Theor Appl Genet (2003) 107:74–83
DOI 10.1007/s00122-003-1234-z

A. De Bustos · N. Jouve

Characterisation and analysis of new HMW-glutenin alleles encoded by the Glu-R1 locus of Secale cereale

Received: 28 August 2002 / Accepted: 2 January 2003 / Published online: 7 March 2003

Abstract This work reports the molecular characterisation of new alleles of the previously reported *Glu-R1* locus. Wheat lines carrying the chromosome substitution 1R(1D), rye cultivars and related wild species were analysed. Five new x-type and four y-type *Glu-R1* glutenin subunits were isolated and characterised. The coding region of the sequences shows the typical structure of the HMW glutenin genes previously described in wheat, with the N and C-terminal domains flanking the central repetitive region. Tri-, hexa- and nona-peptides found in the central repetitive region of wheat glutenin genes were also present in the rye genes. Duplications and deletions of these motifs are responsible for allelic variation at the *Glu-R1* locus. Orthologous genes (from different genomes) were more closely related than paralogous genes (x- and y-type), supporting the hypothesis of gene duplication before *Triticeae* speciation. Differences in the number and position of cysteine residues identified alleles which in wheat are associated with good dough quality. SDS proteins encoded by some characterised alleles were presumptively identified.

Keywords High-molecular-weight-glutenin · *Glu-R1* · Orthologous genes · *Secale cereale* · *Triticum aestivum*

Introduction

The high-molecular-weight (HWM) glutenins are thought to influence flour quality (Payne 1987; Odenbach and Mahgoub 1988). Exhaustive studies have been performed to fully characterise these proteins. Payne et al. (1981) identified different glutenin subunits using SDS-PAGE electrophoresis. Other authors undertook the molecular characterisation of the *Glu-1* loci (Forde et al. 1985; Thompson et al. 1985; Halford et al. 1987; Anderson and Greene 1989; Halford et al. 1992; De Bustos et al. 2000). These studies have demonstrated that each *Glu-1* locus contains two tightly linked genes (named x and y) coding for protein subunits of different molecular weight. Generally, glutenin genes are organised into three regions, two highly conserved terminal domains, N and C, and a repetitive central region that consists predominantly of repeats coding for tri-, hexa- and nona-peptides. The number of repeats of these peptides leads to the differences between glutenin subunits.

The corresponding rye proteins were detected by comparing the endosperm protein patterns of wheat and rye using SDS-PAGE electrophoresis. The gene *Glu-R1* was located on the long arm of chromosome 1R (Lawrence and Shepherd 1981; Singh and Shepherd 1984), corresponding to the site of the orthologous genes in the homoeologous chromosomes of wheat and barley (Shewry et al. 1983). Recently, the present authors undertook the molecular characterisation of the x- and y-type genes of the *Glu-R1* locus of rye (De Bustos et al. 2001a). The x- and y-type subunits of the *Glu-R1* gene were characterised using primers designed from a nucleotide sequence of the allele *Glu-D1y10,* which recognised upstream and downstream flanking positions of the coding regions of the genes (De Bustos et al. 2001b). These genes reflect the typical structure of glutenin subunits. Sequence alignment of these genes with previously characterised wheat glutenin genes clearly related the x and y genes of rye and wheat. Consequently, a *Glu-R1x* and *Glu-R1y* designation for these genes was proposed.

This work reports the identification of new alleles of the locus *Glu-R1* isolated from different rye cultivars and their most-closely related wild subspecies. This work was performed in order to clarify this new genetic system and to analyse the relationships of rye HMW subunits with wheat glutenins as a pre-requisite for their use in wheat breeding programmes.

Communicated by J.W. Snape

A. De Bustos · N. Jouve (✉)
Department of Cell Biology and Genetics, University of Alcalá,
28871 Alcalá de Henares (Madrid), Spain,
e-mail: nicolas.jouve@uah.es
Tel.: +34-918854750
Fax: +34-918854799

Materials and methods

Plant material

The characterisation of new *Glu-R1* alleles was carried out in the following plant material: *Triticum aestivum* L. cv 'S-149-3' [2n = 20" + 1R (1D)"], a selected line belonging to a working collection of wheat derived from a cross between the cultivar 'Anza' and the Triticale 'GA-250' (Rubio et al. 1999); the addition line 1R of 'Imperial' rye into common wheat 'Chinese Spring'; the *Secale cereale* L. cultivars 'Imperial', 'Petkus' and 'Smith'; and the wild related subspecies *ancestrale, dighoricum, segetale* and *afghanicum*. Wild material was kindly supplied by Dr. K. Hammer of the Germplasm Bank at Gatersleben (Germany). The wheat cultivars 'Chinese Spring', 'Hope', 'Bezostaya' and the wheat line '7841' (Rubio et al. 1999; De Bustos et al. 2001a) were used as testers in different experiments.

HMW glutenin analyses and molecular characterisation of genes

HMW glutenin analyses were performed using 7.5% SDS-PAGE as previously described by Singh and Shepherd (1991). DNA was purified from young leaves following the method of Sharp et al. (1988). The isolation and characterisation of genes was performed as previously reported by De Bustos et al. (2001a) with some modifications: PCR amplification of genes was performed using primers designed to amplify the *Glu-D1y10* allele (De Bustos et al. 2000, 2001b) except for the type y gene of addition line 1R. PCR reactions were carried out using PfuTurbo DNA polymerase (Stratagene), a high fidelity PCR polymerase. Reactions (in 50-μl volumes containing 100 ng of genomic DNA, 0.2 μM of each primer, 100 μM of each dNTP, 2.5 U of PfuTurbo DNA polymerase and 1 x PfuTurbo DNA polymerase buffer) were performed in a 2400 PCR thermocycler system (Perkin Elmer). The PCR cycle consisted of one cycle of 5 min at 95 °C, 35 cycles of 30 s at 94 °C, 1 min at 60 °C, 3 min at 72 °C and a final extension cycle of 10 min at 72 °C. The conditions for amplifying the type y gene of addition line 1R were the same except that the annealing temperature was 50 °C. The following primers were used: L: 5'CAACCAATCTC-CACAATC3'; R: 5'CTGCAGAGAGTTCTATCA3'. Cloning and sequencing of PCR products was carried out as previously described (De Bustos et al. 2001a). Sequence alignment was performed using the Clustal W 1.5 programme (Thompson et al. 1994). The Phylip programme, version 3.5, (Felsenstein 1993) was used to build phylogenetic trees.

