Mol Gen Genet (1987) 209:326–332

MGG

Identification of a transposon-like insertion in a *Glu-1* allele of wheat

N.P. Harberd*, R.B. Flavell, and R.D. Thompson**
Molecular Genetics Department, Plant Breeding Institute, Maris Lane, Trumpington, Cambridge, CB2 2LQ, UK

Summary. The *Glu-1* locus, present on the long arms of the group 1 chromosomes of wheat, codes for a group of storage protein polypeptides termed high molecular weight (HMW) subunits of glutenin. Hexaploid wheat varieties carry a 'silent' *Glu-1y* allele on chromosome 1A, no polypeptide being attributable to this locus. When two such alleles from different varieties were compared, one was found to contain an 8 kb insertion of DNA, termed Wis-2, interrupting the coding sequence. The insertion site is flanked by a 5 bp duplication. The two ends of Wis-2 contain similar sequences over 500 bp long and its termini contain almost the same short sequences but in opposite orientation. These terminal sequences are related to those of several 'retroposon'-type transposable elements found in other organisms.

Key words: HMW glutenin gene – Wheat – Transposon

Introduction

The high molecular weight (HMW) glutenin genes encode a small group of HMW polypeptide subunits which comprise part of the glutenin aggregate found in wheat endosperm (Kasarda et al. 1976). These subunits play a major role in determining the breadmaking quality of wheat flour (Payne et al. 1984), and genetic variation in this group of polypeptides has been subject to intensive investigation (Payne et al. 1980, 1981, 1982). Isolation of a HMW glutenin cDNA has allowed identification of restriction fragments containing HMW glutenin genes in digests of DNA from a number of wheat varieties. The chromosomal origins of these fragments have been determined by the use of null-isomic-tetrasomic and intervarietal chromosome substitution lines (Thompson et al. 1983; Harberd et al. 1986). Each of the 1A, 1B and 1D chromosomes of wheat carries two closely linked genes, one specifying an 'x' and the other specifying a 'y' type HMW glutenin subunit (Payne et al. 1981; Harberd et al. 1986).

The HMW glutenin genes located on the 1A chromosome of wheat (Thompson et al. 1983) are of particular interest. In endosperm of the variety Chinese Spring four HMW glutenin subunits are synthesized, two of which are specified by chromosome 1B, the other two by chromosome 1D (Payne et al. 1981). Thus the HMW glutenin genes on chromosome 1A of Chinese Spring do not direct the synthesis of a HMW glutenin polypeptide. This block to expression appears to be at the level of transcription since there are no detectable HMW glutenin transcription products derived from chromosome 1A of Chinese Spring wheat (Thompson et al. 1983).

In some other wheat varieties, for example Cheyenne, chromosome 1A does specify a HMW glutenin polypeptide. This polypeptide is of the x type (Payne et al. 1983). In modern wheat varieties the y type subunit from chromosome 1A is always absent and the *Glu-Aly* sublocus gene encoding it is silent. The DNA sequence of a cloned *Glu-Aly* sublocus gene from the variety Cheyenne has been reported recently (Forde et al. 1985). As expected, this gene appears to be silent because it has a stop codon in the appropriate reading frame and no translation product corresponding to it was found in hybrid-selection translation experiments.

This paper describes the characterization of a cloned wheat DNA fragment containing the *Glu-Aly* sublocus gene from the variety Chinese Spring. The partial gene sequence obtained is compared with that of its allelic counterpart from the variety Cheyenne. This comparison permits the identification of two kinds of difference between these allelic genes. Firstly, the genes differ by the presence of an inserted DNA segment in the coding region of *Glu-Aly* of Chinese Spring and not present in *Glu-Aly* of Cheyenne. Secondly, they differ in the number of repeat units found in their central repetitive regions (Thompson et al. 1985; Harberd et al. 1986). These observations provide information on the divergence of DNA sequences during the relatively short period of time since the separation of these two varieties.

Materials and methods

A cloned wheat DNA *Eco*RI fragment derived from chromosome 1A of Chinese Spring wheat and containing the *Glu-Aly* sublocus gene was isolated and identified using procedures described previously (Thompson et al. 1985; Harberd et al. 1986). Restriction enzyme site mapping, Southern hybridizations, and DNA sequencing were performed as described in previous publications (Thompson et al. 1985, 1983). DNA sequence data were compiled using the

* *Present address:* Department of Genetics, 345 Mulford Hall, University of California at Berkeley, CA 94720, USA
** *Present address:* Max-Planck-Institut für Züchtungsforschung, Egelspfad, D-5000 Köln 30, Federal Republic of Germany
Offprint requests to: R. Flavell

