison showed little conservation where 12 amino acids were different within the first 70 amino acid sequence that forms the transit peptide. Eleven were common to both diploid wheats: Thr 14 → Ala, Ser 17 → Gly, Val 18 → Ile, Pro 25 → Ala, Leu 30 → Val, Asn 34 → Val, - 54 → Gln, Pro 58 → Ala, Phe 61 → Gly, Asp 62 → Thr, Met 68 → Val; whereas only one appeared in *T. urartu* (Val 5 → Ala). The amino acid change at position 61 was also common to durum wheat and emmer, but in the latter species the Asp62 was replaced by Asn.

The remaining differences were randomly placed in the mature protein, although the

central region of the sequence (220-350) was very conservative and no modification was found. The amino acid changes found between *Wx-A^{u}1* and *Wx-A1a* were: Ser 75 → Gly, Ala 105 → Pro, Ile 133 → Val, Val 138 → Ile, Val 139 → Ala, Arg 141 → Glu, Tyr 151 → Phe, Gln 191 → Leu, His 214 → Tyr, Ile 358 → Thr, Thr 364 → Ala, Thr 451 → Ser, Trp 454 → Arg, Asp 575 → His and Leu 599 → Met. Three of these changes were exclusive to the Wx-A^{u}1 sequence and were not detected in the Wx-A^{m}1 sequence (Ser 75 → Gly, Val 138 → Ile and Thr 364 → Ala). This latter sequence also showed three exclusive amino acid changes (Lys 361 → Asn, Asp 368 → Asn and Ala 377 → Val). Although further studies need to be undertaken, these changes in the sequence of the waxy proteins could cause activity differences and thus changes in starch composition. Consequently the *Wx-A^{u}1* allele was considered novel and tentatively named *Wx-A^{u}1a*.

The phylogenetic relationships between the *Wx-A^{u}1* gene and the rest of the *Wx-A1* genes analysed (*Wx-A^{m}1* from einkorn and *Wx-A1* from polyploid wheats) were evaluated using both nucleotide and amino acid sequences (Figure 3). In both cases (gene and protein), the novel allele showed a higher homology with the *Wx-A^{m}1* than to those present in polyploid wheats. This is contrary to the theory that suggests that the A genome of the polyploid wheats is derived from *T. urartu* and not from einkorn (see Salamini et al. 2002 for revision). However, these data should be used cautiously because, due to the time that has passed between the generation of the wild emmer wheat and present day, the *Wx* gene of *T. urartu* could have evolved in a different way. This suggests that the variation in this wild species could be very different to that present in modern wheat and thus could be a good candidate as a source of this or other genes that might be used in the breeding of durum or common wheat.

In conclusion, several sequence variations were found in the *T. urartu* sequence,

which translate to amino acid changes and that may have led to the synthesis of novel starches in wheat, thus creating the potential for wheat breeders to manipulate starch synthesis through conventional breeding or transgenic modification. The production of wheats containing novel starches will reduce or eliminate the need for costly post-harvest modifications.

Acknowledgements

This research was supported by grant AGL2010-19643-C02-01 from the Spanish Ministry of Science and Innovation, co-financed with European Regional Development Fund (FEDER) from the European Union.

References

Alvarez JB, L Caballero, S Nadal, MC Ramírez, A Martín 2009 Development and gluten strength evaluation of introgression lines of *Triticum urartu in* durum wheat. Cereal Res Commun 37: 243-248.

Baldwin PM 2001 Starch granule-associated protein and polypeptides: a review. Starch/Stärcke 53: 475-503.

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

Castagna R, S Gnocchi, M Perenzin, M Heun 1997 Genetic variability of the wild diploid wheat *Triticum urartu* revealed by RFLP and RAPD markers. Theor Appl Genet 94: 424-430.

Dvorak J, PE McGuire, B Cassidy 1988 Apparent source of the A genomes of wheats inferred from polymorphism in abundance and restriction fragment length of repeated nucleotide sequences. Genome 30: 680-689.

Felsenstein J 1985 Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39: 783-791.

Guzman C, L Caballero, JB Alvarez 2009 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.

Guzman C, L Caballero, JB Alvarez 2011 Molecular characterization of the Wx-B1 allelic variants identified in cultivated emmer wheat and comparison with those of durum wheat. Mol Breeding 28: 403-411.