Southern experiments

Clones of genes isolated from all material analysed were digested using *Stu*I and *Nsi*I restriction endonucleases. Wheat HMW subunits Dx5 and Dy10 were amplified by PCR, and cloned and digested with the same enzymes as above. Fragments obtained after endonuclease digestion were fractionated in a 0.8% agarose gel and blotted onto nylon membranes (Roche). *Stu*I and *Nsi*I fragments from Dx5 and Dy10 alleles were digoxigenin labelled (High Prime, Roche) and used as probes. Hybridisation was carried out at 68 °C. LumiAnalyst (Roche) was used for detection.

Results and discussion

Sequence isolation and characterisation

Primers based on published sequences of different alleles of glutenin genes (De Bustos et al. 2000, 2001b) are capable of amplifying orthologous genes in related species (De Bustos et al. 2001a). In previous work, the authors characterised glutenin genes encoded by chromosome 1R of rye. Using the same strategy, new alleles of this locus were identified in the present work. Primers for amplifying the Dy10 allele also amplified products of different size in wheat lines carrying a substitution of chromosome 1R(1D), in rye cultivars and in related wild species. These PCR products were cloned and characterised as previously described (De Bustos et al. 2001a).

Initially, fragment size was taken as a reference to distinguish between the x and y gene types, the x-type being larger than the y-type. Hybridisation experiments were carried out to corroborate this assumption. Clones of PCR products were digested using *Nsi*I and *Stu*I endonucleases. The recognition site of these restriction enzymes, flanking the coding region, is highly conserved between glutenin genes. After digestion, fragments were separated on agarose gels (Fig. 1a) and blotted onto a membrane for hybridisation with the labelled probes Dx5 and Dy10 cloned from the *Glu-1* genes of wheat to test homology with presumptive x- and y-type sequences (Fig. 1b and c respectively). As expected, strong signals were obtained in upper bands with the Dx5 probe, whereas the Dy10 probe hybridised preferentially with smaller fragments. Since the hybridisation signal was obtained even under conditions of high stringency, this experiment also reveals the close relationship between the sequences coding for wheat glutenins and orthologous sequences amplified from genomic DNA of rye.

Subcloning and sequencing of the amplification products was performed as described by De Bustos et al. (2001a). Both x- and y-type genes were found in all plant material analysed except in the 1R rye-wheat addition line, in which only the x-type was clearly identified (see explanation below). The coding sequences of x-type genes ranged from 2,229 bp to 2,435 bp, except for the x gene of 'S-149' where a stop codon was found at position 1,389 bp in the coding sequence (Table 1). For the y-type genes, the size of the coding sequence ranged from 2,136 bp to 2,211. These values are, to some extent, different from those of the wheat orthologous genes. Therefore, x-type rye genes are smaller than wheat genes and y-type rye genes are bigger than wheat y-type genes.

Analysis of sequence and structure of coding region

As expected, the complete nucleotide sequence of these clones showed high homology to wheat genes (data not shown). All lacked introns and showed typical glutenin structure (Anderson and Greene 1989). They contain the 'TATA' box at 85 bp upstream of the ATG codon, identical in position and sequence to the 'TATA' box found in wheat glutenins. The coding sequence is also similar to that of other HMW glutenins with their N and C terminal domains and central repetitive region (Fig. 2). The N-terminal region domain starts after the 21-residue signal peptide, which is almost identical in all glutenins (De Bustos et al. 2001a). This region is composed of 86

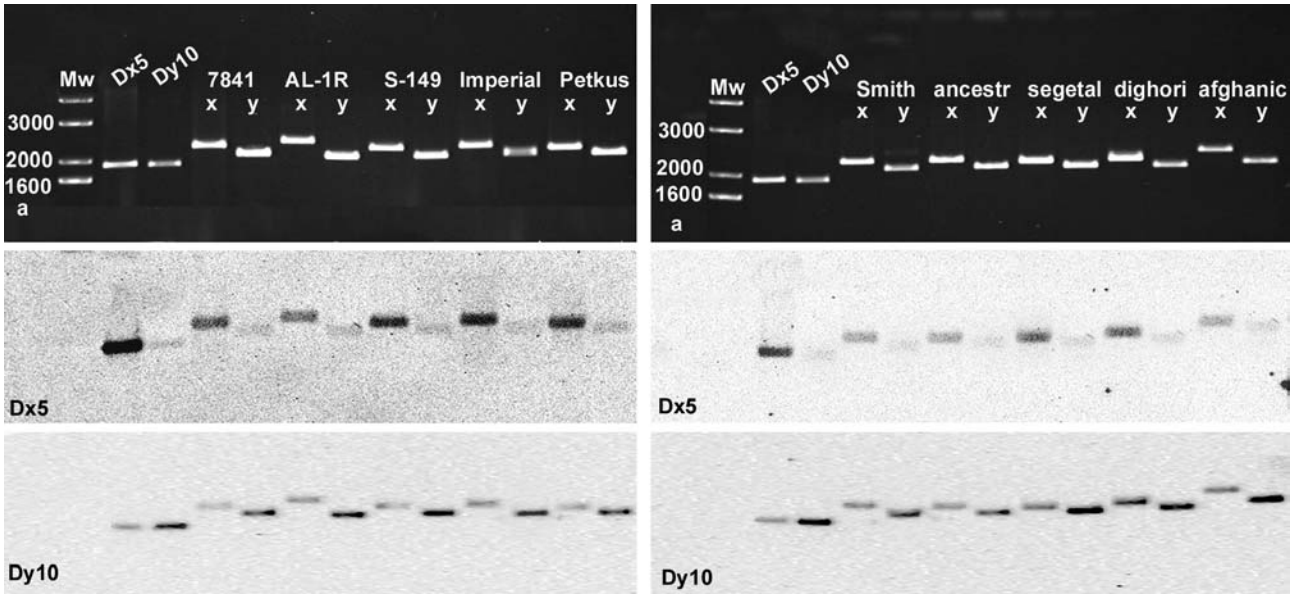

Fig. 1 Hybridisation of clones obtained in the material analysed with Dx5 and Dy10 wheat sequences. *a* Separation of fragments after *Nsi*I and *Stu*I restriction digest of clones. *Lane 1* in both gels contains the molecular marker. *Lanes 2 and 3* contain Dx5 and Dy10 wheat sequences, respectively. Gels were blotted onto a membrane and sequentially hybridised with Dx5 and Dy10 probes at high stringency temperature