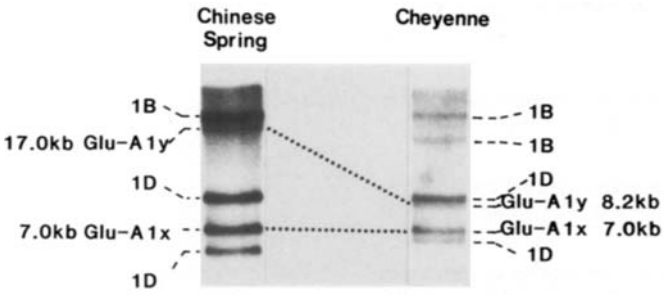


Fig. 1. Hybridization of the HMW glutenin cDNA pTag1290 to *Eco*RI-digested DNA from the wheat varieties Chinese Spring and Cheyenne. Chromosomal locations of fragments (1A, 1B, 1D) and gene designations (*Glu-Alx, Glu-Aly*) were determined as described in Harberd et al. (1986)

Staden programs (Staden 1982a, b) on a VAX-VMS computer.

Results

Restriction fragment variation between the 1A chromosomes of wheat varieties Chinese Spring and Cheyenne

Hybridization of an HMW glutenin cDNA to *Eco*RI digests of nullisomic-tetrasomic wheat DNAs has revealed the existence of two hybridizing fragments (17.0 kb and 7.0 kb) derived from chromosome 1A of the variety Chinese Spring (Thompson et al. 1983; Harberd et al. 1986). The variety Cheyenne also has two hybridizing *Eco*RI fragments derived from chromosome 1A, but one of them is smaller (8.2 kb) than its counterpart in Chinese Spring (17.0 kb, Fig. 1; Forde et al. 1985; Harberd et al. 1986). The DNA

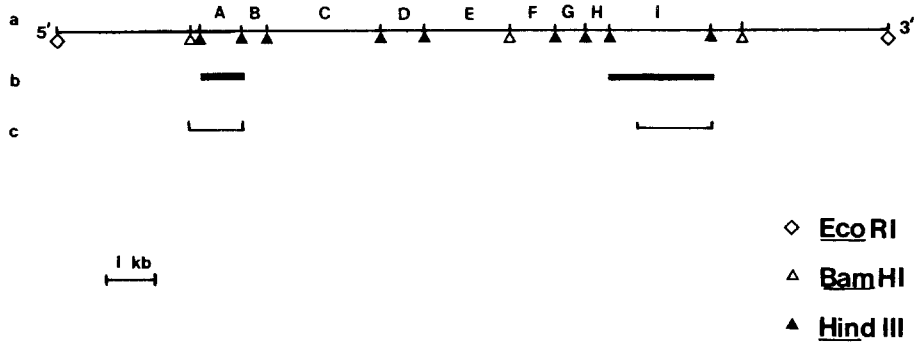
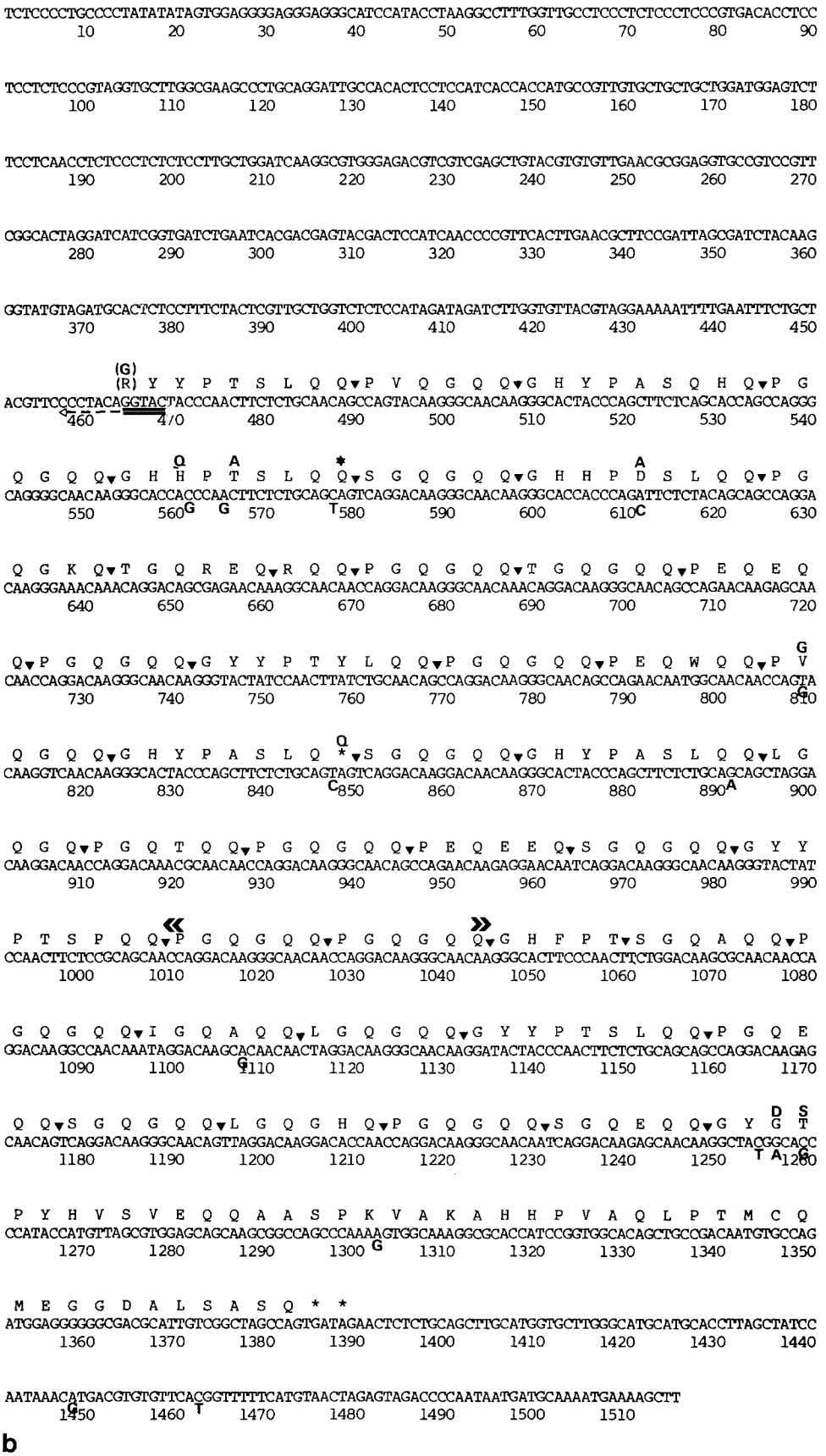