Johnson BL 1975 Identification of the apparent B-genome donor of wheat. Can J Genet Cytol 17: 21-39.

Mason-Gamer RJ, FW Clifford, EA Kellogg 1998 Granule-Bound Starch Synthase: structure, function and phylogenetic utility. Mol Biol Evol 15: 1658-1673.

Miller TE 1987 Systematic and evolution. In: Lupton, F.G.H. (ed). Wheat breeding: its scientific basis. Chapman & Hall, London, pp. 1-30.

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

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

Rodríguez-Quijano M, MT Nieto-Taladriz, JM Carrillo 1998 Polymorphism of waxy proteins in Iberian hexaploid wheats. Plant Breeding 117: 341-344.

Saito M, M Konda, P Vrinten, K Nakamura, T Nakamura 2004 Molecular comparison of waxy null alleles in common wheat and identification of a unique null allele. Theor Appl Genet 108: 1205-1211.

Saito M, T Nakamura 2005 Two point mutations identified in emmer wheat generate null Wx-A1 alleles. Theor Appl Genet 110: 276-282.

Salamini F, H Özkan, A Brandolini, R Schäfer-Pregl, W Martin 2002 Genetics and geography of wild cereal domestication in the Near East. Nat Genet Reviews 3: 429-440.

Stacey J, P Isaac 1994 Isolation of DNA from plants. In: Isaac PG (ed.) Methods in Molecular Biology: Protocols for Nucleic Acid Analysis by Non-Radiactive probes. Totawa: Humana Press, pp. 9-15.

Tamura K, M Nei, S Kumar 2004 Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Nat Ac Sci 101: 11030-11035.

Tamura K, D Peterson, N Peterson, G Stecher, M Nei, S Kumar 2011 MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biol Evol 28: 2731-2739.

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

Vrinten P, T Nakamura, M Yamamori 1999 Molecular characterization of waxy mutations in wheat. Mol Gen Genet 261: 463-471.

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

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

Yamamori M 2009 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.

Yamamori M, K Yamamoto 2011 Effects of two novel *Wx-A1* alleles of common wheat

(*Triticum aestivum* L.) on amylose and starch properties. J Cereal Sci 54: 229-235.

Yan L, M Bhave 2000 Sequences of the waxy loci of wheat: utility in analysis of waxy proteins and developing molecular markers. Biochem Genet 38: 391-411.

Yan L, M Bhave, R Fairclough, C Konik, S Rahman, R Appels 2000 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.

Zeng M, CF Morris, II Batey, CW Wrigley 1997 Sources of variation for starch gelatinization, pasting, and gelation properties in wheat. Cereal Chem 74: 63-71.

Zuckerkandl E, L Pauling 1965 Evolutionary divergence and convergence in proteins. In: Bryson, V., Vogel, H.J. (Eds.), Evolving Genes and Proteins. Academic Press, New York, pp. 97-166.

Table 1. Size of the different exons and introns in the sequence evaluated.

	Species					
	Common wheat	Durum wheat	Wild emmer wheat	Emmer wheat	Einkorn wheat	*T. urartu*
Exon 1	321	321	321	313*	324	324
Exon 2	81	81	81	81	81	81
Exon 3	99	99	99	99	99	99
Exon 4	154	154	154	154	154	154
Exon 5	101	101	101	101	101	101
Exon 6	354	354	354	354	354	354
Exon 7	180	180	180	180	180	180
Exon 8	192	192	192	192	192	192
Exon 9	87	87	87	87	87	87
Exon 10	129	129	129	129	129	129
Exon 11	117	117	117	44*	117	117
Intron 1	82	82	82	82	88	89
Intron 2	84	84	84	84	88	88
Intron 3	109	109	109	109	118	118
Intron 4	125	125	125	125	143	125
Intron 5	99	99	99	99	81	81
Intron 6	91	91	91	91	89	89
Intron 7	95	95	95	95	81	81
Intron 8	90	90	90	90	83	80
Intron 9	98	98	98	98	99	99
Intron 10	93	93	93	93	115	115

* Sequence incomplete.

