Theor Appl Genet (2013) 126:1703–1711
DOI 10.1007/s00122-013-2085-x

ORIGINAL PAPER

Molecular characterisation of the amino- and carboxyl-domains in different *Glu-A1x* alleles of *Triticum urartu* Thum. ex Gandil.

Juan B. Alvarez · M. Victoria Gutiérrez · Carlos Guzmán · Luis M. Martín

Received: 9 October 2012 / Accepted: 9 March 2013 / Published online: 24 March 2013

Abstract The wild diploid wheat (*Triticum urartu* Thum. ex Gandil.) is a potential gene source for wheat breeding, as this species has been identified as the A-genome donor in polyploid wheats. One important wheat breeding trait is bread-making quality, which is associated in bread wheat (*T. aestivum* ssp. *aestivum* L. em. Thell.) with the high-molecular-weight glutenin subunits. In *T. urartu*, these proteins are encoded by the *Glu-A1x* and *Glu-A1Ay* genes at the *Glu-A^{u}1* locus. The *Glu-A1x* genes of 12 *Glu-A^{u}1* allelic variants previously detected in this species were analysed using PCR amplification and sequencing. Data showed wide diversity for the *Glu-A1x* alleles in *T. urartu*, which also showed clear differences to the bread wheat alleles. This variation could enlarge the high-quality genetic pool of modern wheat and be used to diversify the bread-making quality in durum (*T. turgidum* ssp. *durum* Desf. em. Husn.) and common wheat.

Introduction

Wheat quality can be defined as the ability of a variety to produce flour suitable for a specific product. Consequently, this parameter is a variable that will depend on consumer preferences, the kind of product developed and the process used in the manufacturing of the product. This trait is strongly associated with grain components and their physico-chemical properties. Among these components, the endosperm storage proteins have been repeatedly associated with gluten strength (see Wrigley et al. 2006 for reviews). These proteins, also named prolamins, make up a complex mixture that can range between 50 components for hexaploid wheat and about 20 for diploid species (Shewry et al. 1988) and are mainly constituted by two fractions: gliadins, divided in α-, β-, γ- and ω-gliadins; and glutenins, classified in high- and low-molecular-weight subunits (HMWGs and LMWGs, respectively). In bread wheat (*Triticum aestivum* ssp. *aestivum* L. em. Thell.), multiple analyses have revealed the main role of HMWGs in flour quality (see Wrigley et al. 2006 for reviews).

Genetic studies have shown that the HMWGs are encoded by the complex loci *Glu-1* located on the long arm of each chromosome of the homoeologous group-1, called *Glu-A1*, *Glu-B1* and *Glu-D1*, respectively (Payne 1987). Each locus consists of two closely linked genes that encode for two types of subunits of different molecular weight and hence different mobility in SDS-PAGE: the *x*-type subunit, with higher molecular weight and lower mobility in SDS-PAGE; and the *y*-type subunit, smaller and with greater electrophoretic mobility (Harberd et al. 1986). The typical structure of the HMWGs comprises three distinct domains: two non-repetitive domains, amino-terminal (N-terminal) and carboxyl-terminal (C-terminal) where most cysteine residues involved in the disulphide bonds are found, together with one central repetitive domain that comprises hexa-, nona- and tri-peptide motifs, with the last only present in *x*-type subunits (Shewry et al. 1992). Both non-repetitive domains present a structure based on α-helices.

Communicated by A. Graner.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-013-2085-x) contains supplementary material, which is available to authorized users.

J. B. Alvarez (✉) · M. V. Gutiérrez · C. Guzmán · L. M. Martín
Departamento de Genética, Escuela Técnica Superior de Ingeniería Agronómica y de Montes, Universidad de Córdoba, Edificio Gregor Mendel, Campus de Rabanales, CeiA3, 14071 Córdoba, Spain
e-mail: jb.alvarez@uco.es

The N-terminal domains of the *x*-type subunits have 81–89 residues, whereas the C-terminal domains include 42 residues. In contrast, the repetitive domain shows one variable length with β-turn structure, responsible for the main differences in size of these genes (Shewry et al. 2002). This structure has been associated with the properties of gluten, mainly due to the number and distribution of disulphide bonds present in the N- and C-terminal domains, together with the properties and interactions of the repetitive domain.

In recent times, due to the high and stable grain yields for cultivated wheats, quality improvement has become the main role in most wheat breeding programmes. Therefore, the search for species that could be useful in contributing genes for wheat improvement is of great importance (Jauhar 1993). Of these species, the putative diploid ancestors of the genomes of wheat have proved to be good sources of useful genes (Srivastava and Damania 1989). One of them is *T. urartu* Thum. ex Gandil. ($2n = 2\times = 14$; A^uA^u), a wild wheat restricted to the Fertile Crescent region, in Armenia, Azerbaijan, Iran, Iraq, Lebanon, Syria and Turkey (Johnson 1975; Miller 1987). *Triticum urartu* has been proposed as a donor of the A genome in polyploid wheat species (Dvorak et al. 1993).