Fig. 2. a Restriction endonuclease map of the pEco5 fragment containing the *Glu-Aly* gene from Chinese Spring. A to I represent the *Hin*dIII and *Bam*HI-*Hin*dIII fragments used for the genome organization studies (see Fig. 5). Fragments A (0.8 kb) and I (2.0 kb) hybridize to the pTag1290 HMW glutenin cDNA probe, as represented by the *solid bars* (**b**); the fragments between A and I contain DNA from the insert. The *thin bars* (**c**) show the regions sequenced

```
  A  S  P  Q  Q▼P  G  K  W  Q  E▼L  G  Q  G  Q  Q▼W  Y  Y  P  T  S  L  Q  Q▼P  G  Q  G
AAGCTTCTCCACAACAGCCAGGAAAATGGCAAGAACTAGGACAAGGGCAACAATGGTACTATCCAACTTCTCTGCAGCAGCCAGGACAAG
       10        20        30        40        50        60        70        80        90

                                                                                  L
  Q  Q▼G  Y  Y  R  T  S  L  Q  Q▼P  G  Q  R  Q  Q▼G  Y  Y  R  T  S  R  Q  Q▼P  G  Q  G
GGCAACAAGGGTACTACCGAACTTCTCTGCAGCAGCCAGGACAAAGGCAACAAGGGTACTACCGAACTTCTCGGCAGCAGCCAGGACAAG
       100       110       120       130       140       150       160  T    170       180

  Q  Q▼I  G  Q  W  Q  Q▼G  Y  Y  P  T  S  P  Q  H▼P  G  Q  G  Q  Q▼P  G  Q  V  Q  K▼I
GGCAACAGATAGGACAATGGCAACAAGGGTACTACCCAACTTCTCCACAGCACCCAGGACAAGGGCAACAACCAGGACAAGTGCAAAAAA
       190       200       210       220      G230      240       250       260       270

  G  Q  G  Q  Q▼P  E  K  G  Q  Q▼L  G  Q  E  Q  Q▼I  G  Q  G  Q  Q▼P  E  Q  G  Q  Q▼P
TAGGACAAGGGCAACAACCAGAAAAAGGGCAACAACTAGGACAAGAGCAACAAATAGGACAAGGGCAACAACCAGAACAAGGGCAACAAC
       280       290       300       310       320       330       340       350       360

                                                                                  W
  G  Q  G  Q  Q▼P  G  Q  G  Q  Q▼G  Y  Y  P  T  S  L  Q  Q▼P  G  Q  G  Q  Q▼P  G  Q  *
CAGGACAAGGGCAACAACCAGGACAAGGGCAACAAGGGTACTACCCAACTTCTCTGCAGCAGCCAGGACAAGGGCAACAACCAGGACAAT
       370       380       390       400       410       420       430       440       450

  Q  Q▼P  G  Q  G  Q  Q▼G  Y
GACAACAACCAGGACAAGGGCAACAAGGGTACTGTTGGAAATATGCCCTAGAGGCAATAATAAATTGATTATTATTATATTTCCTNGTTC
G      460       470       480  - - 490 ->       500       510       520       530       540

ATGATAATCGTTTATTATCCATGCTAGAATTGTATTGATAGGAAACTCAGATACATGTGTGGATACATAGACAACACCATGTCCCTAGTA
       550       560       570       580       590       600       610       620       630

AGCCTCTAGTTGACTAGCTCGTTAATCAATAGATGGTTACGGTTTCCTGACCATGGACATTGGATGTCATTGATAACGGGATCACATCAT
       640       650       660       670       680       690       700       710       720

TAGGAGAATGATGTGATGAACAAGACCCAATCCTAAGCCTAGCACAAGATCATGTAGTTCGTNNNCTAAAGCTT
       730       740       750       760       770       780       790
a
```

