Theor Appl Genet (2001) 103:37–44 

**C.C. Hsia · O.D. Anderson**

# Isolation and characterization of wheat ω-gliadin genes

Received: 24 August 2000 / Accepted: 15 December 2000

**Abstract** The DNA sequences of two full-length wheat ω-gliadin prolamin genes (ωF20b and ωG3) containing significant 5′ and 3′ flanking DNA sequences are reported. The ωF20b DNA sequence contains an open reading frame encoding a 30,460-Dalton protein, whereas the ωG3 sequence would encode a putative 39,210-Dalton protein except for a stop codon at amino-acid residue position 165. These two ω-gliadin genes are closely related and are of the ARQ-/ARE-variant type as categorized by the derived N-terminal amino-acid sequences and amino-acid compositions. The ω-gliadins were believed be related to the ω-secalins of rye and the C-hordeins of barley, and analyses of these complete ω-gliadin sequences confirm this close relationship. Although the ω-type sequences from all three species are closely related, in this analysis the rye and barley ω-type sequences are the most similar in a pairwise comparison. A comparison of ω-gliadin flanking sequences with respect to that of their orthologs and with respect to wheat gliadin genes suggests the conservation of flanking DNA necessary for gene function. Sequence data for members of all major wheat prolamin families are now available.

**Keywords** Omega-gliadin · Gliadins · Wheat · Sulfur-poor prolamins · Storage proteins

## Introduction

In wheat, the prolamin seed-storage proteins include the high-molecular-weight (HMW) glutenins and the large gliadin-family group. The latter is composed mainly of the α-, γ-, and ω-gliadins and the low-molecular-weight (LMW) glutenins (additional minor subfamilies will be reported elsewhere). The majority of γ-gliadin and ω-gliadin genes are encoded by the *Gli-1* loci on the short arm of chromosome 1 and tightly linked to the LMW-glutenin genes encoded by the *Glu-3* loci (Tatham and Shewry 1995). A characteristic of the prolamins is a repeat domain rich in glutamine and proline residues. Of the four main wheat prolamin families, the ω-gliadins differ from the others in that they generally have no cysteines and at most one methionine, and are therefore termed ”sulfur-poor” (Tatham and Shewry 1995). The absence of cysteine residues also means that the ω-gliadins are not participants in intra- and inter-molecular disulfide bond formation as are the other wheat prolamins. No ω-gliadin sequences have yet been reported for wheat. Sulfur-poor orthologs of wheat ω-gliadins include the ω-secalins of rye and the C-hordeins of barley. Full-length sequences of ω-secalin (Hull et al. 1991; Clarke et al. 1996) and C-hordein genes (Entwistle 1988; Entwistle et al. 1991; Sainova et al. 1993) both show a coding region that is almost entirely composed of repetitive motifs.

We report the full-length sequences of two closely related ω-gliadin genomic clones of 3,789 bp and 3,925 bp, respectively, that include the complete coding region and significant flanking DNA sequences. DNA sequence data from each of the major families of wheat seed-storage proteins are now available, and a more complete comparative sequence analysis of wheat storage-protein genes is possible.

Communicated by J. Dvorak

C.C. Hsia · O.D. Anderson (✉)
U.S. Department of Agriculture, Agricultural Research Service, Western Regional Research Center, 800 Buchanan Street, Albany, CA 94710, USA
e-mail: oandersn@pw.usda.gov
Tel.: +1-510-559-5773, Fax: +1-510-559-5777

## Material and methods

The ω-gliadin clones were isolated from wheat genomic lambda bacteriophage (λ) libraries as described in Anderson et al. (1997). The initial screen of the libraries was with a γ-gliadin probe (clone γ13; Anderson et al. 2001). *Eco*RI fragments of the DNA clones were subcloned into plasmid RVIIΔ7-Z2 (a vector modified from RVIIΔ7; Lynn et al. 1983) and transformed into the *Escherichia coli* Sure cell line (Stratagene, Inc.). Sequencing of both strands was completed by primer walking using ABI PRISM DyeTerminator and BigDye Terminator Cycle Sequencing Ready Reaction