Caption figures

Fig. 1.- Alignment of genomic DNA sequences of the *Wx-A1* (common, durum, wild emmer and emmer), *Wx-A^{m}1* (einkorn) and *Wx-A^{u}1* (*T. urartu*) *gene.*

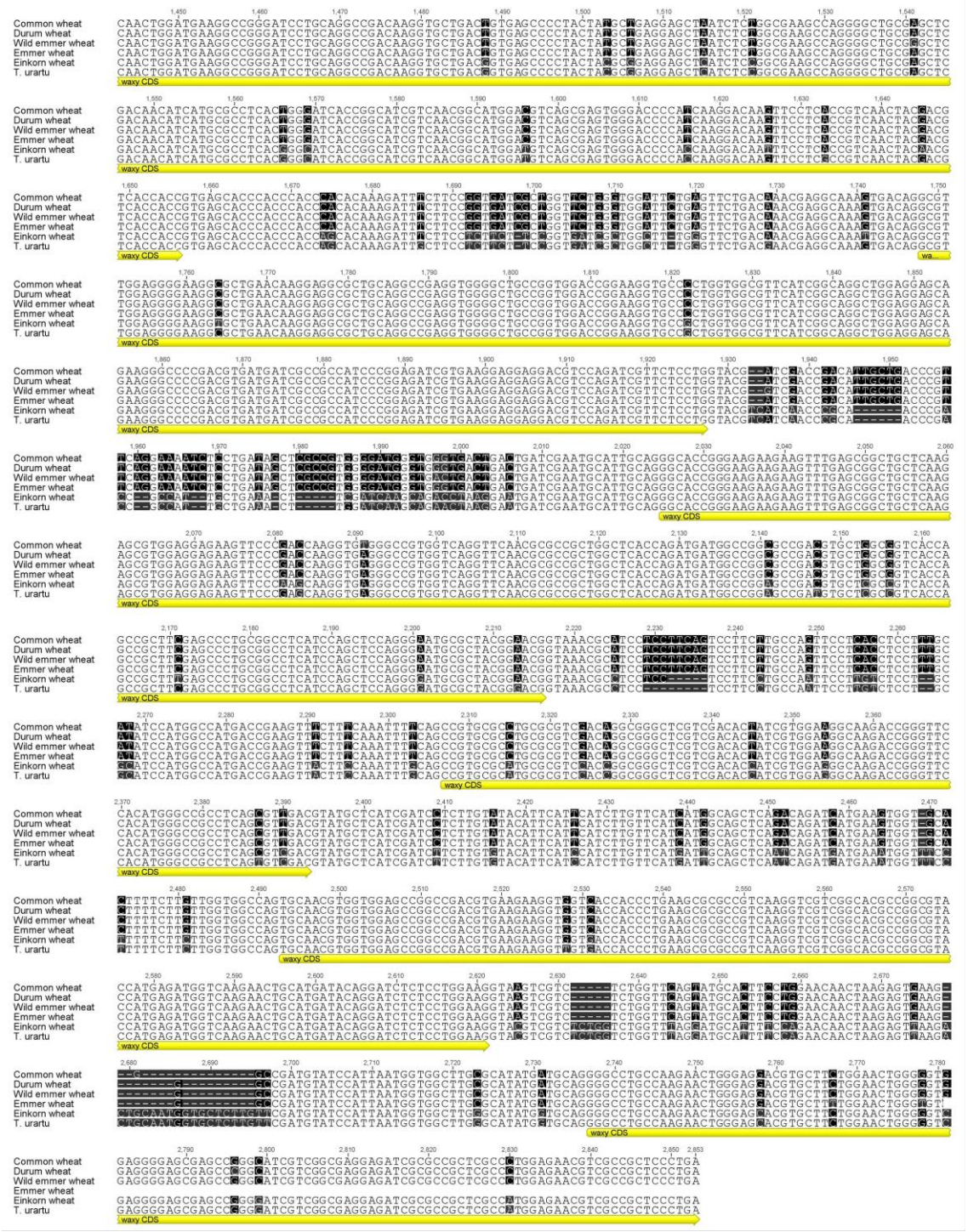

Fig. 2.- Alignment of deduced protein sequences of the *Wx-A1* gene from polyploid