Fig. 3a, b. DNA sequences of **a** the 0.8 kb *Hin*dIII fragment (A) and **b** the 2.0 kb *Hin*dIII fragment (I) from the pEco5 insert containing the *Glu-Aly* sublocus gene from Chinese Spring. The DNA sequence is shown together with a one-letter code translation of the coding regions of sequence. The positions of hexamer and monomer repeat units within the coding sequence are indicated. ▼. Apparent nucleotide and amino acid substitution differences between the *Glu-Aly* alleles of Chinese Spring and Cheyenne (sequence of Cheyenne *Glu-Aly* from Forde et al. 1985) are shown. The Cheyenne sequence is given where different from Chinese Spring. The DNA duplications flanking the insertion are *doubly underlined* and the inverted terminal sequences of the insertion are *underlined*. The *double chevron* (**b**) indicates the position of the hexamer repeat unit duplicated in the Chinese Spring *Glu-Aly* sequence with respect to the Cheyenne *Glu-Aly* sequence



b

Fig. 3 (cont.)

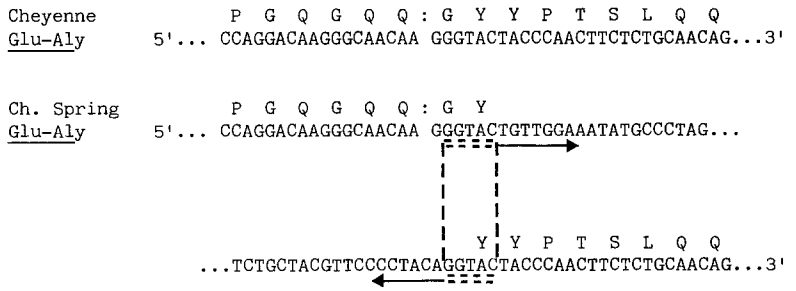


Fig. 4. Comparison of the DNA and amino acid sequences of the *Glu-Aly* alleles of Cheyenne and Chinese Spring in the region of insertion showing the 5 bp target DNA duplication. Sequence of *Glu-Aly* from Cheyenne is reported in Forde et al. (1985)

sequence of the HMW glutenin gene present on the 8.2 kb *Eco*RI fragment from Cheyenne has recently been determined (Forde et al. 1985). This gene is thought to be the *Glu-Aly* sublocus gene of that variety (Forde et al. 1985; Harberd et al. 1986). In order to compare allelic genes from two varieties we isolated and characterized the *Glu-Aly* sublocus gene from Chinese Spring by cloning the 17.0 kb *Eco*RI fragment from that variety.

Cloning and restriction mapping of the Glu-Aly sublocus from Chinese Spring

A wheat genomic DNA library was constructed by ligation of partial *Eco*RI digest fragments of Chinese Spring DNA into the *Eco*RI site of the λ cloning vector Charon 34 (Thompson et al. 1985). Several clones containing a 17.0 kb *Eco*RI fragment which hybridized to the HMW glutenin cDNA were isolated from this library. The 17.0 kb fragment from one of these was subcloned into pUC9 and shown to be derived from chromosome 1A of Chinese Spring by comparison of the hybridization of the HMW glutenin cDNA to *Eco*RI, *Bam*HI and *Hin*dIII digests of the subclone to that of wheat DNA digested with the same restriction enzymes (data not shown, see Thompson et al. 1985). The clone containing this fragment is designated pEco5.

A restriction map of this fragment was determined (Fig. 2). The regions of the fragment containing the HMW glutenin sequence were identified by hybridization of digests of the cloned DNA (and further subclones of it) to ^{32}P-labelled HMW glutenin cDNA and also to ^{32}P-labelled poly(A)$^+$ RNA from endosperm. Interestingly, two regions of the clone hybridize to these probes, and these regions are separated by approximately 8.0 kb of non-hybridizing DNA. The two regions of hybridization are confined to *Hin*dIII fragments of 0.8 kb and 2.0 kb. These two fragments were subcloned for DNA sequencing, and the regions sequenced are indicated in Fig. 2.