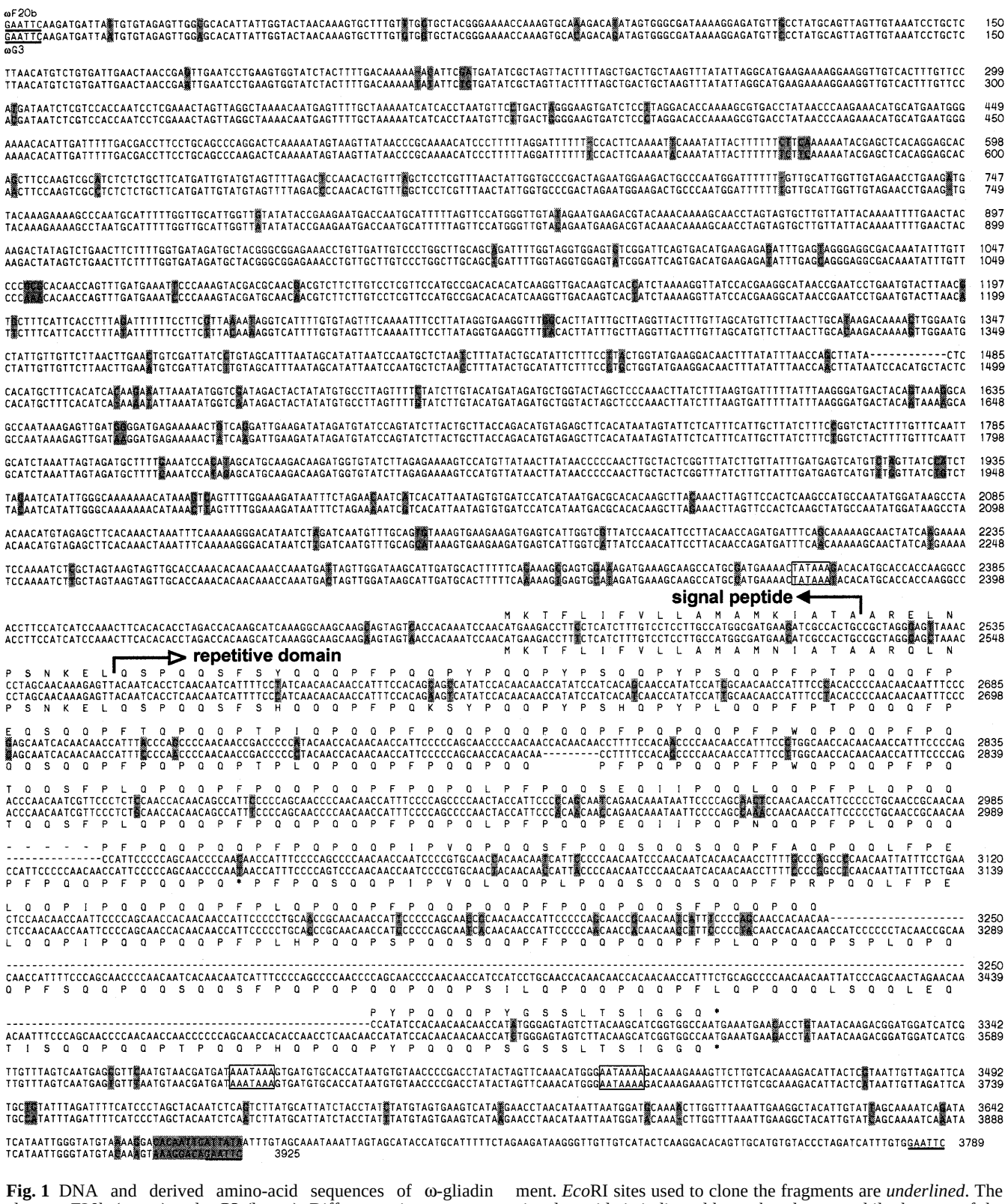



**Fig. 1** DNA and derived amino-acid sequences of ω-gliadin clones ωF20b (upper) and ωG3 (lower). Differences in sequences are shaded in *gray*. A putative TATA box and polyadenylation signals are *boxed*. Stop codons are indicated by *asterisks*. *Dashes* indicate gaps introduced into the sequence of ωF20b to allow alignment. *Eco*RI sites used to clone the fragments are *underlined*. The signal peptide is indicated by a *closed arrow* while the start of the repeat domain is indicated by an *open arrow*. Genbank accession numbers: ωF20b, AF280605; ωG3, AF280606

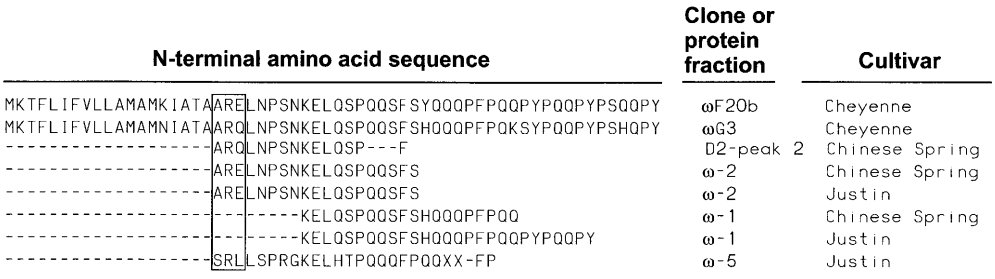



**Fig. 2** N-terminal sequences of ω-gliadins. Alignment of the first 60 residues of derived polypeptides of clones ωF20b and ωG3 are compared with sequences from Masci et al. (1999; D2-peak 2) and of Kasarda et al. (1983; final five sequences). The *box* identifies the initial amino acids of some mature polypeptides

Kits for sequence analysis on an ABI PRISM 310 capillary DNA sequencer. Sequences were assembled using AutoAssembler v.1.4.0 (PE Biosystems) and analyzed using the Lasergene software (DNAstar, Inc.).

## Results

### Isolation and sequencing of ω-gliadin clones

A known γ-gliadin clone (γ13; Anderson et al. 2001) was used to isolate additional γ-gliadin clones from cv Cheyenne λ libraries. Most of the new clones hybridized strongly to the γ-gliadin probe. However, two clones hybridized less intensely and were suspected to be ω-gliadins (ωF20b and ωG3) since Sabelli and Shewry (1991) suggested that cross-hybridization between γ-gliadins and ω-gliadins occurs. Additionally, data from N-terminal sequencing and amino-acid analysis of ω-gliadin proteins (Kasarda et al. 1983), as well as published DNA sequences of rye orthologs (ω-secalins) and barley orthologs (C-hordeins), indicate that the ω-gliadin repetitive domain is similar to that of γ-gliadins.

No ω-gliadin gene sequences have been previously reported although many laboratories, including ours, have attempted such isolations. Attempts in our laboratory to clone the *Eco*RI fragments carrying the putative ω-gliadin genes from the original λ clones into various traditional vectors and hosts resulted in random deletions of the inserts. Efforts to stabilize the inserts for cloning succeeded when the plasmid RVIIΔ7-Z2 and the *E. coli* Sure cell line were used. RVIIΔ7-Z2 was known to stabilize inserts that are difficult to clone (Anderson, unpublished observation) and the *E. coli* Sure cell line is engineered to be DNA-repair deficient and to reduce homologous recombination for increased insert stability.