The variability of the endosperm storage proteins in one wide collection (169 accessions) of *T. urartu* was evaluated by SDS-PAGE analysis (Caballero et al. 2008, 2009; Martín et al. 2008). For the HMWGs, these studies permitted detection of up to 17 allelic variants, which were catalogued from *Glu-A^{u}1-af* to *Glu-A^{u}1-av* alleles (McIntosh et al. 2011). Of these 17 alleles, 12 were detected in 162 accessions obtained from the National Small Grain Collections (Aberdeen, Idaho, USA), which were originally collected in Lebanon (91 accessions), Turkey (66 accessions), Armenia (two accessions), Iraq and Iran (one accession each).

The main goal of the current study was the molecular characterisation of the *Glu-A1x* gene of 12 allelic variants present at the *Glu-A1* locus in *T. urartu*, together with their comparison with the main *Glu-A1x* alleles present in durum (*T. turgidum* ssp. *durum* Desf. em. Husn.) and bread wheat.

Materials and methods

Plant material

Twelve accessions of *T. urartu* representative of the HMWGs variation detected in 162 accessions by Caballero et al. (2008) were analysed. All these accessions were obtained from the National Small Grain collection (Aberdeen, Idaho, USA). The durum wheat cultivar *Alaga* (subunit Ax1) was used as standard for comparison.

DNA extraction and PCR amplification

DNA isolation was carried out from young leaf tissue using the CTAB method (Stacey and Isaac 1994).

The primers reported by D'Ovidio et al. (1995) were used to amplify the complete coding sequence of the *Glu-A1x* gene (AxF: 5′-AGATGACTAAGCGGTTGGTTC-3′; and AxR: 5′-AAGCGAGCTGCAGAGAGTTC-3′). PCR mixtures were carried out in 20 µl final volume using 1× Optibuffer (Bioline), 150 ng of template DNA, 0.6 µM of each primer, 2.5 mM $MgCl_2$, 0.2 *m*M of each deoxyribonucleotide, 1× Hi-Spec Additive and 2 U of BIO-X-ACT Short DNA Polymerase (Bioline). DNA was subjected to an initial denaturation step at 94 °C for 5 min, followed of one step at 62 °C for 2 min. The amplification conditions were for 30 cycles at 94 °C for 30 s, 62 °C for 1 min and 72 °C for 2 min and 30 s followed by a final incubation step at 72 °C for 10 min.

The amplicons were separated on 1 % agarose gels in TBE buffer. These amplicons were digested with *DdeI* endonuclease and separated in PAGE gel with discontinuous Tris–HCl buffer system (pH: 6.8/8.8) at a polyacrylamide concentration of 10 % (w/v, C: 1.68 %).

DNA sequencing analysis

A set of specific primers was designed on the *Glu-A1-1b* sequence (NCBI ID: EF055262) to amplify the two extremes of the *Glu-A1x* gene. The Fragment 1, with size of 495 bp, includes the N-terminal domain together with the beginning of the repetitive domain (AxtN1F: 5′-CTGCTCCACGCTAACATGGTATG-3′ and AxtN1R: 5′-TGACTAAGCGGTTGGTTCTTTT-3′); whereas the Fragment 2 (336 bp) includes the bottom of repetitive domain and C-terminal domain (AxtCF: 5′-CCTGATTGTTGC TTTTGTCC-3′ and AxtCR: 5′-AATGGCTGCAATCAGG GTACTA-3′).

PCR mixtures for the specific primers were carried out in 20 µl final volume using 1× *Taq* PCR buffer (Promega), 75 ng of template DNA, 0.4 µM of each forward and reverse primer, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP and 1 U of *Taq* DNA polymerase (Promega). For the amplification, genomic DNA was subjected to an initial denaturation step at 95 °C for 5 min, following to 35 cycles at 95 °C for 1 min, 60 °C for 1 min and 72 °C for 1 min followed by a final incubation step at 72 °C for 8 min. The amplicons were separated and excised from a 1 % agarose gel, purified using the Favorgen Gel/PCR Purification Kit (Favorgen Biotech Corp.) and cloned into pGEM T-easy vector (Promega, Madison, WI, USA) for sequencing. Inserts were sequenced from at least three different clones using an ABI Prism 310 Genetic Analyzer (Applied Biosystems, Carlsban, CA, USA).

Data analysis

DNA analyses were conducted by DNAsp ver. 5.0 (Librado and Rozas 2009). Nucleotide diversity was estimated as theta (θ), the number of segregating (polymorphic) sites (Watterson 1975), and pi (π), the average number of nucleotide differences per site between two sequences (Nei 1987). Tests of neutrality were performed using Tajima's *D* statistic (1989). The synonymous (K_s) and non-synonymous (K_a) substitution rates and the relation K_a/K_s were computed using DNAsp ver. 5.0 (Librado and Rozas 2009). Divergence times were calculated by the mean divergence times (2.7 million years ago, MYA) between the A and D genomes estimated by Dvorak and Akhunov (2005).

Phylogenetic tree was constructed with MEGA5 software (Tamura et al. 2011) using the concatenate sequence obtained from the merge of the Fragments 1 and 2 of each line. Neighbour-joining cluster with all sequences analysed was generated using the maximum composite likelihood method (Tamura et al. 2004) and one bootstrap consensus from 1,000 replicates was used (Felsenstein 1985).