DNA sequencing of selected regions of the Chinese Spring Glu-Aly gene

DNA from the subcloned 0.8 and 2.0 kb *Hin*dIII fragments was further subcloned into M13 vectors and sequenced by the dideoxy method (Sanger et al. 1980). Clones for sequencing were constructed by sonication (Deininger 1983), Bal31 deletion (Poncz et al. 1982) or by generation of specific restriction endonuclease fragments followed by cloning in M13. The complete sequence of the 0.8 kb fragment and the sequence of part of the 2.0 kb fragment are as shown (Fig. 3a, b). The gene sequences found in these fragments are almost identical to regions of the *Glu-Aly* sublocus gene of Cheyenne. For example, the 5′ *Hin*dIII site of the 0.8 kb fragment is also found in the Cheyenne gene (*Hin*dIII cleaves between nucleotide 1244 and 1245 of the sequence published in Forde et al. 1985), and the sequences continue to be nearly identical for a further 479 nucleotides 3′ to this site (Fig. 3a). However, at nucleotide 483 or 484 (1724 or 1725 of the sequence of Forde et al. 1985) there is an abrupt change. Sequence 3′ to this nucleotide no longer encodes an HMW glutenin subunit and is unrelated to the coding sequence found in this region of the Cheyenne *Glu-Aly* sublocus gene sequence. The Chinese Spring chromosome 1A fragment returns to HMW glutenin coding sequence some 8 kb further downstream of this break at nucleotide 464 of the sequence obtained from the 2.0 kb *Hin*dIII fragment (Fig. 3b). Significantly, it is a continuation of the *Glu-Aly* sequence from exactly the position where it had previously been interrupted. The sequence then continues without interruption to the 3′ *Hin*dIII site located in the 3′ untranslated region of the gene (Forde et al. 1985). It seems likely that the discontinuity in the coding sequence of the *Glu-Aly* sublocus gene of Chinese Spring has been caused by the insertion of an 8 kb segment of unrelated DNA into the gene (Fig. 4). This insertion has caused a short 'target' duplication since the 5 or 6 bp of coding sequence on either side of the insertion are duplicated with respect to this region of the Cheyenne *Glu-Aly* (Fig. 4).

There are a number of other differences between the sequences of the Chinese Spring and Cheyenne *Glu-Aly* sublocus alleles. Several nucleotide substitutions are apparent and the sequences also differ by the presence of one extra repeat unit in the central repetitive section of the Chinese Spring allele (Fig. 3b). However, the major difference between these two alleles is due to the presence of an unrelated DNA segment inserted into the Chinese Spring *Glu-Aly* allele which is not present in the Cheyenne *Glu-Aly* allele.

Estimation of multiplicity of sequences from within the Chinese Spring Glu-Aly insertion in the wheat genome by DNA hybridization

In order to investigate the nature of the DNA segment inserted into the Chinese Spring *Glu-Aly* allele, DNA fragments derived from this region were hybridized to digests of total wheat DNA. The DNA fragments labelled A to I inclusive (defined by *Hin*dIII and *Bam*HI sites, Fig. 2) were purified from digests of subclones of the *Glu-Al* sublocus gene. These fragments cover the entire inserted DNA segment, together with some HMW glutenin coding sequences in those fragments covering the termini of that insert. The fragments were nick translated and hybridized to digests of total wheat DNA (Fig. 5) and to digests of known quantities of pEco5 (the plasmid containing the wheat chromosome 1A fragment carrying the *Glu-Aly* sublocus gene) on the same filter (not shown). The resulting