A total of 3,789 bp and 3,925 bp were sequenced for the ωF20b and ωG3 clones, respectively (Fig. 1). Approximately 2.5 kbp of each clone was 5′ sequence and 0.4–0.5 kbp was 3′ flanking sequenced. The coding region totalled 840 bp for ωF20b and 1,073 bp for ωG3. ωF20b has an open reading frame throughout the entire coding region. Clone ωG3 has an in-frame stop codon (position 3,026 in Fig. 1) and is assumed to be a pseudogene.

The presumptive ω-gliadin sequences were analyzed using BLASTN (Altschul et al. 1997) and the search yielded the closest homologies with ω-secalins and C-hordeins, confirming the identity of the new sequences

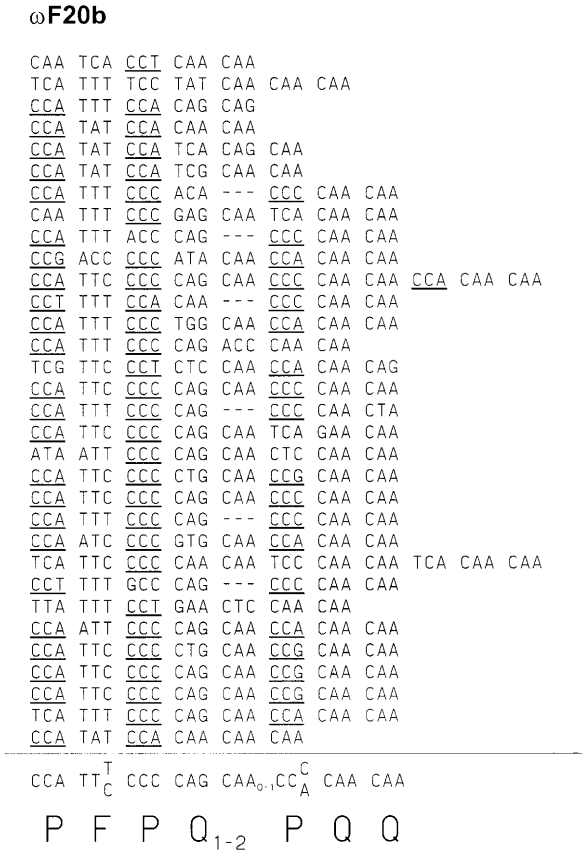



**Fig. 3** Alignment of ωF20b repetitive motifs by DNA sequence. *Dashes* indicate gaps introduced to emphasize motifs. Proline codons are *underlined*. A proposed consensus DNA motif is shown below along with its derived amino-acid sequence

as ω-gliadins. Significantly lower match scores were obtained with γ-gliadins, oat avenins and α-gliadins. BLASTN and BLASTX homology searches of the more distal 5′ and 3′ flanking DNAs yielded no significant matches to known sequences.

The general protein structure of an ω-gliadin is simple: a 19-residue putative signal peptide followed by a 10–11-residue non-repetitive N-terminus, a repetitive region encompassing 90–96% of the protein and a 10–11 residue C-terminus. As expected from amino-acid analyses of ω-gliadins (Kasarda et al. 1983), there are no cysteine or methionine residues encoded by either of the clones sequenced. In addition, Tatham and Shewry (1995) noted that ω-gliadins have few charged amino ac-


**Fig. 4** Repetitive domain of all reported S-poor prolamins. Repeat units are arrayed vertically to show repeat motifs. Stop codons are indicated by *periods*. The repetitive domains of the ω-secalins are consolidated because pSec1B and pSec2B (Hull et al. 1991) are nearly identical to that of pSec1 (Clarke et al. 1996). Differences in amino-acid residues from the pSec1 sequence are indicated by *slashes* separating pSec1 from the other two: the pSec1 residue is to the left of the slash; pSec1B residues are to the right of the slash; pSec2B residues are to the right of the slash and *shaded*

ids, consistent with the finding that ωF20b encodes just seven charged residues out of 261 total residues. Gene ωG3 would encode 13 charged residues out of 338 except for the internal stop codon. Whereas the sulfur-rich gliadins have polyglutamine homopolymers or glutamine-rich regions separate from the repeat domains, all glutamine residues in the ω-gliadins are found within the repeat units that make up the repetitive domain. The repetitive domains of the two encoded amino acid sequences are quite similar: ωF20b has an additional PQQ in one repeat unit (position 2,770–2,778) that is otherwise identical to the repeat unit found in ωG3 at the same position. Sequence ωG3 has an extra repeat unit, PFPQQ, at positions 2,990–3,004 and an extra eight repeat units towards the end of the C-terminus (starting at position 3,250). In addition to differences in the numbers and positions of the repeat units, ωF20b and ωG3 differ by 23 residues in their coding region: 21 changes result from single base alterations and two changes are due to alterations in two bases.