Results

PCR analysis and digestion patterns of *x*-type genes present at the *Glu-A1* locus

In a previous work (Caballero et al. 2008), 162 accessions of *T. urartu* were analysed by SDS-PAGE means. In the current study, the *Glu-A1x* gene from 12 of those *T. urartu* accessions carrying different alleles was amplified. All lines analysed presented amplification products with size around 2.3 kbp. Small differences were detected among them as was shown in the gel (Fig. 1). In order to evaluate in detail, these differences in the PCR products were digested with specific endonucleases, specifically with *Dde*I restriction enzyme (Fig. 2). This endonuclease permits differentiation of the N- and C-terminal domains from the repetitive domain, and variation between the alleles evaluated could be observed in both regions. Differences were also detected in the repetitive domain, which were related with sequence size. As these genes are intronless, the gene size is equivalent to polypeptide size, which was the basis for the previous classification in SDS-PAGE (Caballero et al. 2008).

Two new pairs of primers were designed to amplify both extremes of the gene. The first pair produced one amplicon (Fragment 1) that included the signal peptide, the N-terminal domain and the first 174 nucleotides of the repetitive domain. The second pair produced one amplicon (Fragment 2) that included the last 210 nucleotides of the repetitive domain and the C-terminal domain (Fig. 2). Both

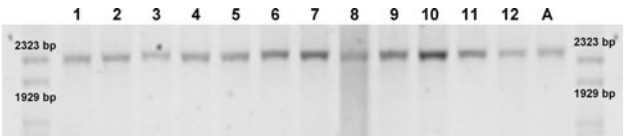

Fig. 1 PCR amplification of the complete coding region of *Glu-A1x* genes. *Lanes* are as follows: *1* PI 428319 (*Glu-A1-1af*), *2* PI 428242 (*Glu-A1-1ag*), *3* PI 428240 (*Glu-A1-1ah*), *4* PI 428335 (*Glu-A1-1ai*), *5* PI 538741 (*Glu-A1-1aj*), *6* PI 428230 (*Glu-A1-1ak*), *7* PI 428253 (*Glu-A1-1al*), *8* PI 427328 (*Glu-A1-1am*), *9* PI 428327 (*Glu-A1-1an*), *10* PI 428256 (*Glu-A1-1ao*), *11* PI 428224 (*Glu-A1-1ap*), *12* PI 428228 (*Glu-A1-1aq*), and *A* cv. Alaga (*Glu-A1-1a*)

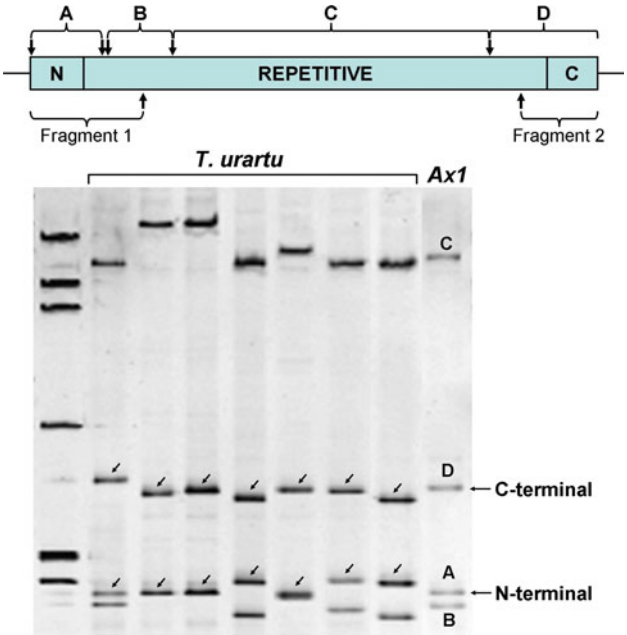

Fig. 2 PAGE separation of PCR products from the *Glu-A1x* genes digested with *Dde*I

fragments were cloned and sequenced. Fragment 1 was 495 bp in size with the exception of alleles *Glu-A1-1af*, *Glu-A1-1ai* and *Glu-A1-1aj* of 486 bp and the *Glu-A1-1am* allele of 477 bp (Table 1). In the first case this different size is a consequence of one deletion of nine nucleotides inside the N-terminal domain, whereas the second case (*Glu-A1-1am*) was due to a deletion of 18 nucleotides in the repetitive domain (Supplementary Material, FS1). One similar deletion was detected in the sequence of the *Glu-A1-1b* allele that coded the subunit Ax2* in wheat (Supplementary Material, FS1). However, the Fragment 2 size was equal in all sequences analysed (Supplementary Material, FS2). These sequences were concatenated and are available in the GenBank database.