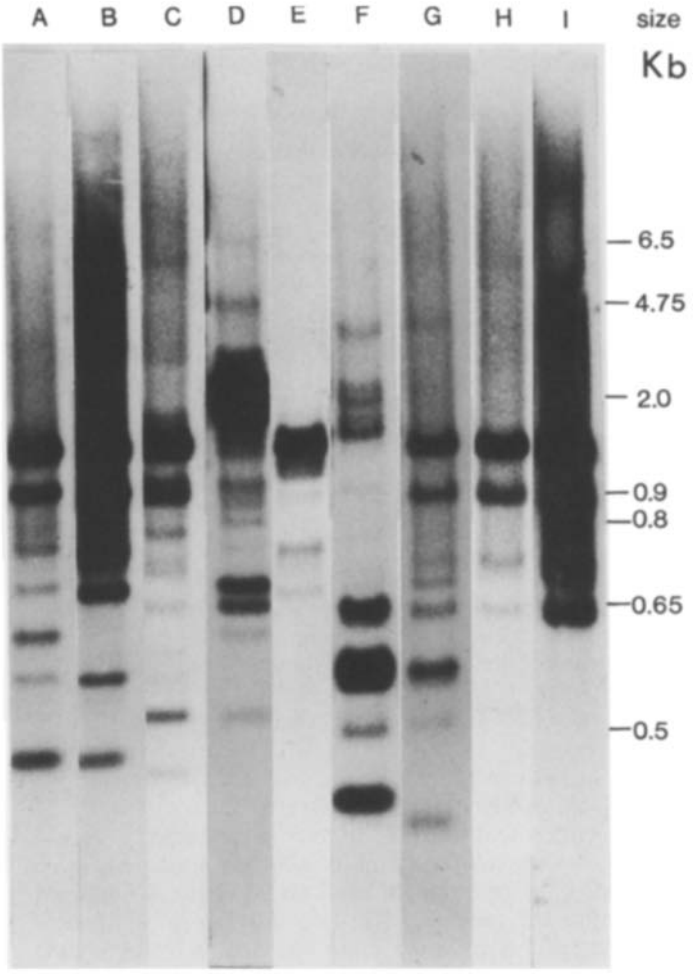


Fig. 5. Hybridization of fragments A–I from pEco5 (see Fig. 2) to *Hin*dIII digests of wheat DNA

	5' ENDS	3' ENDS
Wis-2 (wheat)	TGTTGGAAA----	-TT--CC-ACA
Wire-1 (wheat)	TGTTGG----C--	-TT--CC-ACA
BS-1 (maize)	TGTT-G-AA-C--	--T--C--ACA
Cin1 (maize)	TGTTGG----C--	-----CC-ACA
Ty1 (yeast)	TGTTGGAA-A-A-	---------CA
IAP (mouse)	TGTTGG-A--C--	-T------ACA
Copia (Drosophila)	TGTTGGAA-A-A-	------C-ACA
B104 (Drosophila)	TGTT---A-A---	--------ACA
MMSV (mouse)	TG---GA----A-	-----C---CA

Fig. 6. Similarities between the terminal nucleotides of certain transposable elements from various organisms. The nucleotides included are those held in common between three or more of the elements; – represents a non-conserved nucleotide. Wis-2 is the element described in this paper. Wire-1 is another wheat retroposon-like element recently characterised (N. Harris and R. Flavell, in preparation). Bs-1 is from maize (see Johns et al. 1985). Cin-1 is a dispersed repeat from maize (Shepherd et al. 1984). Ty1 is a transposable element in *Saccharomyces cerevisiae* (Farabaugh and Fink 1980). IAP is the intracisternal A-particle gene of mouse which belongs to a group of endogenous proretroviral-like elements with long terminal repeats (Kuff et al. 1983). Copia is a transposable element of *Drosophila* comprising up to 5% of the *Drosophila* genome (Levis et al. 1980). B104 is a dispersed repeat of *Drosophila* (Scherer et al. 1982). MMSV is the Moloney virus of mouse which has long terminal repeats (Dhar et al. 1980)

hybridization patterns to the total wheat DNA are highly complex. Each probe reveals many hybridizing fragments in wheat DNA, and at least some of the sequences contained on these fragments are present in many more than 100 copies per haploid wheat genome. Interestingly, fragments A, B and C appear to hybridise to the same major chromosomal DNA fragments as fragments G, H and I. Also all these fragments hybridise to both of the *Bam*HI fragments of pEco5 (see Fig. 2). Therefore, related sequences are found at both ends of the DNA inserted into the *Glu-Aly* gene, a conclusion we have confirmed by cross-hybridisation of isolated pEco5 fragments (data not shown). The duplicated sequences must extend over more than 500 bp. From the 700 bp of sequence determined at the 5′ and 3′ ends, which cannot be matched, it appears that the duplicated sequences are not inverted with respect to one another, but in the same orientation.

Discussion

In a previous study DNA restriction fragment length polymorphisms at the *Glu-1* loci of five different wheat varieties were examined and alleles previously distinguishable on the basis of the HMW subunit they encode were often found to be distinguishable on the basis of their associated restriction enzyme sites (Harberd et al. 1986). In this paper a detailed comparison of sections of two allelic *Glu-Aly* sublocus genes from the varieties Chinese Spring and Cheyenne is presented. It is clear that the major difference between these two genes is the presence of an 8 kb segment of unrelated DNA inserted into the coding sequence of the Chinese Spring gene but absent from its counterpart in Cheyenne. This observation cannot be attributed to artefacts arising during cloning of the *Glu-Aly* allele from Chinese Spring, since digests of the cloned fragment produce fragments of sizes predicted from hybridization of the HMW glutenin cDNA to digests of DNA from nullisomic-tetrasomic wheat lines (data not shown; Thompson et al. 1985).

The presence or absence of the insert seems not to be the only determinant of whether *Glu-Aly* is expressed in these two varieties. There is no detectable HMW glutenin transcript derived from chromosome 1A of Chinese Spring (Thompson et al. 1983) but it seems that the *Glu-Aly* sublocus gene of Cheyenne, which lacks the inserted DNA segment, is also not expressed (Forde et al. 1985). The absence of a 1Ay type HMW subunit throughout the hexaploid and tetraploid wheats suggests that the *Glu-Aly* gene is usually silent in cultivated lines.