The amino-acid composition of the translated coding regions of these new ω-gliadin genes is also consistent with data obtained by Kasarda et al. (1983) for ω-gliadin proteins: as predicted for ω-gliadins, proline, glutamine and phenylalanine residues comprise approximately 80% of the total amino acids compared to 50–60% for the other wheat gliadins. Kasarda et al. (1983) used the N-terminal protein sequences to classify the ω-gliadins based on the first three amino acids of the mature protein. By this criterion, these new ω-gliadin sequences belong to the ARQ-/ARE-type protein which has a 4:3:1 ratio of Q:P:F for total amino acids. The ARQ-/ARE-type has a similar ratio to the S-poor prolamins of rye and barley, and the genes may be encoded on the 1 A and 1D chromosomes (Tatham and Shewry 1995; Masci et al. 1999). The SRL-type has a different Q:P:F ratio (5:2:1) and is characteristic of the ω-gliadins encoded by the 1B chromosome (Tatham and Shewry 1995; Dupont et al. 2000). N-terminal amino-acid sequences derived from ω-gliadin genes ωF20b and ωG3 are identical (Fig. 2), or nearly identical, to N-terminal peptide sequences from the 1D-encoded ω-gliadins (Kasarda et al. 1983; Masci et al. 1999; Dupont et al. 2000), and therefore supports the origin of these two genes from the 1D chromosome.

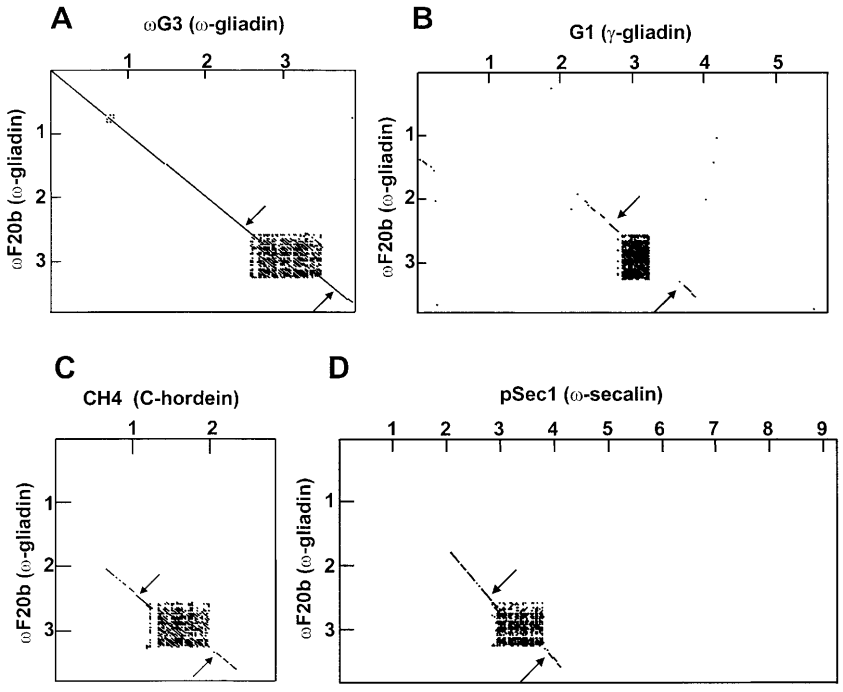



**Fig. 5A–D** Dot matrix homology plots of the ωF20b DNA sequence (ω-gliadin) compared to selected prolamin gene sequences. All matrices were generated using the Megalign module in the Lasergene (DNAstar, Inc.) software. Parameters were set to find an 80% match over a 20-bp window except for the plot against G1 (γ-gliadin) in which a 70% match over a 30-bp window was used to better determine the limits of conservation between ω-gliadin and γ-gliadin sequences. *Arrows above the diagonals* indicate the positions of start codons while the *arrows below the diagonals* indicate stop codons. **A** ωF20b vs ωG3 (ω-gliadin). **B** ωF20b vs G1 (γ-gliadin). **C** ωF20b vs CH4 (C-hordein). **D** ωF20b vs pSec1 (ω-secalin)

The ωG3 gene is assumed to be a pseudogene by the definition of containing an in-frame stop codon. Since this is the only obvious defect in either the coding or flanking DNA sequences, it is not excluded that ωG3 may express mRNA that only produces an N-terminal portion of an ω-gliadin protein. Pseudogenes are common in the cereal prolamins, although their distribution seems to vary with each prolamin gene family. For example, the α-gliadin family is estimated to be composed of 50% pseudogenes (Anderson and Greene 1997) which is hypothesized to be caused by the high frequency of glutamine codons (CAA and CAG). A nonsense mutation can occur from a C→T transition, the predominant type of single-base change. Though we only have two examples of ω-gliadin genes, both have a high percentage of glutamine codons (average 40%) in the repetitive domain, which comprise almost the entire coding region. It may be anticipated that the ω-gliadin family also has a large percentage of pseudogenes; but more sequences are needed to confirm this speculation.

## Repeat structure

The DNA sequence of the repeat domain of ωF20b is arrayed vertically in Fig. 3 to emphasize the motif structure. A DNA consensus, CCA $TT^{T}/_{C}$ CCC CAG $CAA_{0-1}$ $CC^{C}/_{A}$ CAA CAA, based on the most frequently occurring nucleotide for each position in the codon, is suggested below the array. Variations from the DNA consensus are usually due to single base-pair changes. The derived amino-acid motif, $PFPQ_{1-2}PQQ$, is shown at the bottom of the figure. This peptide motif is similar to that proposed for the ω-secalins and C-hordeins, PQQPFPQQ (Tatham and Shewry 1995), and for the γ-gliadins, $PFPQ_{1-2}(PQQ)_{1-2}$ (Anderson et al. 2001). It is not known whether this is representative of all ω-gliadins or if ωF20b and ωG3, whose isolation was a result of hybridization with a γ-gliadin probe, represent members of the ω-gliadin family whose repeats are the most closely related to γ-gliadin repeats.