The frequency and distribution of these alleles was very variable. *Glu-A1-1af* was the most frequent (43.3 %) and was found in 71 accessions (64 from Lebanon and seven from Turkey). The other two alleles with one deletion in the N-terminal domain (*Glu-A1-1ai* and *Glu-A1-1aj*) were

Table 1 Origin and sequence characteristics of the *Glu-A1x* genes evaluated

Allele[1]	Ay subunit[2]	Accession	Size (bp)		NCBI code
			Fragment 1	Fragment 2	
Glu-A1-1af	Active	PI 428319	486	336	JX102646
Glu-A1-1ag	Non-active	PI 428242	495	336	JX102638
Glu-A1-1ah	Active	PI 428240	495	336	JX102644
Glu-A1-1ai	Active	PI 428335	486	336	JX102636
Glu-A1-1aj	Active	PI 538741	486	336	JX102643
Glu-A1-1ak	Non-active	PI 428230	495	336	JX102637
Glu-A1-1al	Active	PI 428253	495	336	JX102639
Glu-A1-1am	Active	PI 427328	477	336	JX102640
Glu-A1-1an	Active	PI 428327	495	336	JX102635
Glu-A1-1ao	Active	PI 428256	495	336	JX102645
Glu-A1-1ap	Non-active	PI 428224	495	336	JX102642
Glu-A1-1aq	Active	PI 428228	495	336	JX102641

[1] According to McIntosh et al. (2011)

[2] According to Caballero et al. (2008)

detected in 19 and five accessions, respectively; all were from Lebanon, with exception of one Turkish accession for *Glu-A1-1ai*. *Glu-A1-1am* was only found in one accession from Iraq (Table 2).

Sequence comparative analysis

A summary of DNA polymorphism found in the *Glu-A1x* sequences evaluated is shown in Table 3. These data were evaluated using the concatenate sequence (Fragment 1 + Fragment 2), without the repetitive domain that was not sequenced. Up to 27 polymorphic sites were detected in the overall sequences, with 28 mutations detected (η); Fragment 1 was ≈threefold more variable than Fragment 2 (Table 3). Of these mutations, 11 were synonymous (silent mutations) while the rest were non-synonymous, which implied changes in the amino acid sequences. A large degree of variation was detected in the N-terminal domain, where 11 mutations were found (Table 3). However, the highest number of non-synonymous substitutions was in the repetitive domain (five each for Fragments 1 and 2). The use of both fragments permitted detection of up to nine different haplotypes or alleles among the 12 sequences analysed (Table 3), the other differences detected by SDS-PAGE (Caballero et al. 2008) were likely due to differences in the repetitive domain (gene size).

Two statistics, π (Nei 1987) and θ (Watterson 1975), were used to estimate nucleotide diversity (Table 3). Both values were similar in all cases, which is associated with a drift–mutation balance. The values for the C-terminal domain were lower than in the rest of the gene, which is in consonance with the low values of diversity observed in this domain compared to the other domains of the gene. Tajima's *D* values were not significant in any case, which was consistent with a neutral equilibrium (Table 3).

Table 2 Distribution of the alleles evaluated in the current study

Allele	Total	Armenia	Iran	Iraq	Lebanon	Turkey
Glu-A1-1af	71	–	–	–	64	7
Glu-A1-1ag	30	2	–	–	3	25
Glu-A1-1ah	26	–	–	–	–	26
Glu-A1-1ai	19	–	–	–	18	1
Glu-A1-1aj	5	–	–	–	5	–
Glu-A1-1ak	5	–	1	–	–	4
Glu-A1-1al	1	–	–	1	–	–
Glu-A1-1am	1	–	–	1	–	–
Glu-A1-1an	1	–	–	–	1	–
Glu-A1-1ao	1	–	–	–	–	1
Glu-A1-1ap	1	–	–	–	–	1
Glu-A1-1aq	1	–	–	–	–	1
Overall	162	2	1	2	91	66

Amino acid sequence comparison

The alignment of the deduced proteins from the *T. urartu* sequences, and their comparison with the amino acid sequence of the Ax1 and Ax2* subunits detected in wheat are shown in Fig. 3. The main differences were the presence of two InDels: one in the amino-terminal domain located among the 83 and 85 residues (Phe-Tyr-Pro) and other in the repetitive domain among the 132–137 residues (Gly-Gln-Gly-Gln-Gln-Pro). The other changes were single nucleotide polymorphisms (SNPs).

Table 3 DNA polymorphism and test statistics for selection of the 12 sequences from *T. urartu* evaluated

	Fragment 1				Fragment 2			Total
	Signal peptide	N-terminal	Repetitive	Complete	Repetitive	C-terminal	Complete	
Region	1–63	64–321	322–495	1–495	496–705	706–831	496–831	1–831
Size	63	258	174	495	210	126	336	831
k	0.79	3.65	1.33	5.77	2.12	0.58	2.7	8.47
η	2	11	7	20	6	2	8	28
s	2	11	7	20	5	2	7	27
SS	1	7	2	10	1	0	1	11
NSS	1	4	5	10	5	2	7	17
h	3	7	6	8	6	3	7	9
$\theta \times 10^{-3}$	10.5	14.7	10.7	13.1	9.5	5.3	7.9	11.0
$\pi \times 10^{-3}$	12.5	14.6	14.9	14.2	10.1	4.6	8.0	11.1
D	0.554	0.010	−1.100	−0.343	0.258	−0.382	0.072	−0.227

k, average number of nucleotide differences; η, total number of mutations; s, number of polymorphic sites; SS, synonymous substitutions; NSS, non-synonymous substitutions; h, number of haplotypes; θ, Watterson's estimate; π, nucleotide diversity; D, Tajima's estimate D test

The non-synonymous substitutions observed in all *T. urartu* lines together with comparison with the Ax1 and Ax2* subunits used as standards are shown in Table 4. Three of these changes (two in Fragment 1, positions 62 and 63; and one in Fragment 2, position 62) were distinctive among wheat and *T. urartu* (Table 4). For Fragment 1, the residues 62 and 63 were equivalent in wheat (Gly) and each presented two alternatives in *T. urartu*. Eight alleles showed the change Gly62 → Ile while four had the change Gly62 → Val; for the residue 63, 11 alleles had the change Gly63 → Val and one had Gly63 → Ile (Table 3). The change in Fragment 2 was more variable and had three options in *T. urartu*: in seven alleles it was Ser62 → Pro, in four there was no change (Ser62 → Ser) and in one was Ser62 → Thr. The other changes appeared only in some alleles of *T. urartu* (Table 4), with most alleles similar to the wheat subunits used as standards.