Flanking each end of the DNA segment inserted into *Glu-Aly* of Chinese Spring is a small (5 or 6 bp) duplication). Such duplications of a short sequence at the site of insertion are a common feature of DNA insertion events (Saedler and Nevers 1985). The insertion interrupts one of the nine amino acid encoding repeat units from the central repetitive section of the *Glu-Aly* gene (Thompson et al. 1985; Forde et al. 1985). The comparable repeat unit from *Glu-Aly* of Cheyenne (nucleotides 1718 to 1744 in the sequence of Forde et al. 1985) is of normal length and does

not contain any adjacent duplications of a few base pairs which might be expected to result from excision of a transposable element (Schwarz-Sommer et al. 1985). It is likely therefore that the Cheyenne *Glu-Aly* allele represents a progenitor allele into which a DNA segment has been inserted resulting in an allele of the type represented by the Chinese Spring *Glu-Aly* allele. The inserted DNA segment has been called Wis-2 (Wheat insertion sequence-2). Wis-1 is an insertion sequence found upstream of an α-amylase gene (R.A. Martienssen and D.C. Baulcombe, in preparation).

The Wis-2 DNA inserted into the Chinese Spring *Glu-Aly* gene has the terminal bases 5′ TGTTGG and CCTACA 3′ which are inverted duplicates of one another, with the exception of 1 bp. Terminal inverted repeats are a common characteristic of many transposable elements (Freeling 1984; Finnegan 1985). However, these particular short terminal inverted repeats are closely related to those found in numerous other transposable elements which are propagated by reverse transcription of an RNA intermediate (Finnegan 1985). The terminal homologies are illustrated in Fig. 6. The terminal sequences of Wis-2 are identical or very similar to those of elements in yeast (Ty-1) (Farabaugh and Fink 1980), *Drosophila* (copia) (Levis et al. 1980), and mammalian retroviruses (Kuff et al. 1983; Rogers 1985) all of which have 5′ TG-CA 3′ at their ends. All these elements are also characterised by having direct repeats, several hundred base pairs long, at their termini each of which carries the TGTTGG-CCTACA type of inverted repeat (Finnegan 1985). The results in Fig. 5 and the sequence in Fig. 3 imply that Wis-2 in the *Glu-Aly* allele of Chinese Spring also has long direct repeats at its termini. These features therefore suggest that this wheat insertion sequence is of the 'retroposon type' having been inserted into the *Glu-Aly* gene following reverse transcription of an RNA transcript from another copy elsewhere in the genome. Another example of a wheat element with long terminal repeats each of which has the terminal inverted repeats TGTTGG-CCAACA has recently been characterised by N. Harris and R. Flavell (see Fig. 6; Flavell 1986 and unpublished results) and other examples have been described in maize (see Fig. 6; Johns et al. 1985; Shepherd et al. 1984; Flavell 1986). It could be therefore that many of the dispersed repeats in plant genomes are of this type.

There are many other sequences in the wheat genome related to those of Wis-2 inserted into the *Glu-Aly* allele of Chinese Spring (Fig. 5). However, most of them appear to be organised differently from those in Wis-2 although it is possible that much of the variation between Wis-2 and the other copies indicated by Fig. 5 is due to mutations in restriction enzyme recognition sites rather than the existence of other copies organised in unrelated structures.

The coding regions of HMW glutenin gene sequences contain an extensive central section consisting of tandemly repetitive sequences. The length of this region appears to vary considerably between HMW glutenin genes (Thompson et al. 1985; Harberd et al. 1986). The sequence of the repetitive region of *Glu-Aly* from Chinese Spring differs from that of Cheyenne by the presence of one extra repeat unit. Two six amino acid encoding repeat units are found between nucleotides 1009–1044 (Fig. 3b) in *Glu-Aly* of Chinese Spring, compared with only one of these (nucleotides 2264–2281, Forde et al. 1985) at the same position in *Glu-Aly* of Cheyenne. The two units in Chinese Spring are exact copies of one another. Of course we cannot be certain that this recent duplication did not arise during molecular cloning of the *Glu-Aly* gene fragment rather than at some point in the evolutionary divergence of the two varieties. Nevertheless this observation provides evidence that this region of repetitive DNA has a propensity for rapid change via deletion and/or duplication of repeat units. Mechanisms by which this change might come about have been proposed (Thompson et al. 1983).

The molecular characterization of this locus aids our understanding of the evolution of an agronomically important gene family in bread wheat. The number of active HMW glutenin genes varies between three and five amongst hexaploid wheat varieties. It will be of interest to compare other inactive *Glu-1* alleles, or indeed, inactive members of other gene families in hexaploid wheat, to see if common mechanisms of inactivation are involved.

Acknowledgements. N.P.H. was supported by an M.R.C. Training Fellowship in Recombinant DNA Technology and a grant from the Perry Foundation. This work was supported in part by a research contract with the E.E.C.

