Theor Appl Genet (1997) 95:50–58

O. D. Anderson · J. C. Litts · F. C. Greene

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

Received: 1 October 1996 / Accepted: 20 December 1996

Abstract A detailed study of the organization of the wheat α-gliadin gene family is described. In the first stage of this study recombinant libraries of sufficient size and quality were constructed to allow examination of large, closely-related, gene families. A large number of α-gliadin clones were then isolated in a partial screen of these libraries. A set of these clones were sequenced and ten new gene sequences are now reported. Including these new sequences, 20 of the 27 known α-gliadin genomic and cDNA sequences are from cv Cheyenne, making this cultivar the best studied for gliadin family organization. An analysis of the DNA flanking the coding regions shows divergences which may indicate the functional ends of the α-gliadin genes. High-resolution Southern analyses on DNA from aneuploid and chromosome-substitution lines allow most of the 20 distinct α-gliadin *Hin*dIII restriction fragments in cvs Cheyenne and Chinese Spring to be assigned to specific genomes. It is estimated that as many as 150 α-gliadin genes occur in the cultivar Cheyenne, and the limitations of such estimates are discussed.

Key words Wheat · α-Gliadins · Storage proteins · Prolamines · Gene family · Sequence conservation · Southern analysis

Communicated by G. E. Hart

O. D. Anderson (✉) · J. C. Litts[1] · F. C. Greene[2]
USDA, ARS, Western Regional Research Center,
800 Buchanan Street, Albany, CA 94710, USA

Present addresses:
[1] Onyx Pharmaceuticals, 3031 Research Drive,
Richmond, CA 94806, USA
[2] USDA, ARS, North Atlantic Area, 600 East Mermaid Lane,
Philadelphia, PA 19118, USA

Introduction

Gliadins, the major class of wheat seed proteins made in the endosperm, are mainly monomeric, are high in proline and glutamine, and contribute to dough physical characteristics (Pomeranz 1988). The gliadins were traditionally divided into three groups, the α-, γ-, and ω-gliadins, which are now known to be related by evolution to low-molecular-weight glutenins (Shewry and Tatham 1990) to form the gliadin gene superfamily. Of the four gliadin subgroups, the α-gliadins are typically the most abundant, comprising 15–30% of the seed protein of most cultivars. The α-gliadin genes are found at the *Gli-2* loci on the group-6 homoeologous chromosomes, while all other wheat prolamine genes occur on the group-1 chromosomes (Payne 1987). Estimates of gene copy numbers in different wheat cultivars and ancestral grasses vary from 25 to 150 (Anderson et al. 1984; Harberd et al. 1985; Reeves and Okita 1987), and there is a high degree of variability in the number of gliadin proteins and the restriction fragment length polymorphism among different cultivars (D'Ovidio et al. 1992). These differences are believed to be due to duplications and deletions of chromosome segments, probably generated by unequal crossing-over, and by gene conversion events. D'Ovidio et al. (1991) have described one such deletion of a block of α-gliadin genes, and the existence of closely related α-gliadin sequence sub-families has been described (Anderson 1991; Anderson et al. 1991). Besides their importance as a major protein component of the human diet and their relation to wheat quality, the α-gliadins are a major initiator of intestinal damage in coeliac disease (Shewry et al. 1992).

Several partial and complete wheat α-gliadin cDNA sequences have been reported: six from cv Cheyenne (Kasarda et al. 1984; Okita et al. 1985) and one from cv Chinese Spring (Garcia-Maroto et al. 1990). In addition, nine α-gliadin genomic sequences are known: five from cv Yamhill (Anderson et al. 1984; Rafalski et al.

1984; Sumner-Smith et al. 1985; Reeves and Okita 1987), one from the wild diploid wheat ancestor *Triticum urartu* (Reeves and Okita 1987), and three from cv Cheyenne (Anderson 1991). Since these sequences comprise only a fraction of the total, and are derived from four different sources, a more thorough description of this gene family, and their derived proteins, is still lacking.

Our research has concentrated on the storage proteins of the wheat cultivar Cheyenne, known for good bread-making quality. Among the genes characterized thus far are the complete set of high-molecular-weight subunit genes (Anderson et al. 1988), and additional sequences from the other gliadin families (Anderson, unpublished). Our objectives are to understand the structure and modes of evolution of these families, characterize the variability among the three genomes of bread wheat, and determine how these proteins affect flour and dough properties. In the following report we present part of a detailed characterization of the α-gliadin gene family organization in the cultivar Cheyenne. We report copy number determinations, chromosome assignments of restriction fragments and individual genes, the sequence of ten new α-gliadin genes, and an analysis of the patterns of flanking DNA divergence. The accompanying paper (Anderson and Greene 1997) analyzes the coding DNA and protein sequences.