Arrays of the peptide motifs of the S-poor prolamins are shown in Fig. 4. Our proposed alignment of peptide motifs for the ω-secalins is based on analysis of the DNA structure, and thus is slightly different from the 32-residue repeat unit composed of four octapeptides, PQQIIPQQ PQQPFPLQ $PQQPF^{P}/_{S}QQ$ PQQPFPQQ, proposed by Hull et al. (1991) for rye secalins. The length of the repeat region of the ω-gliadins is comparable to those of ω-secalins and C-hordeins, and most of these repeat units in the ω-gliadins are variants of the consensus octapeptide motif. Those which differ from this motif typically have an extra PQQ/SQQ or begin with a repeat peptide containing 1–2 isoleucine residues. A repeat unit starting with double isoleucine is also found throughout the ω-secalins and C-hordeins, indicating an origin before speciation. This repeat unit may have initially been a single-copy repeat that underwent duplication events and spread through the repeat domain. The same mechanism could also have produced multiple copies of the repeat units containing a tryptophan codon, PFPWQPQQ, in the C-hordein, CH4, which is found in only one copy in the other C-hordeins and the ω-gliadins.

## Dot plots

A homology matrix of the two ω-gliadin clones shows that the sequences are closely related throughout their length (Fig. 5A) except for the final 14 bp of ωG3. Pair-

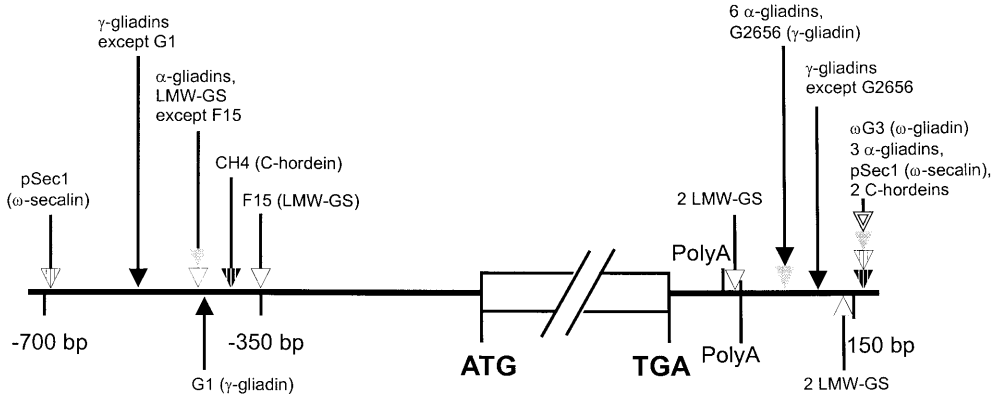



**Fig. 6** Divergence points in flanking DNA sequences of members of the gliadin superfamily and S-poor prolamins. All reported gliadin genes and orthologs of ω-gliadins with significant flanking DNA sequences were compared pairwise to ω-gliadin ωF20b to determine the points of divergence. The *rectangular box* represents the ωF20b coding region and the *lines* represent the flanking sequences. *Arrows* indicate the point of divergence of a sequence from ωF20b, or divergence within a 50-bp region when more than one sequence diverges from ωF20b at approximately the same position. The position of the upstream divergence sites was calculated from the initiation codon (indicated by ATG). The position of downstream divergence sites was calculated from the second polyadenylation (*Poly A*) signal. Points of divergence at the 5′ end are (1) approximately –690 bp: ω-secalin pSec1 (Clarke et al. 1996); (2) approximately –550 bp: γ-gliadins γ13, G6, G2656 (Anderson et al. 2000) and pW1020 (Scheets and Hedgcoth 1988); (3) approximately –450 bp: α-gliadins CNN5, CNN10, CNNE18 C (Anderson et al. 1997), Yam2 (Anderson et al. 1984), OKURARTU (Reeves and Okita 1987), W8142 (Sumner-Smith et al. 1983) and LMW-GS F23 (Cassidy et al. 1998) and LMWG-1D1 (Colot et al. 1989); (4) approximately –440 bp: γ-gliadin G1 (Anderson et al. 2000); (5) approximately –430 bp: C-hordein CH4 (Sainova et al. 1993); (6) approximately –350 bp: LMW-GS F15 (Cassidy et al. 1998). Sequences diverging at the 3′ end are (1) about 15 bp before the polyA site: LMW-GS F14 and F24 (Cassidy et al. 1998); (2) about 15 bp after the polyA site: α-gliadins CNN5, CNN10, CNNE18 C, CNN113 (Anderson et al. 1997), Yam2 and W8233 (Rafalski et al. 1984); (3) about 85 bp after the polyA site: γ-gliadins L311 A and L311B (Rafalski 1986) and pW1020, G1 and G6; (4) about 140 bp after the polyA site: LMW-GS F23 and LMWG-1D1 (Cassidy et al. 1998); (5) about 170 bp after the polyA site: ω-gliadin ωG3 (this paper); α-gliadins CNN18, W8142 and OKURARTU; C-hordeins CH4 and λhor1–14 (Entwistle 1988)