Table 1 Plant material used in the study, gene type and name of clones isolated coding for HMW glutenins. The size of coding sequence and number of amino acids of deduced proteins are indicated. Sequences have been deposited in the NCBI Genebank and the accession numbers are shown

Species	Cultivar	Gene type	Clone	Size of coding sequence (bp)	Size of putative protein (residues)	Genebank accession
T. aestivum	S-149	x	pAHTas149-3	1,389(2,295)	462	AJ314772
T. aestivum	S-149	y	pAHTas149-4	2,139	713	AJ314783
T. aestivum	Ad. line 1R	x	PAHTas1Rx	2,433	811	AJ314784
T. aestivum	Ad. line 1R	y	PAHTas1R11	2,160	720	AJ314785
S. cereale ssp. *cereale*	Smith	x	PAHSccs1	2,262	754	AJ314779
S. cereale ssp. *cereale*	Smith	y	PAHSccs5	2,121	707	AJ314767
S. cereale ssp. *cereale*	Petkus	x	PAHSccp19	2,262	754	AJ314778
S. cereale ssp. *cereale*	Petkus	y	PAHSccp3	2,139	713	AJ314780
S. cereale ssp. *cereale*	Imperial	x	PAHScci48	2,343	781	AJ314782
S. cereale ssp. *cereale*	Imperial	y	PAHScci10	2,139	713	AJ314781
S. cereale ssp. *ancestrale*		x	PAHSca2	2,262	754	AJ314773
S. cereale ssp. *ancestrale*		y	PAHSca7	2139	713	AJ314774
S. cereale ssp. *segetale*		x	PAHScs3	2,262	754	AJ314768
S. cereale ssp. *segetale*		y	PAHScs1	2,139	713	AJ314777
S. cereale ssp. *dighoricum*		x	PAHScd2	2,229	743	AJ314776
S. cereale ssp. *dighoricum*		y	PAHScd1	2,139	713	AJ314775
S. cereale ssp. *afghanicum*		x	PAHScaf6	2,433	811	AJ314769
S. cereale ssp. *afghanicum*		y	PAHScaf98	2,211	737	AJ314770

and 104 amino acids in the x- and y-types respectively. However, the C-terminal domain has 42 residues and finishes with the two typical glutenin stop codons in both gene types.

The central repetitive region is composed of the basic repeat motifs described years ago for glutenin genes (Halford et al. 1987): a hexapeptide of PGQGQQ, a nonapeptide GYYPTSPQQ with minor residue changes, and a highly conserved GQQ tripeptide present only in the x-type (Fig. 3). Features earlier described by Anderson and

Fig. 2 Sequence alignment of the deduced amino-acid sequences for x-type (**A**) and y-type (**B**) genes of the HWM glutenins obtained. The *grey box* at the N-terminal region indicates the sequence of the signal peptide. The *small grey boxes* indicate the first and last residues of the central repetitive region. The cysteine residues have also been boxed. Identical residue in all material analysed have been denoted as *. When one or two samples of different amino acids from the rest of the material indicate. or: respectively

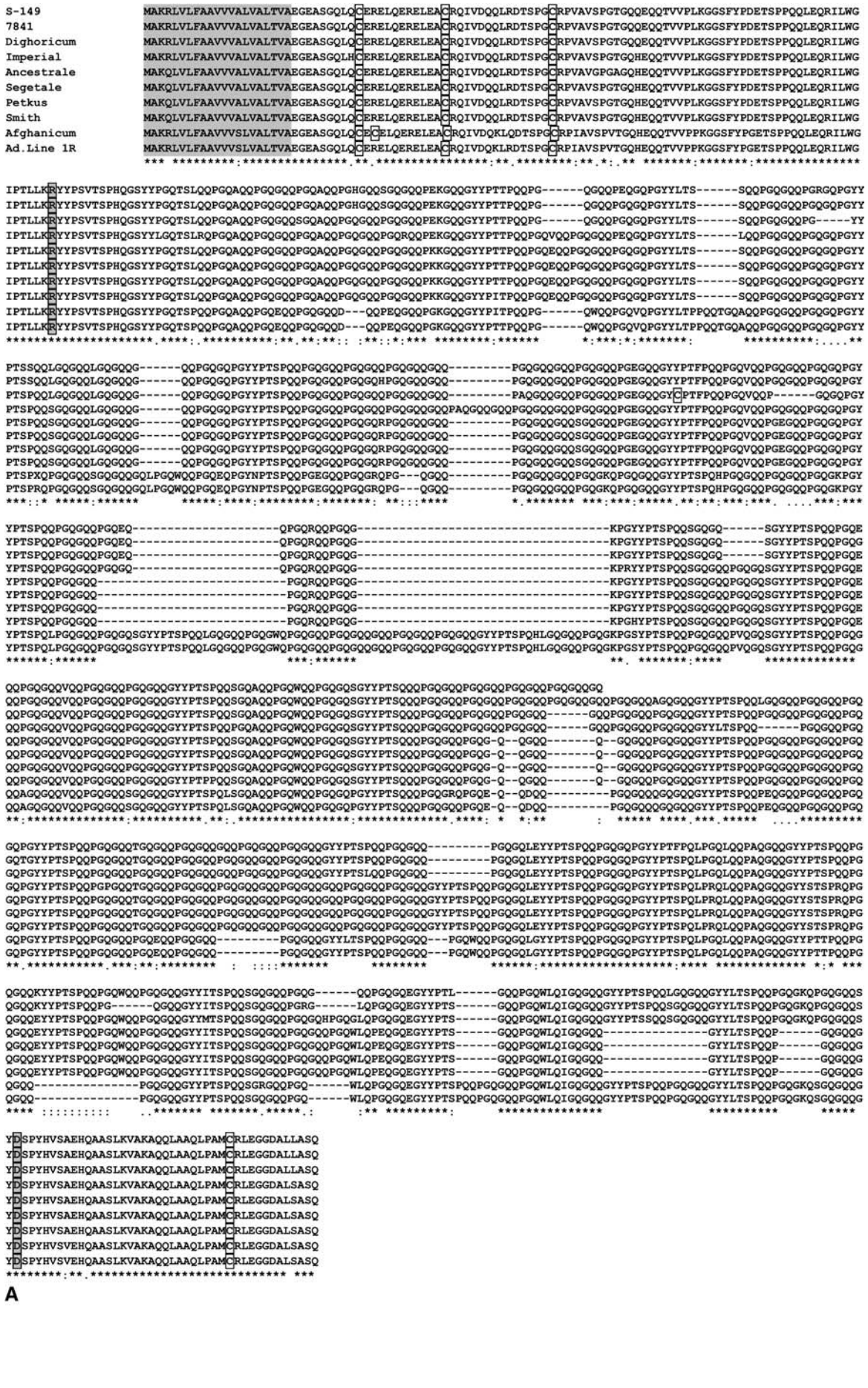
S-149
7841
Dighoricum
Imperial
Ancestrale
Segetale
Petkus
Smith
Afghanicum
Ad.Line 1R
A