Neighbour-joining tree of the *Ax* gene

A phenogram based on the maximum composite likelihood method was constructed using all the *Glu-A1x* sequences evaluated in this study, together with the main alleles (*Glu-A1-1a*, NCBI ID: X61009; *Glu-A1-1b*, NCBI ID: EF055262; and *Glu-A1-1c*, NCBI ID: AF145590) detected for the A genome in tetraploid and hexaploid wheats (Fig. 4). The concatenate sequence (Fragments 1 and 2) was used in all cases. The *T. urartu* sequences were grouped in three clusters. One cluster included the *Glu-A1-1ah*, *Glu-A1-1ak* and *Glu-A1-ao* alleles present in 32 Turkish accessions and in the Iranian accession. Another cluster showed only the *Glu-A1-am* allele detected exclusively in one Iraq accession, and the rest of the evaluated accessions (128) grouped in the third cluster. All of them showed clear separation of the *Glu-A1x* genes of wheat, which were grouped in one separate cluster.

The K_s and K_a substitution rates among *Glu-A1x* genes of *T. urartu* and wheat (subunits Ax1 and Ax2*) were calculated using the concatenate sequence (Fragments 1 and 2). The K_s value (0.0240) was ≈twofold of the K_a value (0.0104), with $K_a/K_s = 0.433$. The divergence rate ($r = 4.08 \times 10^{-8}$ synonymous substitutions/year) was calculated using $K_s = 0.1918$ obtained between the *Glu-A1x* sequences (subunits Ax1 and Ax2*) and *Glu-D1x* sequences (subunits Dx2 and Dx5) and the divergence time estimated by Dvorak and Akhunov (2005) for the separation between the A and D genomes (2.7 MYA). According to this estimation, the divergence time between *T. urartu* and wheat *Glu-A1x* genes was ≈0.6 MYA.

Discussion

The use of relatives and ancestors of wheat has been suggested as a promising strategy for wheat breeding (Srivastava and Damania 1989). Most of these species are wild, which implies an additional effort for breeders because undesirable characteristics can be incorporated during the crossing process together with useful traits. However, because the process of crossing and selection is slow, it is important to characterise the allelic variants of the genes of interest in these species before their incorporation into modern wheat for a successful breeding process.

One important aim in wheat breeding is improving bread-making quality, which is related to the seed storage proteins associated with gluten strength. Several studies