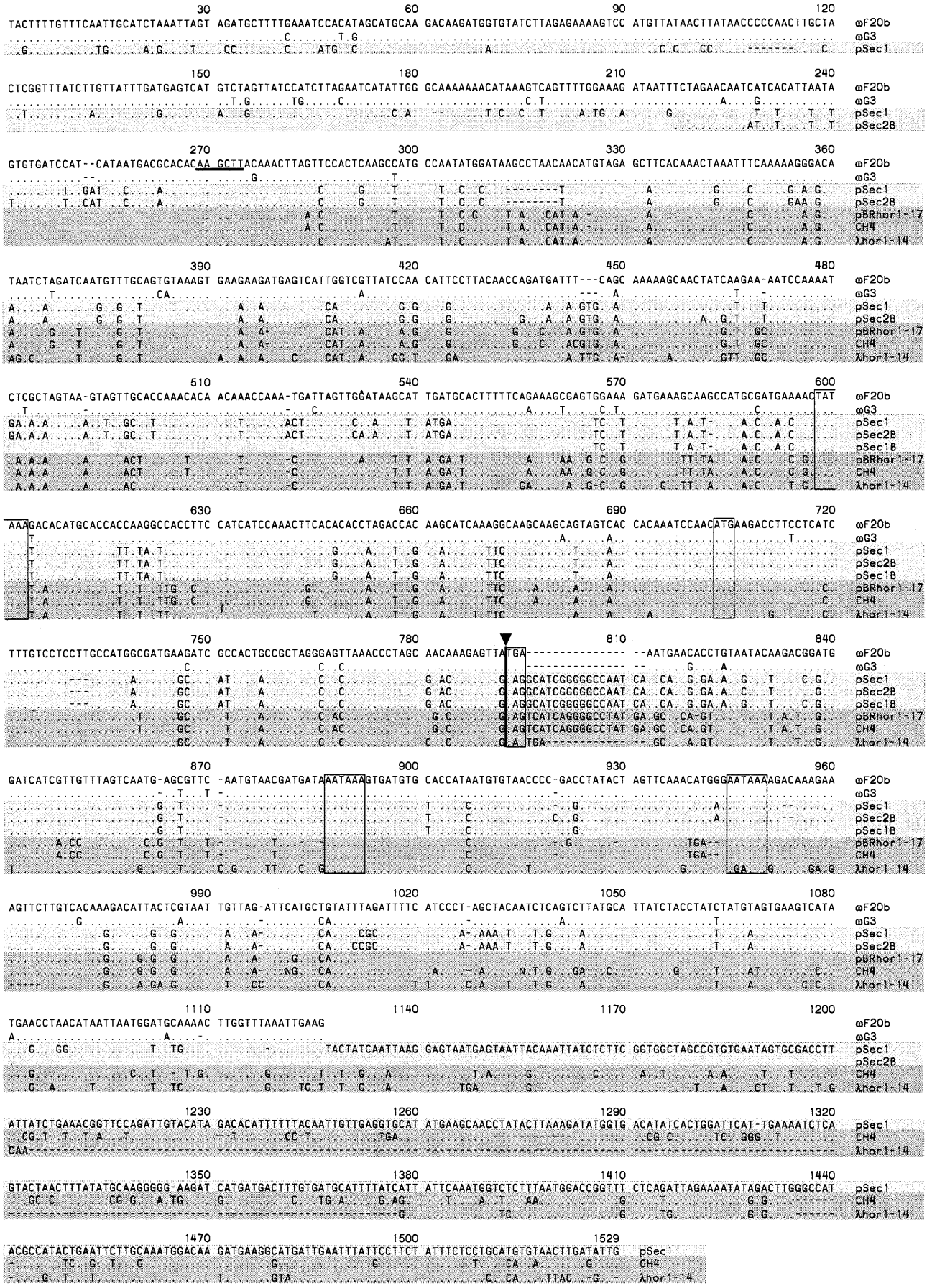
All gliadin-type prolamins in the present analysis (gliadins, LMW-glutenins, ω-secalins, and C-hordeins) share similarity in flanking DNA sequences from at least –350 bp from the ATG codon and to the polyadenylation site (Fig. 6). These conserved sequences may indicate the limits of the regulatory sequences necessary for a functional prolamin gene as we previously speculated (Anderson and Greene 1997).

## Comparison of ω-orthologous noncoding sequences

Points of sequence divergence among ω-type prolamins were also analyzed in a Clustal analysis (Higgins and Sharp 1989) of S-poor prolamin sequences with the repetitive domains removed (Fig. 7). All sequences analyzed have a conserved *Hin*dIII site at approximately –530 bp (position 269 in Fig. 7). One C-hordein clone, CH4 (Sainova et al. 1993), has sequence data beyond this point but also diverges from the other S-poor prolamin sequences at the *Hin*dIII site (diverged sequence not shown in Fig. 7). The ω-gliadins (ωF20b and ωG3) and the ω-secalin pSec1 (Clarke et al. 1996), diverge from each other 270-bp upstream of the conserved *Hin*dIII site or about –800 bp from the start codon. On the 3′ flank, all three groups of S-poor prolamins show conservation to 176 bp (position 1,126) after the polyadenylation site

wise comparisons are also shown between ωF20b and the other ω-type prolamines plus γ-gliadin clone G1 (Figs. 5B–D). In addition, enough flanking DNA sequences were available for divergence points to be determined between ωF20b and at least one member of each of the other gliadin families (Fig. 6). For 5′ flanking sequences, most members of the same gliadin family diverged from ωF20b within 50 bp of each other. The exceptions were the γ-gliadins, for which all but the G1 γ-gliadin diverged at about –550 bp, and the LMW-GS for which all but the F15 LMW-GS diverged at about –450 bp (Cassidy et al. 1998; Fig. 6). For the 3′ flanking sequences, members of each family did not diverge within the same 50 bp, but subgroups within the family seemed to be clustered with respect to certain divergence points. All sequences diverged by 176 bp 3′ of the polyadenylation site.