Fig. 2

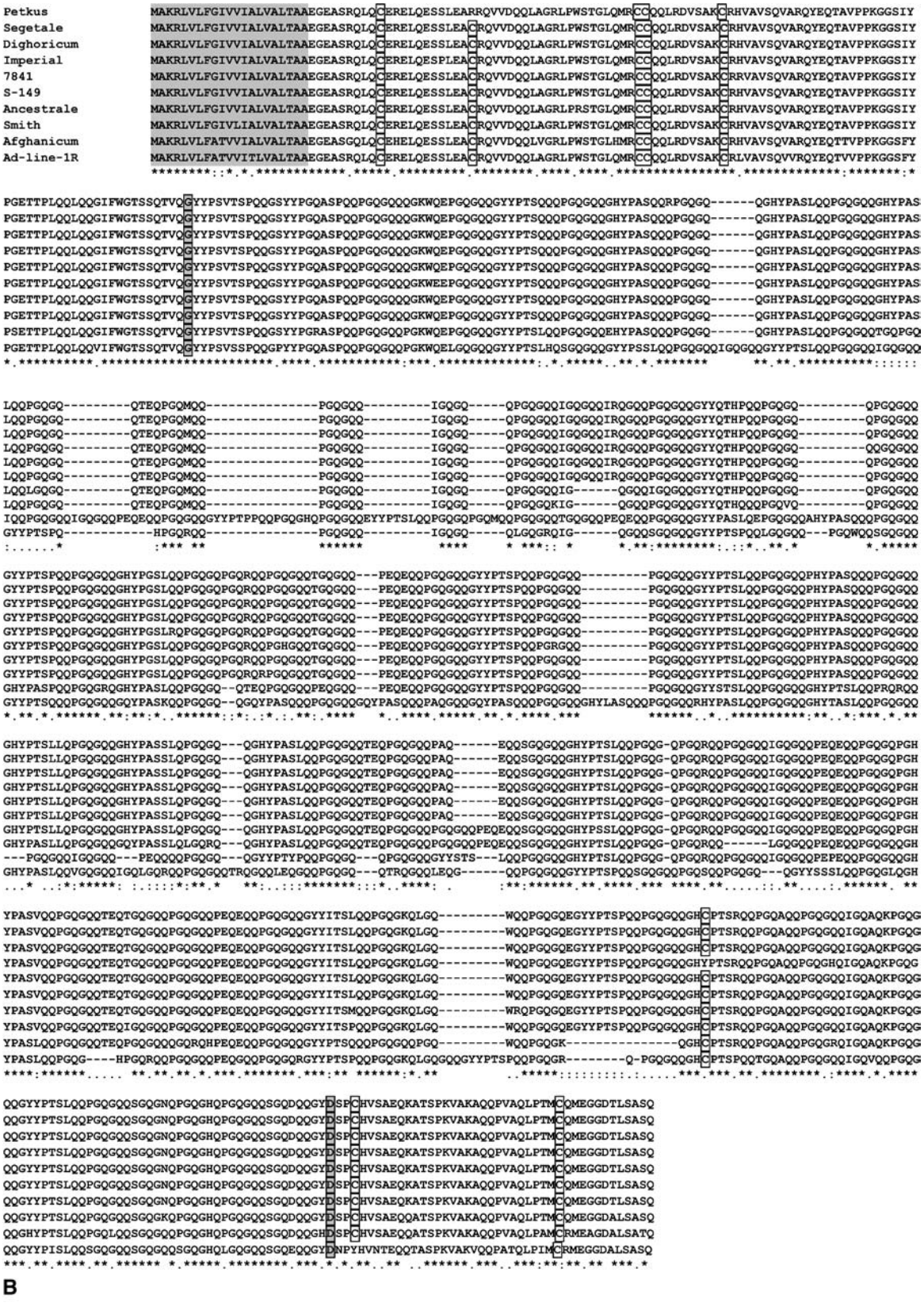

Fig. 2 (Legend see page 76)