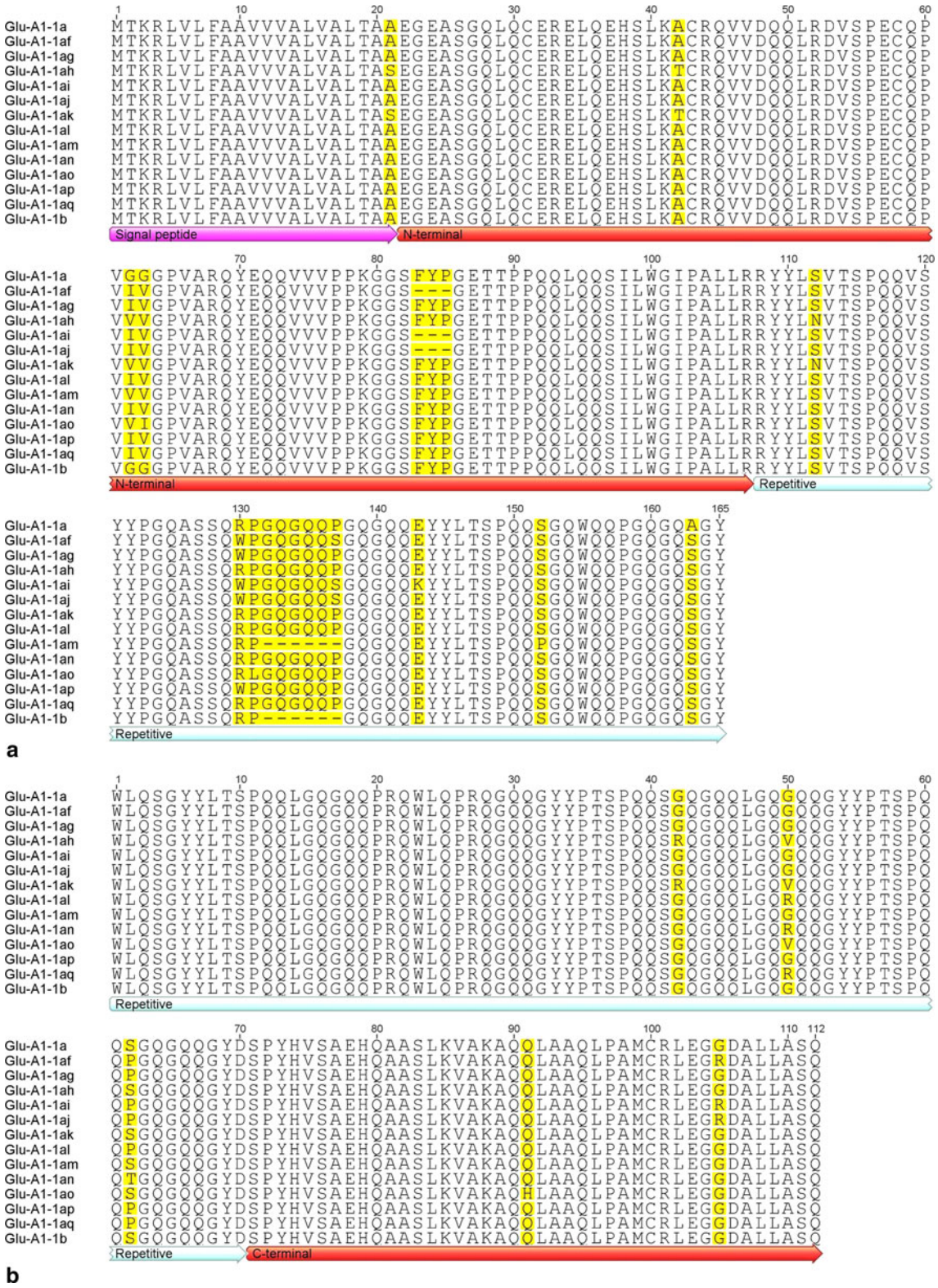

Fig. 3 Comparison of the aminoacid sequence of the Fragment 1 (**a**) and Fragment 2 (**b**) of the *Glu-A1x* alleles from *T. urartu* with the *Glu-A1-1a* allele (subunit Ax1) and the *Glu-A1-1b* allele (subunit Ax2*) of wheat

Table 4 Amino acid comparison among wheat (subunit Ax1 and Ax2*) and *T. urartu*, and changes inside of *T. urartu* lines

Position	Wheat	*T. urartu*		Sequence
	a.a.	a.a	change	
Fragment 1				
21	Ala	Ala	Ser	*Glu-A1-1ah*
42	Ala	Ala	Thr	*Glu-A1-1ah, Glu-A1-1ak*
62	Gly	Ile	Val	*Glu-A1-1ah, Glu-A1-1ak, Glu-A1-1am, Glu-A1-1ao*
63	Gly	Val	Ile	*Glu-A1-1ao*
112	Ser	Ser	Asn	*Glu-A1-1ah, Glu-A1-1ak*
130	Arg	Arg	Trp	*Glu-A1-1af, Glu-A1-1ag, Glu-A1-1ai, Glu-A1-1aj, Glu-A1-1ap*
131	Pro	Pro	Leu	*Glu-A1-1ao*
137	Pro/-[1]	Pro	Ser	*Glu-A1-1af, Glu-A1-1ai, Glu-A1-1aj*
143	Glu	Glu	Lys	*Glu-A1-1ai*
152	Ser	Ser	Pro	*Glu-A1-1am*
163	Ala/Ser[1]	Ser	–	–
Fragment 2				
42	Gly	Gly	Arg	*Glu-A1-1ah, Glu-A1-1ak*
50	Gly	Gly	Val	*Glu-A1-1ah, Glu-A1-1ak, Glu-A1-1ao*
			Arg	*Glu-A1-1al, Glu-A1-1an, Glu-A1-1aq*
62	Ser	Pro	Ser	*Glu-A1-1ah, Glu-A1-1ak, Glu-A1-1am, Glu-A1-1ao*
			Thr	*Glu-A1-1an*
91	Gln	Gln	His	*Glu-A1-1ao*
105	Gly	Gly	Arg	*Glu-A1-1af, Glu-A1-1ah, Glu-A1-1ai*