Materials and methods

Isolation of genomic DNA

Wheat DNA was isolated from mature embryo cell nuclei. Embryos were obtained by disrupting 100 g of mature seeds for 5 s in a Waring Blendor at the highest setting. Large pieces were pooled and re-run until seeds were completely broken up. The sample was then sifted through a series of 12, 20, and 30 mesh-per-inch screens (we used Tyler sieves). Embryos were collected on the 30 mesh-per-inch screen. Chaff was blown away with a hair dryer. At this stage the embryos could be stored at 4°C for several weeks with no loss of germination ability. To continue purification, embryos were stirred into 100 ml of cold 2M sucrose. Embryos and chaff float, while endosperm fragments sink. The top layer was removed, collected on a nylon screen, and washed with water. This layer was then placed on a step gradient of 50 ml 1M sucrose and 50 ml 2 M sucrose in a 150-ml Corex tube and spun in a HS-4 swinging-bucket rotor (DuPont) at 2500 rpm for 10 min at room temperature. The embryos were removed from the 1–2 M sucrose interface, placed in water, and washed onto a nylon screen.

Nuclei were prepared by grinding 5–20 g of embryos in 0.5–1.0 ml of ice-cold Honda buffer (0.45 M Sucrose, 2.5% Ficoll 400 000, 5% Dextran 40 000, 25 mM Tris pH 7.5, 5 mM $MgCl_2$, 14 mM 2-β-mercaptoethanol, 0.25% Triton X-100) in a pre-chilled mortar. The paste was then diluted with the same buffer and filtered through a 100-μm nylon mesh. Grinding was repeated with the residue and the two filtrates then pooled. The resultant 10–15 ml of extract was transferred to a Dounce homogenizer and a Dounce type-B pestle was passed four times through the sample. The samples were then layered on a three-step Percoll gradient (10 ml 40%, 4 ml 60%, and 4 ml 90%) in buffer A (0.45 M sucrose, 25 mM Tris pH 7.7, 10 mM $MgCl_2$) and spun at 5000 rpm for 20 min at 4°C to pellet starch. The nuclei were collected from the 40–60% and 60–90% interfaces, washed with 3–4 vol of buffer A, pelleted by spinning at 5000 rpm in an HB4 rotor (Dupont) for 10 min, and re-suspended in 5 ml of TE (10 mM Tris pH 7.7, 1 mM EDTA).

To extract DNA, 1 vol of 2× lysis buffer (20 mM Tris pH 8.0, 100 mM EDTA, 200 mM NaCl, 2% sodium laurylsarcosine, and 200 μg/ml Proteinase K mixed just before use) was added to the suspension. The lysate was wrapped in foil, incubated for 2 h at 50°C, extracted twice with phenol, twice with phenol/chloroform, and once with chloroform. If the second phenol interface was large, the phenol extraction was repeated 1–2 times. One-third vol of 10 M NH_4Ac and 1 vol of ice-cold ethanol were added. The DNA was spooled out and re-suspended at approximately 200 μg/ml in TE. If the re-suspension step proceeded too slowly, the sample was incubated at 40–50°C to assist hydration.

Genomic libraries and isolation of α-gliadin clones

Two types of genomic lambda libraries were constructed with cv Cheyenne embryo genomic DNA. The first library type used a partial *Eco*RI digest fractionated on a 10–40% sucrose gradient in 20 mM Tris pH 8.0, 5 mM EDTA, 1 M NaCl. The fractions containing DNA in the 10–20 kb size range were ligated to *Eco*RI arms of the lambda vector λSep6-lac5 (Maniatis et al. 1982), packaged into phage particles (Stratagene, La Jolla, Calif.), and plated on the recA line 1046 (Maniatis et al. 1982). A second set of subgenomic *Eco*RI libraries was constructed from complete digests of Cheyenne genomic DNA. DNA was separated into ten size fractions ranging from 3–5 to 20+ kb on sucrose gradients as described above. Individual fractions were precipitated with 2.5 vol of ethanol. DNA from different fractions were individually ligated to *Eco*RI + *Bam*HI-cut lambda EMBL4, packaged into phage particles, and plated on P2392 cells (Stratagene).