References

Deininger PL (1983) Rapid subcloning of sonicated DNA: Application to shotgun DNA sequence analysis. Anal Biochem 129:216–223

Dhar R, McClements WL, Enquist LW, Van de Woude GF (1980) Nucleotide sequence of integrated Moloney sarcoma provirus long terminal repeats and their host viral junctions. Proc Natl Acad Sci USA 77:2939–2941

Farabaugh PJ, Fink GR (1980) Insertion of eukaryotic transposable element Ty1 creates a 5 base pair duplication. Nature 286:352–356

Finnegan D (1985) Transposable elements in eukaryotes. Int Rev Cytol 93:281–326

Flavell RB (1986) Repetitive DNA and chromosome evolution in plants. Philos Trans R Soc Lond [Biol] 312:227–242

Forde J, Malipica J-M, Halford NG, Shewry PR, Anderson OD, Green 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

Freeling M (1984) Plant transposable elements. Annu Rev Plant Physiol 35:277–298

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

Johns MA, Mottinger J, Freeling M (1985) A low copy number, *copia*-like transposon in maize. EMBO J 4:1093–1102

Kasarda DD, Bernardin JE, Nimmo CC (1976) Wheat proteins. In: Pomeranz Y (ed) Advances in cereal science and technology, vol 1. American Association for Cereal Chemistry, St. Paul, Minnesota, pp 158–236

Kuff EL, Fenestra A, Lueders K, Smith L, Hawley R, Hozumi N, Shulman M (1983) The intracisternal A-particle gene belongs to a group of endogenous proretroviral-like elements present in about 100 copies per haploid genome of mouse. Proc Natl Acad Sci USA 80:1992–1996

Levis R, Dunsmuir P, Rubin GM (1980) Terminal repeats of the *Drosophila* transposable element *copia:* nucleotide sequence and genomic organisation. Cell 21:581–588

Payne PI, Law CN, Mudd EE (1980) Control by homoeologous group 1 chromosomes of the high-molecular-weight subunits of glutenin, a major protein of wheat endosperm. Theor Appl Genet 58:113–120

Payne PI, Holt LM, Law CN (1981) Structural and genetical studies on the high-molecular-weight subunits of wheat gluten. 1. Allelic variation in subunits amongst varieties of wheat (*Triticum aestivum*). Theor Appl Genet 60:229–236

Payne PI, Holt LM, Worland AJ, Law CN (1982) Structural and genetical studies on the high-molecular-weight subunits of wheat glutenin. 3. Telocentric mapping of the subunit genes on the long arms of the homoeologous group 1 chromosomes. Theor Appl Genet 63:129–138

Payne PI, Holt LM, Thompson RD, Bartels D, Harberd NP, Harris PA, Law CN (1983) The high-molecular-weight subunit of glutenin: classical genetics, molecular genetics and the relationship to bread-making quality. Proceedings 6th International Wheat Genetics Symposium, Kyoto, Japan, pp 827–834

Payne PI, Holt LM, Jackson EA, Law CN (1984) Wheat storage proteins: their genetics and their potential for manipulation by plant breeding. Philos Trans R Soc Lond [Biol] 304:359–371

Poncz M, Solowiejczyk D, Ballantine M, Schwartz F, Surrey S (1982) "Non-random" DNA sequence analysis in bacteriophage M13 by the dideoxy chain termination method. Proc Natl Acad Sci USA 79:4298–4302

Rogers JH (1985) The origin and evolution of retroposons. Int Rev Cytol 93:187–279

Saedler H, Nevers P (1985) Transposition in plants: a molecular model. EMBO J 4:585–590

Sanger F, Coulson AR, Barrell BG, Smith AGH, Roe BA (1980) Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J Mol Biol 143:161–178

Scherer G, Tschudi C, Perera J, Delius H (1982) B104, a new dispersed repeated gene family in *Drosophila melanogaster* and its analogies with retroviruses. J Mol Biol 157:435–451

Schwarz-Sommer Zs, Gierl A, Cuypers H, Peterson PA, Saedler H (1985) Plant transposable elements generate the DNA sequence diversity needed in evolution. EMBO J 4:591–597

Shepherd NS, Schwarz-Sommer Z, vel Spalve JB, Gupta M, Wienand U, Saedler H (1984) Similarity of the Cin 1 repetitive family of *Zea mays* to eukaryotic transposable elements. Nature, 307:185–187

Staden R (1982a) Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. Nucleic Acids Res 10:4731–4751

Staden R (1982b) An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences. Nucleic Acids Res 10:2951–2961

Thompson RD, Bartels D, Harberd NP, Flavell RB (1983) Characterization 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 (1985) Nucleotide sequence of a gene from chromosome 1D of wheat encoding a HMW glutenin subunit. Nucleic Acids Res 13:6887–6846

Communicated by H. Saedler

Received December 19, 1986 / March 29, 1987