[1] Subunit Ax1 and Ax2*, respectively

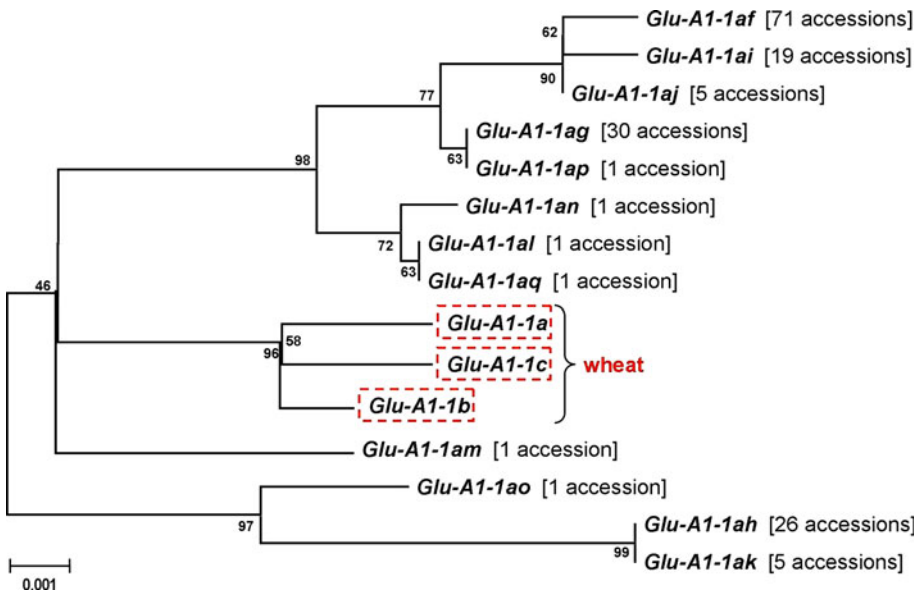

Fig. 4 Neighbour-joining tree based on maximum composite likelihood method of concatenate sequence (Fragment 1 + Fragment 2) of the *Glu-A1x* gene of *T. urartu* and wheat. *Numbers in nodes* indicated bootstrap estimates from 1,000 replications. The number of accessions carrying each allele is based in the previous study of Caballero et al. (2008) and is indicated between *brackets*

have suggested that the effects of these proteins are related to their number or polymorphism. According to the variation of the number of subunits, Ciaffi et al. (1995) showed that the presence of active Ay subunits could have a positive effect on bread-making quality in durum wheat. Rogers et al. (1997) reported positive effects by introducing a *Glu-A1* allele expressing both Ax and Ay subunits from *T. monococcum* ssp. *aegilopoides* Link em. Thell. (syn. *T. boeoticum* Boiss. ssp. *thaoudar*). These active subunits are always absent in bread wheat, while they are expressed in wild diploid (*T. urartu*) and tetraploid wheats (Waines and Payne 1987; Ciaffi et al. 1993).

Polymorphism effects within these proteins have been shown to be determinants in the variation of bread-making quality. In this respect, it is important to emphasise that variation for the *Glu-A1* locus is low and asymmetrically distributed. Most modern wheats present one of the three main alleles for this locus (*Glu-A1a*, *Glu-A1b* and *Glu-A1c*). The rest of the alleles detected could be catalogued as rare (≤5 %) or very rare (≤1 %) according to the classification of Marshall and Brown (1975). Consequently, the search of new *Glu-A1* alleles, together with their incorporation into the genetic pool of cultivated wheat, is very important for enlarging the gluten properties and diversifying the uses of this crop in the food industry.

One source of the possible new alleles is *T. urartu*, a wild diploid wheat determined to be the possible ancestor of the A genome in wheat (Dvorak et al. 1993). The variability of this species for HMWGs was evaluated in a wide collection obtained from two germplasm banks (National Small Grain Collections, Aberdeen, USA; and the Institute for Plant Genetics and Crop Plant Research, Gatersleben, Germany). This analysis led to detection of important variation that could be used to generate new materials (Caballero et al. 2008). In this respect, the positive effect on gluten strength of the introgression of *T. urartu* into durum wheat was evaluated by Alvarez et al. (2009), showing great potential of this species in quality breeding in wheat.

Although the additional presence of the Ay subunits is desirable, the role of the Ax subunits is not as important. However, as these genes are closely linked, both of them are incorporated together in any classic breeding process. The active and inactive *Glu-A1y* alleles found by Caballero et al. (2008) were analysed using a PCR–RFLP technique by Gutiérrez et al. (2011), which showed that their nucleotide sequences clearly differed from those of cultivated wheat. For this reason, in the current study, only the *Glu-A1x* gene of the 12 alleles previously identified by Caballero et al. (2008) were characterised using sequencing of their non-repetitive domains. In these domains all the cysteine residues are present, which are mainly responsible for the inter-bonds among different HMWGs—the degree of stable unions between these proteins determines gluten properties and thus bread-making quality (Shewry et al. 2002). Nine of these alleles sequenced showed active Ay subunits. Although the number of cysteine residues was similar to the Ax1 and Ax2* subunits (three in the N-domain and one in the C-domain), these alleles showed some variations in the amino- and carboxyl-sequences (three and two amino acid changes, respectively). However, some alleles showed differences in the repetitive domain that could influence the elastic properties of these repetitive domains. The main changes were Pro121 → Leu, Pro137 → Ser or Ser152 → Pro.

Analysis of the DNA polymorphism in the *T. urartu* sequences showed that the divergence among the A^u and A sequences evaluated in the current study could have begun approximately 600,000 years ago. This date has been established as the time of the generation of tetraploid wheat, which varies according to different authors between 0.36 MYA (Dvorak and Akhunov 2005) and 0.50 MYA (Huang et al. 2002). During this time the evolution of *T. urartu* could have generated new variation for the A genome, which could be used for breeding purposes because it is unlikely that this variation could be detected in cultivated wheat. However, the current variation present in cultivated wheat could be derived from one small part of the *T. urartu* germplasm, which is in consonance with the great homology described in *T. urartu* alleles with the most frequent subunits (Ax1 and Ax2*) in wheat.

In conclusion, although further studies should be carried out, particularly introgression to test the effects of these alleles, our data showed wide diversity for the *Glu-A1x* alleles in *T. urartu*, which also showed clear differences to wheat alleles. This opens the possibility of increasing the variation of the *Glu-A1* locus in wheat using *T. urartu* as a source and, consequently, to diversify the bread-making quality of wheat by variation in the number or internal characteristics of HMWG subunits.