type x				type y		
'Smith'	Tripeptide	Hexapeptide	Nonapeptide	'Smith'	Hexapeptide	Nonapeptide
RYYPSVTSPHQ				GYYPSVTSPQQ		
GSYYPGQTSLQQ				GSYYPGQASPQQ		
PGQAQQ				PGQGQQ		
PGQQGQQ				QGKWQE		
PGQAQQ		------ (1R,af)		PGQGQQ GYYPTSQQQ		
PGQGQQ ---	DDQ (1R,af)			PGQGQQ GHYPASQQQ		
PGQQGQQ				PGQGQQ GHYPASLQQ		
PKKGQQ GYYPTTPQQ				PGQGQQ GHYPASLQQ	------ (af)	--------- (af)
PGQEQQ		------ (7,1R,af)		PGQGQQ		
PGQGQQ				TEQ		
PGQGQP GYYLTSSQQ				PGQMQQ		
------		TGQAQQ (1R,af)		PGQGQQ		
PGQGQQ				IGQGQQ		
PGQGQP GYYPTSPQQ		P----- (d)		PGQGQK		
------		PGQGQQ (af)		IGQGQQ	------ (af)	
SGQGQQ				------	IRQGQQ (P,I,S,s,d,7)	
LGQGQQ GQQ				PGQGQQ GYYQTHQQQ		
------		PGQWQQ (1R)		PGQVQQ		
PGQGQP GYYPTSPQQ				PGQGQQ GYYPTSPQQ		
PGQGQQ				PGQGQQ GHYPGSLQQ		--------- (af)
PGQGQR				PGQGQ-		
PGQGQQ GQQ	--- (1R,af)			PGQRQR		
PGQGQQ GQQ				PGQGQQ		
SGQGQQ ---	GQQ (I)			TGQGQQ		
------		PGQGQQ (I)		PEQEQQ		
PGEGQQ GYYPTFPQQ				PGQGQQ GYYPTSPQQ		
PGQVQQ				PGQGQQ	------ (af)	
PGEGQQ		------ (d)		PGQGQQ GYYPTSLQQ		
PGQGQP GYYPTSPQQ				PGQGQQ PHYPASQQQ		
PGQGQQ				PGQGQQ GHYPASLLQ		
------		PGQEQQ (d,I,7)		PGQGQQ GQYPASSLQ	------ (af)	--------- (af)
PGQRQQ		------ (1R,af)		LGQRQQ GHYPASLQQ	------ (af)	--------- (af)
PGQGKP GYYPTSPQQ				PGQGQQ		
SGQGQQ ---------		------ (d)	GYYPTSPQQ (7)	TEQ		
PGQGQS GYYPTSPQQ			--------- (7,1R,af)	PGQGQQ		
PGQEQQ				PGQGQQ	------ (P,I,S,s,d,7)	
PGQGQQ VQQ				PEQEQQ		
PGQGQQ		------ (7)		SGQGQQ GHYPTSLQQ		
PGQGQQ GYYPTSPQQ				PGQGQ-		
SGQAQQ				PGQRQQ ---------		GYYSTSLQQ (af)
PGQWQQ ---------			GSYPTSPQQ (1R,af)	------ ---------	PGQGQQ (af)	GHYPTSLQQ (af)
------		PGQGQQ (1R,af)		------	PGQGQQ (P,I,S,a,s,d,af)	
PGQGQS GYYPTSQQQ				LGQGQQ ---------		GYYPTYPQQ (af)
PGQGQQ				PEQEQQ		
------		PGQGQQ (d,I,7,1R)		------ ---------	PGQGQQ (af)	GYYSTSLQQ (af)
------		PGQGQQ (7)		------ ---------	PGQGQQ (af)	GHYPTSLQQ (af)
PGQGQQ GQQ				------	PGQGQ- (af)	
QGQGQQ		------ (1R)		------	PGQRQQ (af)	
PGQGQQ GYYPTSPQQ				------	PGQGQQ (af)	
PGQGQQ				------	IGQGQQ (af)	
PGQGQQ		------ (I)		------	PEPEQQ (af)	
PGQGQP GYYPTSPQQ				PGQGQQ GHYPASVQQ		
PGQGQQ				PGQGQQ		
TGQGQQ ---	DQQ (1R,af)	------ (I)		TEQ		
PGQGQQ		------ (7)		IGQGQQ		
PGQGQQ GQQ------			GYYPTSPQQ (1R,af)	PGQGQQ		
PGQGQQ GQQ	--- (d,I,7,1R,af)	------ (d,7)		PEQEQQ		
PGQGQQ		------ (d,I)		PGQGQQ GYYITSPQQ		
PGQGQQ GYYPTSPQQ				PGQGKQ		
------		PGQGQQ (d,I,7,1R,af)		LGQWQQ		
------		PGQEQQ (1R,af)		PGQGQE GYYPTSPQQ	------ (af)	--------- (af)
------		PGQGQQ (1R,af)		PGQGQQ GHCPTSRQQ		
PGQGQL EYYPTSPQQ				PGQAQQ		
PGQGQP GYYPTSPQL			--------- (1R,af)	PGQGQQ		
PRQLQQ				IGQAQK		
PAQGQQ GYYSTSPRQ				PGQGQQ GYYPTSLQQ		
PGQGQQ EYYPTSPQQ				PGQGQQ		
PGQWQQ ---	WQQ (I)	------ (d)		SGQGKQ		
PGQGQQ GYYITSPQQ				PGQGHQ		
SGQGQQ				PGQGQQ		
PGQGQQ ---------			GYYPTSPQQ (1R,af)	SGQDQQ		
------		SGQGQQ (1R,af)		GYD		
PGQWLQ		------ (d,7)				
PEQGQE GYYPTSGQQ						
------		PGQGQQ (1R,af)				
PGQWLQ						
IGQGQQ GYYLTSPQQ						
------ ---------		SGQGQQ (d,I,7,1R,af)	GYYLTSPQQ (d,I,7,1R,af)			
------		PGQGKQ (d,I,7,1R,af)				
PGQGQQ						
GYD						

Fig. 3 Repetitive region of the x- and y-types. The amino-acid residues of x- and y-types of cultivar 'Smith' were ordered by repeat classes. Peptide differences of the plant material referred to as the 'Smith' sequence is indicated. Gaps were manually introduced. S: S-149, 1R: addition line 1R, 7: '7841', P: 'Petkus', I: 'Imperial', a: *ancestrale,* d: *dighoricum,* s: *segetale,* af: *afghanicum*

Greene (1989) for wheat genes are also found in rye genes: changes in the repetitive region involve mainly the hexapeptide motif, or a hexapeptide accompanied by either the tripeptide or nona-peptide motifs. Also, the tripeptide and nona-peptide motifs always occur between hexapeptides, and runs of hexapeptide or a hexapeptide plus one of the other two motifs are a common feature in both the x- and y-types. However, contrary to that observed in wheat glutenin genes, only in the y-type sequences of rye are the deletion or duplication of tripeptides and nona-peptides in concert with adjoining hexapeptide motifs. Similarly, hexapeptides flanking nona-peptides with similar residue preferences have been observed in y-type genes, but not in the x-type genes of rye. Finally, a tripeptide TEQ was found three times in the y-type, which contrasts with y-type wheat genes in which tripeptides have been not detected (Halford et al. 1987; Anderson and Greene 1989).

Differences found in the repetitive region could be of applied interest if a relationship between the repetitive structure and visco-elasticity of the dough can be verified. Previous studies have demonstrated that repetitive motifs form regular β-turns (Tatham et al. 1984, 1985) that might contribute to the elastic mechanism of gluten (Shewry et al. 1989). Flavell et al. (1989) and Goldsbrough et al. (1989) have suggested that differences in dough quality rendered by Dy10 and Dy12 wheat alleles are due to the structure of β-turns, which depend on the amino-acid composition of the proteins from both alleles. However, Halford et al. (1992) found insufficient evidence in the repetitive domain of the Ax1 and Ax2* alleles, which affects the conformation of proteins, to explain differences in dough quality. Thus, a deeper analysis of the relationship between protein structure and visco-elasticity is needed to clearly relate the composition of the repetitive domain and dough quality.