**Fig. 7** Alignment of DNA sequences of S-poor prolamins from wheat, rye and barley. All flanking sequence and coding sequence up to but not including the repetitive domain are compared. Wheat ω-gliadin (ωF20b and ωG3) sequences are *unshaded*, rye ω-secalin sequences (pSec1, pSec2B, and pSec1B) are shaded in *light gray*, and barley C-hordein sequences (pBRhor1–17, CH4, and λhor1–14) are shaded in *dark gray*. The *vertical bar* after position 792 indicates the removed repetitive and C-terminal domains. The 5′ flanking sequences of ωF20b, ωG3 and CH4 and the 3′ flanking sequences of ωF20b, ωG3, pSec1, CH4 and λhor1–14 beyond the point of sequence divergence are not shown. Putative TATA box and polyadenylation sites, and the initiation and stop codons are *boxed*. *Dashes* indicate gaps introduced into the sequences to optimize alignments. Dots in the sequence represent bases identical to ωF20b. The 5′ *Hin*dIII site from where the barley C-hordein sequences diverge from the other S-poor prolamins is *underlined* (position 269). The ωF20b sequence is used as the sequence for comparison up to the point where it no longer shares sequence similarity to its orthologs (position 1,126). The pSec1 sequence is used as the sequence for comparison after position 1,126 ▶


30 60 90 120
TACTTTTGTTTCAATTGCATCTAAATTAGT AGATGCTTTTGAAATCCACATAGCATGCAA GACAAGATGGTGTATCTTAGAGAAAAGTCC ATGTTATAACTTATAACCCCCAACTTGCTA ωF20b
ωG3
pSec1
150 180 210 240
CTCGGTTTATCTTGTTATTTGATGAGTCAT GTCTAGTTATCCATCTTAGAATCATATTGG GCAAAAAAACATAAAGTCAGTTTTGGAAAG ATAATTTCTAGAACAATCATCACATTAATA ωF20b
ωG3
pSec1
pSec2B
270 300 330 360
GTGTGATCCAT--CATAATGACGCACACAA GCTTACAAACTTAGTTCCACTCAAGCCATG CCAATATGGATAAGCCTAACAACATGTAGA GCTTCACAAACTAAATTTCAAAAAGGGACA ωF20b
ωG3
pSec1
pSec2B
pBRhor1-17
CH4
λhor1-14
390 420 450 480
TAATCTAGATCAATGTTTGCAGTGTAAAGT GAAGAAGATGAGTCATTGGTCGTTATCCAA CATTCCTTACAACCAGATGATTT---CAGC AAAAAGCAACTATCAAGAA-AATCCAAAAT ωF20b
ωG3
pSec1
pSec2B
pBRhor1-17
CH4
λhor1-14
510 540 570 600
CTCGCTAGTAA-GTAGTTGCACCAAACACA ACAAACCAAA-TGATTAGTTGGATAAGCAT TGATGCACTTTTTCAGAAAGCGAGTGGAAA GATGAAAGCAAGCCATGCGATGAAAACTAT ωF20b
ωG3
pSec1
pSec2B
pSec1B
pBRhor1-17
CH4
λhor1-14
630 660 690 720
AAAGACACATGCACCACCAAGGCCACCTTC CATCATCCAAACTTCACACACCTAGACCAC AAGCATCAAAGGCAAGCAAGCAGTAGTCAC CACAAATCCAACATGAAGACCTTCCTCATC ωF20b
ωG3
pSec1
pSec2B
pSec1B
pBRhor1-17
CH4
λhor1-14
750 780 810 840
TTTGTCCTCCTTGCCATGGCGATGAAGATC GCCACTGCCGCTAGGGAGTTAAACCCTAGC AACAAAGAGTTATGA---------- --AATGAACACCTGTAATACAAGACGGATG ωF20b
ωG3
pSec1
pSec2B
pSec1B
pBRhor1-17
CH4
λhor1-14
870 900 930 960
GATCATCGTTGTTTAGTCAATG-AGCGTTC -AATGTAACGATGATAAATAAAGTGATGTG CACCATAATGTGTAACCCC-GACCTATACT AGTTCAAACATGGGAATAAAAGACAAAGAA ωF20b
ωG3
pSec1
pSec2B
pSec1B
pBRhor1-17
CH4
λhor1-14
990 1020 1050 1080
AGTTCTTGTCACAAAGACATTACTCGTAAT TGTTAG-ATTCATGCTGTATTTAGATTTTC ATCCCT-AGCTACAATCTCAGTCTTATGCA TTATCTACCTATCTATGTAGTGAAGTCATA ωF20b
ωG3
pSec1
pSec2B
pBRhor1-17
CH4
λhor1-14
1110 1140 1170 1200
TGAACCTAACATAATTAATGGATGCAAAAC TTGGTTTAAATTGAAG ωF20b
ωG3
pSec1
pSec2B
CH4
λhor1-14
1230 1260 1290 1320
ATTATCTGAAACGGTTCCAGATTGTACATA GACACATTTTTTACAATTGTTGAGGTGCAT ATGAAGCAACCTATACTTAAAGATATGGTG ACATATCACTGGATTCAT-TGAAAATCTCA pSec1
CH4
λhor1-14
1350 1380 1410 1440
GTACTAACTTTATATGCAAGGGGG-AAGAT CATGATGACTTTGTGATGCATTTTATCATT ATTCAAATGGTCTCTTTAATGGACCGGTTT CTCAGATTAGAAAATATAGACTTGGGCCAT pSec1
CH4
λhor1-14
1470 1500 1529
ACGCCATACTGAATTCTTGCAAATGGACAA GATGAAGGCATGATTGAATTTATTCCTTCT ATTTCTCCTGCATGTGTAACTTGATATTG pSec1
CH4
λhor1-14

(positions 945 to 950). The ω-gliadins diverge from the ω-secalins and C-hordeins at this point: the ω-secalin (pSec1) and the C-hordeins (CH4 and λhor1–14; Entwistle 1988) continue to show similarity for another 403 bp and diverge 579 bp (position 1,529) after the polyadenylation site.