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

References

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

Caballero L, Martín MA, Alvarez JB (2008) Allelic variation for the high-and-low-molecular-weight glutenin subunits in wild diploid wheat (*Triticum urartu*) and its comparison with durum wheats. Aust J Agric Res 59:906–910

Caballero L, Martín MA, Alvarez JB (2009) Genetic diversity for seed storage proteins in Lebanon and Turkey populations of wild diploid wheat (*Triticum urartu* Thum. *ex* Gandil.). Genet Resour Crop Evol 56:1117–1124

Ciaffi M, Lafiandra D, Porceddu E, Benedettelli S (1993) Storage-protein variation in wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*) form Jordan and Turkey. I. Electrophoretic characterization of genotypes. Theor Appl Genet 86:474–480

Ciaffi M, Lafiandra D, Turchetta T, Ravaglia S, Bariana H, Gupta RB, MacRitchie F (1995) Breadbaking potential of durum wheat lines expressing both *x*-and *y*-type subunits at the *Glu-A1* locus. Cereal Chem 72:465–469

D'Ovidio R, Masci S, Porceddu E (1995) Development of a set of oligonucleotide primers specific for genes at the *Glu-1* complex loci of wheat. Theor Appl Genet 91:189–194

Dvorak J, Akhunov ED (2005) 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

Dvorak J, Terlizzi P, Zhang HB, Resta P (1993) The evolution of polyploid wheats: identification of the A genome donor species. Genome 36:21–31

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Gutiérrez MV, Guzmán C, Martín LM, Alvarez JB (2011) Molecular characterization of the *Glu-Ay* gene from *Triticum urartu* for its potential use in quality wheat breeding. Plant Genet Resour 9:334–337

Harberd NP, Bartels D, Thompson RD (1986) DNA restriction fragment variation in the gene family encoding high-molecular-weight (HMW) glutenin subunits of wheat. Biochem Genet 24:579–596

Huang S, Sirikhachornkit A, Su X, Faris J, Gill B, Haselkorn R, Gornicki P (2002) Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the *Triticum/Aegilops* complex and the evolutionary history of polyploid wheat. Proc Nat Acad Sci USA 99:8133–8138

Jauhar PP (1993) Alien gene transfer and genetic enrichment of bread wheat. In: Damania AB (ed) Biodiversity and wheat improvement. ICARDA-A Wiley Sayce Publication, Aleppo, pp 103–119

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

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452

Marshall DR, Brown AHD (1975) Optimum sampling strategies in genetic conservation. In: Frankel OH, Hawkes JG (eds) Crop genetic resources for today and tomorrow. Cambridge University Press, Cambridge, pp 53–70

Martín MA, Martín LM, Alvarez JB (2008) Polymorphisms at the *Gli-A^{u}1* and *Gli-A^{u}2* loci in wild diploid wheat (*Triticum urartu*). Euphytica 163:303–307

McIntosh RA, Dubcovsky J, Rogers WJ, Morris C, Appels R, Xia XC (2011) Catalogue of gene symbols for wheat: 2011 Supplement (http://www.shigen.nig.ac.jp/wheat/komugi/genes/macgene/supplement2011.pdf)

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

Nei M (1987) Molecular evolutionary genetics. Columbia Univ Press, New York

Payne PI (1987) Genetics of wheat storage proteins and the effects of allelic variation on bread-making quality. Ann Rev Plant Physiol 38:141–153

Rogers WJ, Miller TE, Payne PI, Seekings JA, Sayers EJ, Holt LM, Law CN (1997) Introduction to bread wheat (*Triticum aestivum* L.) and assessment for bread-making quality of alleles from *T. boeoticum* Boiss. ssp. *thaoudar* at *Glu-A1* encoding two high-molecular-weight subunits of glutenin. Euphytica 63:19–29

Shewry PR, Parmar S, Field JM (1988) Two dimensional electrophoresis of cereal prolamins: applications to biochemical and genetic analyses. Electrophoresis 9:727–737

Shewry PR, Halford NG, Tatham AS (1992) High-molecular-weight subunits of wheat glutenin. J Cereal Sci 15:105–120

Shewry PR, Halford NG, Belton PS, Tatham AS (2002) The structure and properties of gluten: an elastic protein from wheat grain. Philos Trans R Soc Lond B 357:133–142

Srivastava JP, Damania AB (1989) Use of collections in cereal improvement in semi-arid areas. In: Brown AHD, Frankel OH, Marshall DR, Williams JT (eds) The use of plant genetic resources. Cambridge University Press, Cambridge, pp 88–104

Stacey J, Isaac P (1994) Isolation of DNA from plants. In: Isaac PG (ed) Methods in molecular biology: protocols for nucleic acid analysis by non-radioactive probes. Humana Press Inc., Totawa, pp 9–15

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585–595

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

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

Waines JG, Payne PI (1987) Electrophoretic analysis of the high-molecular-weight glutenin subunits of *Triticum monococcum*, *T. urartu*, and the A genome of bread wheat (*T. aestivum*). Theor Appl Genet 74:71–76

Watterson GA (1975) On the number of segregating sites in genetical models without recombination. Theor Pop Biol 7:256–276

Wrigley CW, Bekes F, Bushuk W (eds) (2006) Gliadin and glutenin: the unique balance of wheat quality. AACC International Press, St. Paul