Another important factor associated with dough quality is the number and position of cysteine residues (Tatham et al. 1984; Halford et al. 1987). The number and position of cysteines found in the new x-type rye alleles are almost identical to those in the wheat x-type, confirming previous observations (De Bustos et al. 2001a). The exceptions are an extra cysteine residue close to the first cysteine and found at the N-terminal region of the allele in *S. cereale* ssp. *afghanicum* (Fig. 2a), and perhaps more importantly, the additional cysteine residue of *S. cereale* ssp. *dighoricum* located in the repetitive region, as previously seen in the Dx5 allele of wheat (Anderson et al. 1989). The finding of five cysteine residues in the good quality Dx5 compared to the poorer quality related allele Dx2 which has only four, supports a model of a complex disulphide cross-linked glutenin matrix leading to improved dough visco-elasticity (Anderson and Greene 1989). Similarly, there is an extra cysteine residue in the C-terminal region of y-type rye genes compared to those of wheat. This cysteine seems to be specific for rye y-types and has not been found in y-type subunits of wheat. The remaining cysteine residues are positioned as seen for wheat genes, with the exception

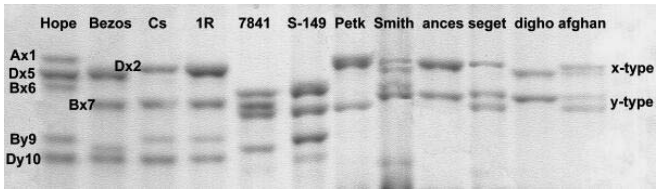

Fig. 4 HMW proteins of the material used in the study analysed by SDS-PAGE. The wheats 'Hope', 'Bezostaya', 'Chinese Spring' (Cs) and line '7841' were used as a glutenin tester. Putative x and y HMW subunits of rye species are indicated on the right

of the y-type rye alleles in 'Petkus' which lack a cysteine at the N-terminal region, and 'Imperial' which is deficient in the cysteine of the repetitive region. As commented above, differences in cysteine residues could mean changes in the conformational structure of proteins that might be related to flour quality. However, no conclusive relationship has been found between the extra cysteines or/and other structural features and quality (Shewry et al. 2001). A way to test the relationship of these proteins with quality would be their use in biotechnological programmes, introducing these genes into wheat and analysing changes in dough quality. At present, in the laboratory of the authors, is realizing the experiments of transformation in wheats of low quality, by means of the introduction of some of the units of the gene *Glu-R1,* which have been characterized in the present work. These experiments might offer some account on the reasons for the low quality of wheat cultivars carrying the rye/wheat substitutions 1R(1A), 1R(1B) and 1R(1D). Also they might be of utility to evaluate comparatively the use of rye glutenin genes to mutate native HMW subunit genes in order to increase cysteine number.

Allelic variation

The allelic variation of HMW glutenin subunits has been sufficiently documented. With the exception of locus *Glu-A1y,* the rest of the glutenin genes have several alleles (Payne and Lawrence 1983). After the characterisation of locus *Glu-R1* (De Bustos et al. 2001a), the aim of this work was to analyse new material for testing allelic variation and therefore help clarify this genetic system.

As shown in Table 1, some clones have the same size, giving the impression that they could be the same allele. A closer look reveals some differences in the x-type. Thus, clones belonging to cultivars 'Petkus' and 'Smith', ssp. *ancestrale* and ssp. *segetale* have the same number of amino acids, but only the x-type clones of 'Petkus' and ssp. *segetale* have identical sequences. These clones code for peptides with amino-acid substitutions (from 2 to 5) with respect to cv 'Smith' and ssp. *ancestrale* (Fig. 2a). Moreover, these differences are due only to single base substitution (data not shown). Since the normal manner of allelic variation in glutenin genes is a change in the number of tri-, hexa- and nona-peptides (Halford et al.

1987) these clones may represent the same allele. A protein of expected molecular weight was seen in samples when SDS separation techniques were used (Fig. 4). This is also the case of the allele found in the addition line 1R and *S. cereale* ssp. *afghanicum,* where only 5 out of 811 residues (0.6%) were different. Sequences found in wheat lines '7841' (De Bustos et al. 2001a) and 'S-149' showed strong similarity. The main difference between them is the presence of two stop codons in 'S-149'. The first stop codon is sited 1,387 bases downstream from the ATG codon. Thus, this allele codes for a protein of 462 residues. This is similar to the early finding of two stop codons in the coding region of the null allele of the *Glu-A1* gene (De Bustos et al. 2000). Also, a silent allele has been recently identified in *Aegilops cylindrica* (Wan et al. 2002), representing the null allele of the *Glu-D1x* gene, due to a single base change resulting in a stop codon. Bearing these precedents in mind, it is proposed that the allele found in 'S-149' might represent the null allele of the *Glu-1R* gene. The other two clones found in cv 'Imperial' and in the wild ssp. *dighoricum* showed differences in the number of repeats of tri-, hexa- and nona-peptides, and consequently represent two new alleles of this gene. On the basis of everything indicated, six different alleles of the x-type of the *Glu-R1* locus were characterised.

Variation in the y-type clones was also observed (Table 1). Seven out of ten sequences analysed, including those of '7841' (De Bustos et al. 2001a), showed the same number of amino acids: 713. Only a few single changes in amino-acid residues (maximum of seven) were detected between the sequences analysed in the rye-wheat substitution lines '7841' and 'S-149', the rye cultivars 'Imperial' and 'Petkus', and the wild subspecies *dighoricum* and *segetale*. No differences were detected between the predicted proteins of ssp. *dighoricum* and ssp. *segetale*. A sequence was also found with the same size in the wild ssp. *ancestrale*. Nevertheless, changes in hexapeptide motifs were detected, so this sequence probably represents another allele of the y-type gene subunit. In the same way, the clone found in the rye cultivar 'Smith' would represent an allele almost identical to those of *ancestrale,* with the minor variation of the absence of a hexapeptide motif. Finally, the clone found in ssp. *afghanicum* represents the allele most separated from the rest of the alleles found at this locus.