wheats, *Wx-A^{m}1* gene from einkorn and *Wx-A^{u}1* from *T. urartu.*

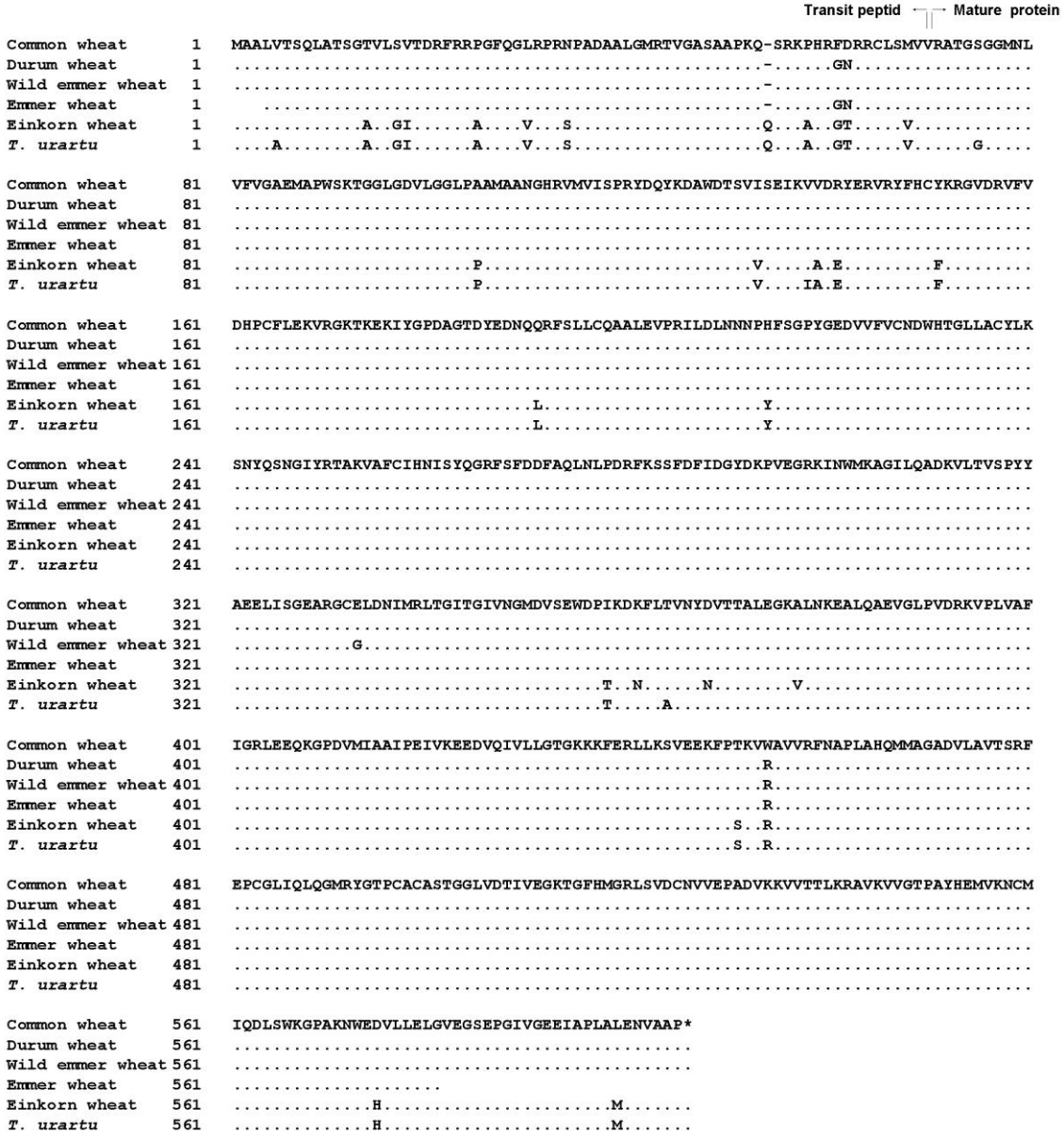

Fig. 3.- Neighbour-joining tree based in the Maximun Composite likelihood method (A) and the Poisson correction method (B) for nucleotide and aminoacid sequences analysed, respectively. Number above node indicated bootstrap estimates from 1,000 replications.