The relatedness of the S-poor prolamins can be estimated by comparison of single changes among the three orthologous sequence patterns in Fig. 7. Such sequence differences can consist of either single base or multiple base changes, deletions or insertions. The comparison in Fig. 7 indicates that rye and barley ω-type sequences are more similar to each other than to the wheat sequences. However, only a limited number of S-poor prolamins from the three different species have been isolated and the wheat sequences reported could represent paralogous, rather than orthologous, S-poor prolamins. More gene sequences from all three species are needed for a better comparison of these species' evolutionary relationships.

**Acknowledgments** The authors thank Renato D'Ovidio, Don Kasarda, and Pat Okubara for reading the manuscript and making suggestions. All experiments complied with U.S. laws.

## References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Anderson OD, Greene FC (1997) The α-gliadin gene family. II. DNA and protein sequence variation, subfamily structure, and origins of pseudogenes. Theor Appl Genet 95:59–65

Anderson OD, Litts JC, Gautier MF, Greene FC (1984) Nucleic acid sequence and chromosome assignment of a wheat storage protein gene. Nucleic Acids Res 12:8129–8144

Anderson OD, Litts JC, Greene FC (1997) The α-gliadin gene family. I. Characterization of ten new wheat α-gliadin genomic clones, evidence for limited sequence conservation of flanking DNA, and Southern analysis of the gene family. Theor Appl Genet 95:50–58

Anderson OD, Hsia CC, Torres V (2001) The wheat γ-gliadin genes: Characterization of ten new sequences and further understanding of γ-gliadin gene family structure. Theor Appl Genet (in press)

Cassidy BG, Dvorak J, Anderson OD (1998) The wheat low-molecular-weight glutenin genes: characterization of six new genes and progress in understanding gene family structure. Theor Appl Genet 96:743–750

Clarke BC, Mukai Y, Appels R (1996) The *Sec-1* locus on the short arm of chromosome 1R of rye (*Secale cereale*). Chromosoma 105:269–275

Colot V, Bartels D, Thompson R, Flavell R (1989) Molecular characterization of an active wheat LMW glutenin gene and its relation to other wheat and barley prolamin genes. Mol Gen Genet 216:81–90

Dupont F, Vensel WH, Chan R, Kasarda DD (2000) Characterization of the 1B-type omega gliadins from *Triticum aestivum* cultivar Butte. Cereal Chem 77:607–614

Entwistle J (1988) Primary structure of a C-hordein gene from barley. Carlsberg Res Commun 53:247–258

Entwistle J, Knudsen S, Muller M, Cameron-Mills V (1991) Amber codon suppression: the in vivo and in vitro analysis of two C-hordein genes from barley. Plant Mol Biol 17:1217–1231

Higgins DG, Sharp PM (1989) Fast and sensitive multiple sequence alignments on a microcomputer. CABIOS 5:151–153

Hull GA, Halford NG, Kreis M, Shewry PR (1991) Isolation and characterization of genes encoding rye prolamins containing a highly repetitive sequence motif. Plant Mol Biol 17:1111–1115

Kasarda DD, Autran JC, Lew EJ-L, Nimmo CC, Shewry PR (1983) N-terminal amino-acid sequences of ω-gliadins and ω-secalins: implications for the evolution of prolamin genes. Biochim Biophy Acta 747:138–150

Lynn DA, Angerer LM, Bruskin AM, Klein WH, Angerer RC (1983) Localization of a family of mRNAs in a single cell type and its precursors in sea urchin embryos. Proc Natl Acad Sci USA 80:2656–2660

Masci S, Egorov TA, Ronchi C, Kuzmicky D, Kasarda DD, Lafiandra D (1999) Evidence for the presence of only one cysteine residue in the D-type low-molecular-weight subunits of wheat glutenin. J Cereal Sci 29:17–25

Rafalski JA (1986) Structure of wheat gamma-gliadin genes. Gene 43:221–229

Rafalski JA, Scheets K, Metzler M, Peterson DM, Hedgcoth C, Söll DG (1984) Developmentally regulated plant genes: the nucleotide sequence of a wheat gliadin genomic clone. EMBO J 3:1409–1415

Reeves CD, Okita TW (1987) Analyses of alpha/beta-type gliadin genes from diploid and hexaploid wheats. Gene 52:257–266

Sabelli PA, Shewry PR (1991) Characterization and organization of gene families at the *Gli-1* loci of bread and durum wheats by restriction fragment analysis. Theor Appl Genet 83:209–216

Sainova OV, Mekhedov SL, Zhelnin LG, Khokhlova TA, Anan'ev EV (1993) Nucleotide sequence of the barley C-hordein gene. Genetika 29:1070–1079

Scheets K, Hedgcoth C (1988) Nucleotide sequence of a gamma gliadin gene: comparisons with other gamma gliadin sequences show the structure of gamma gliadin genes and the general primary structure of gamma gliadins. Plant Sci 57:141–150

Sumner-Smith M, Rafalski JA, Sugiyama T, Stoll M, Söll D (1985) Conservation and variability of wheat alpha/beta-gliadin genes. Nucleic Acids Res 11:3905–3916

Tatham AS, Shewry PR (1995) The S-poor prolamins of wheat, barley and rye. J Cereal Sci 22:1–16