The y-type sequence cloned from addition line 1R differs widely from the other rye y-type genes (Fig. 2b). For this line, no amplification was obtained when the Dy10 allele primers were used. Two new pairs of primers were designed based on the new y-type sequences obtained in this work. After cloning the amplification products obtained with these primers, an exhaustive search was performed that finally led to the finding of a single clone belonging to the y-type included in the study. Nevertheless, a more detailed look at this sequence revealed differences with respect to other rye y-type genes. Finally, alignment of protein sequences including wheat glutenins determined the relationship of this clone

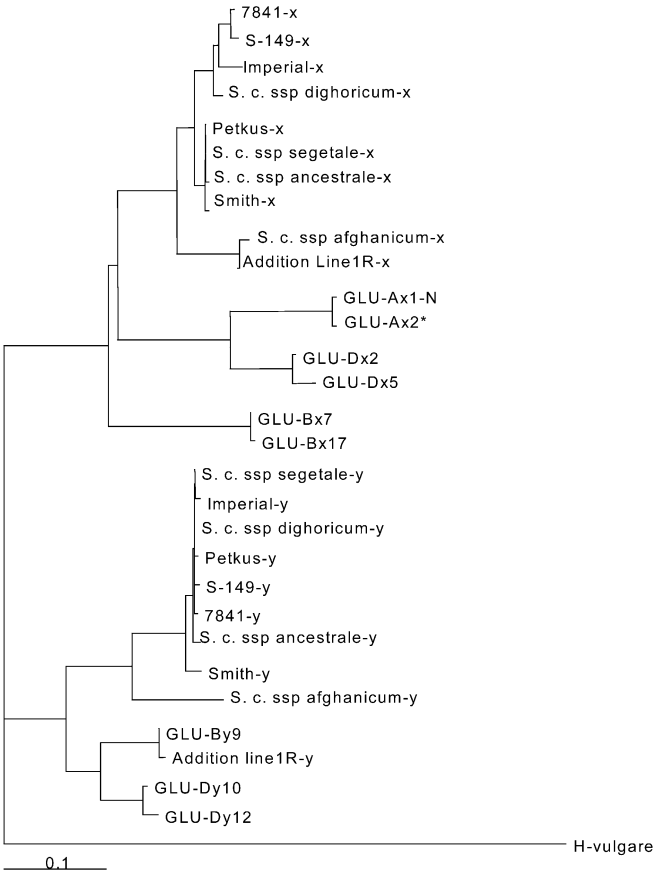

Fig. 5 The phylogenetic tree drawn from the alignment of the deduced amino-acid sequences of glutenin clones obtained from the material analysed and other published sequences of HMW glutenins of wheat. *Hordeum vulgare* was taken as an outgroup

with the *Glu-B1y9* allele of wheat (Fig. 5). Sequence comparison (data not shown) clearly revealed their strong similarity. Nevertheless, the sequence from the 1R rye-wheat addition was 45 nucleotides longer than the *Glu-B1y9* allele, due to the insertion of a hexapeptide and a nona-peptide motif. This result demonstrates the close relationship between the wheat and rye glutenin genes, such that even when using primers based on the rye y-type a wheat y-type subunit was obtained. The amplification product obtained in the addition line might be the sequence belonging to the *Glu-B1y9* subunit present in the wheat 'Chinese Spring' used to obtain this line (Driscoll and Sears 1971). This type of difference was recently noticed by sequencing different alleles of the same glutenin genes in wheat. Thus, Halford et al. (1987) detected differences between the *Glu-B1x7* allele of the cultivar 'Cheyenne' and that already published for 'Chinese Spring' (Thompson et al. 1983). Differences were also found in the present work with respect to the sequences of the coding region of the allele *Glu-D1x5* of 'Cheyenne' (Anderson et al. 1989) and 'Hope' (unpublished).

The inability to amplify and characterize the sequence of y-type *Glu-R1* glutenin subunitsin 1R of 'Imperial' in

the rye-wheat addition line has two explanations. The first depends on the capacity of PCR to specifically amplify the correct sequence, which in turn depends on the conditions needed for the annealing of the designed primers. Secondly, the sequences may have been altered by the presence of retrotransposons that prevent any amplification by PCR, as occurs with the *Glu-A1y* subunit in wheat (Harberd et al. 1987).

The alleles of the rye glutenin genes *Glu-R1x* and *Glu-R1y* seem to present more variability that their corresponding orthologs in wheat. Assuming that changes in the number of tri-, hexa- and nona-peptides are the only mechanism responsible for allelic variation in glutenin subunits, then five x-type and four y-type alleles were isolated. This means that *Glu R1* has more variation than that detected so far in wheat. This result could be related to the allogamous breeding system of *S. cereale* (Allard 1975; Hammer 1990). It is known that the allogamous species show greater variation than autogamous species (Brown 1979). Also, as observed in Fig. 5, orthologous genes from loci *Glu-A1, Glu-B1, Glu-D1* and *Glu-R1* (x- versus x-types, and y- versus y-types) are more closely related than paralogous genes (x- versus y-types). This supports the idea that both types of gene were duplicated before the speciation process of the Triticeae. Anderson et al. (2002) came to the same conclusion in a recent work in which the sequences of the flanking regions of HMW glutenin genes were analysed comparatively.

Acknowledgements The authors thank the MCYT (*Ministerio de Ciencia y Tecnología*) of Spain for financial support of this work (Grant No. AGL2000-0762) and to Adrian Burton for linguistic assistance.

References

Allard RW (1975) The mating system and microevolution. Genetics 79:115–126

Anderson OD, Greene FC (1989) The characterisation and comparative analysis of high-molecular-weight glutenin genes from genomes A and B of a hexaploid bread wheat. Theor Appl Genet 77:689–700

Anderson OD, Greene FC, Yip RE, Halford NG, Shewry PR, Malpica-Romero JM (1989) Nucleotide sequences of the two high-molecular-weight glutenin genes from the D-genome of a hexaploid bread wheat, *Triticum aestivum* L. cv Cheyenne. Nucleic Acids Res 17:461–462

Anderson OD, Larka L, Christoffers MJ, McCue KF, Gustafson JP (2002) Comparison of orthologous and paralogous DNA flanking the wheat high-molecular-weight glutenin genes: sequence conservation and divergence, transposon distribution, and matrix-attachment regions. Genome 45:367–380

Brown ADH (1979). Enzyme polymorphism in plant populations. Theor Pop Biol 15:1–42

De Bustos A, Rubio P, Jouve N (2000) Molecular characterisation of the inactive allele of the gene *Glu-A1* and the development of a set of AS-PCR markers for HMW glutenins of wheat. Theor Appl Genet 100:1085–1094

De Bustos A, Rubio P, Jouve N (2001a) Characterisation of two gene subunits on the 1R chromosome of rye as orthologs of each the *Glu*-1 genes of hexaploid wheat. Theor Appl Genet 103:733–742

De Bustos A, Rubio P, Soler C, García P, Jouve N (2001b) Marker-assisted selection to improve HMW-glutenins in wheat. Euphytica 119:69–73

Driscoll CJ, Sears ER (1971) Individual addition of the chromosomes of 'Imperial' rye to wheat. In: Agronomy abstracts. American Society of Agronomy, Madison, Wisconsin, p 6

Felsenstein J (1993) PHYLIP (phylogeny inference package). Version 3.5c. Distributed by the author, Department of Genetics, University of Washington, Seattle

Flavell RB, Goldsbrough AP, Robert LS, Schnick D, Thompson RD (1989) Genetic variation in wheat HMW glutenin subunits and the molecular basis of breadmaking quality. BioTechnology 7:1281–1285

Forde J, Malpica JM, Halford NG, Shewry PR, Anderson OD, Greene FC, Miflin BJ (1985) The nucleotide sequence of a HMW glutenin subunit gene located on chromosome 1A of wheat (*Triticum aestivum* L.). Nucleic Acids Res 13:6817–6832

Goldsbrough AP, Bulleid NJ, Freedman RB, Flavell RB (1989) Conformational differences between two wheat (*Triticum aestivum*) 'high-molecular-weight' glutenin subunits are due to a short region containing six amino-acid differences. Biochem J 263:837–842

Halford NG, Forde J, Anderson OD, Greene FC, Shewry PR (1987) The nucleotide and deduced amino-acid sequences of an HMW glutenin subunit gene from chromosome 1B of bread wheat (*Triticum aestivum* L.) and comparison with those of genes from chromosomes 1A and 1D. Theor Appl Genet 75:117–126

Halford NG, Field JM, Blair H, Urwin P, Moore K, Robert L, Thompson R, Flavell RB, Thatam, AR, Shewry PR (1992) Analysis of the HMW glutenin subunit encoded by chromosome 1A of bread wheat: (*Triticum aestivum* L.) indicates quantitative effects on grain quality. Theor Appl Genet 83:373–378

Hammer K (1990) Breeding system and phylogenetic relationships in *Secale* L. Biol Zentral bl 109:45–50

Harberd NP, Flavell RB, Thompson RD (1987) Identification of a transposon – like insertion in a *Glu-1* allele of wheat. Mol Gen Genet 209:326–332

Lawrence GJ, Shepherd KW (1981) Inheritance of glutenin protein subunits of wheat. Mapping the nucleolus organiser region, seed protein loci and isozyme loci on chromosome 1R in rye. Theor Appl Genet 60:333–337

Odenbach W, Mahgoub El-S (1988) Relationships between HMW glutenin subunit composition and the sedimentation value in reciprocal sets of inbred backcross lines derived from two winter wheat crosses. In: Miller TE, Koebner RMD (eds) Proc 7th Int Wheat Genet Symp, Cambridge, pp 987–991

Payne PI (1987) Genetics of wheat storage proteins and the effect of allelic variation on breadmaking quality. Annu Rev Plant Physiol 38:141–153

Payne PI, Lawrence GD (1983) Catalogue of alleles for the complex gene loci, *Glu*-A1, *Glu*-B1 and *Glu*-D1, which code for high-molecular-weight subunits of glutenin in hexaploid wheat. Cereal Res Commun 11:29–30

Payne PI, Corfield KG, Holt LM, Blackman JA (1981) Correlation between the inheritance of certain high-molecular-weight subunits of glutenin and breadmaking quality in progenies of six crosses of bread wheat. J Sci Food Agric 32:51–60

Rubio P, Daza L, Jouve N (1999) Meiotic behaviour, chromosome stability and genetic analysis of the preferential transmission of 1B-1R, 1A-1R and 1R(1D) chromosomes in intervarietal hybrids of wheat. Agronomie 19:57–68

Sharp PJ, Desai S, Gale MD (1988) Isozyme variation and RFLPs at the á-amylase loci in wheat. Theor Appl Genet 76:691–699

Shewry PR, Miflin B, Kasarda DD (1983) The structural and evolutionary relationships of the prolamin storage proteins of barley, rye and wheat. Phil Trans R Soc Lond Ser B 304:697–308

Shewry PR, Halford NG, Tatham AS (1989) The high-molecular-weight subunits of wheat, barley and rye: genetics, molecular biology, chemistry and role in wheat gluten structure and functionality. In: Miflin BJ (ed) Oxford surveys of plant

molecular and cell biology, vol 6. Oxford University Press, Oxford, pp 163–219
Shewry PR, Popineau Y, Lafiandra D, Belton P (2001) Wheat glutenin subunits and dough elasticity: findings of the EUROWHEAT project. Trends Food Sci Tech 11:433–441
Singh NK, Shepherd KW (1984) Mapping of the genes controlling high-molecular-weight glutenin subunits of rye on the long arm of chromosome 1R. Genet Res 44:117–123
Singh NK, Shepherd KW (1991) A simplified SDS-PAGE procedure for separating LMW subunits of glutenin. J Cereal Sci 14:203–208
Tatham AS, Shewry PR, Miflin BJ (1984) Wheat gluten elasticity: a similar molecular basis to elastin? FEBS Lett 177:205–208
Tatham AS, Miflin BJ, Shewry PR (1985) The beta-turn conformation in wheat gluten proteins: relationship to gluten elasticity. Cereal Chem 62:405–412
Thompson J D, Higgins DG, Gibson TJ (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting matrix choice. Nucleic Acids Res 22:4673–4680
Thompson RD, Bartels D, Harberd NP, Flavell RB (1983) Characterisation of the multigene family coding for HMW glutenin subunits in wheat using cDNA clones. Theor Appl Genet 67:87–96
Thompson RD, Bartels D, Harberd NP, Flavell RB (1985) Nucleotide sequence of a gene from chromosome 1D of wheat encoding a HMW glutenin subunit. Nucleic Acids Res 13:6833–6846
Wan Y, Wang D, Shewry PR, Halford NG (2002) Isolation and characterization of five novel high-molecular-weight subunits of glutenin genes from *Triticum timopheevi* and *Aegilops cylindrica*. Theor Appl Genet 104:828–839