Clones were isolated from 150-mm agar plates each containing 15 000–20 000 plaques. DNA was transferred from the plaques to nitrocellulose filters (Schleicher and Schuell). Phage DNA was fixed to the filter by denaturing for 15 min with 0.5 M NaOH, 1.5 M NaCl, then neutralized with 1 M Tril pH 7, 3 M NaCl. Filters were then air dried and heated for 1 h at 75°C *in vacuo*. The filters were pre-hybridized overnight in 5×SSC Denhardts (5% Ficoll, 5% BSA, 5% PVP), 0.1% SDS at 65°C. Filters were handled in batches of 20–40 shaken together with rotatory action in a crystallizing dish in a water bath at 65°C. The rotation speed was adjusted until all filters floated free from one another. After hybridization, the filters were washed in the same dish under the same conditions: four washes with 2×SSC, 0.5% SDS, and two washes with 0.5×SSC, 0.5% SDS. Each wash lasted 20–40 min. The probe was either the insert of the cDNA clone pA10 (Kasarda et al. 1984) or the α-gliadin complete coding sequence in clone pY16K (Litts, unpublished). The α-gliadin clones from the λSep6 library were named by CNN (standard abbreviation for cv Cheyenne) plus the clone number; e.g., CNN5. Clones from the subgenomic libraries were named by the CNN, batch, plate, and clone number; e.g., CNNE17C (plate batch E, plate No. 17, plaque No. C).

Lambda DNA was isolated as described by Maniatis et al. (1982). DNA fragments containing α-gliadin sequences were subcloned into M13mp 8, 9, 10 or 11 (Bethesda Research Labs). Complete sequencing of both strands was by the dideoxy method (Sanger et al. 1977) using a Klenow fragment (Pharmacia) or T7 DNA polymerase (Sequenase from US Biochemical). Deletion subclones were constructed by the method of Dale et al. (1985). Gaps in the sequences were filled using unpurified primers (Sanchez-Pescador and Urdea 1984) made on a Beckman System 1 DNA Synthesizer.

Sequence analysis was performed using the Intelligenetics/BIONET and Lasergene (DNAstar) programs.

Southern analysis

Vertical agarose-gel electrophoresis (3 mm thick) was used to separate DNA for Southern analysis. DNA transfer to nitrocellulose membranes, pre-hybridization, and hybridization were performed using the same conditions as in clone isolation except that membranes were probed individually in sealed bags. The α-gliadin probe was the insert from clone pY16K (entire coding sequence). The γ-gliadin probe was the insert from genomic clone pγ13 (coding sequence plus 700 bp of 5′ flank; Torres and Anderson, unpublished).The LMW-glutenin probe was the insert from the cDNA clone pTdUCD1 (Lassner and Dvorak 1986; GENBANK accession No. X51759).

Results and discussion

α-Gliadin clone isolation

The wheat α-gliadin genes are found within (6–20) kb *Eco*RI fragments (Anderson et al. 1984; Fig. 1). In addition, it was known that the other wheat prolamine gene family members were found within clonable-sized *Eco*RI fragments. In order to conduct an extensive survey of the α-gliadins and other prolamine gene families, a set of high-quality genomic libraries was prepared from cv Cheyenne DNA (Table 1) containing a total of 15.9×10^6 individual clones. However, the complete-digest libraries constructed from DNA fragments smaller than 10 kb typically contain two or more independent inserts from different regions in the Cheyenne genome. Thus, the total number of individual inserts in this library set is approximately 30×10^6. These libraries are probably representative of the entire genome since only partial screens have resulted in the isolated of the large numbers of gliadin genes (this report and Anderson, unpublished) and the entire set of six high-molecular-weight glutenin genes of cv Cheyenne (Anderson et al. 1988).

A total of 98 lambda clones giving strong positive hybridization signals with α-gliadin probes were single-plaque purified. Restriction analysis indicated at least 54 distinct clones. The remaining 44 clones are apparently multiple isolates of individual clones or else clones of genes so closely related as to give identical restriction patterns. Figure 1 shows that the distribution of *Eco*RI insert sizes in these clones is similar to the *Eco*RI fragments found by probing Cheyenne genomic DNA. Figure 2 shows the restriction patterns of some of the λ Sep6-CNN library clones cut with *Eco*RI (Fig. 2A) and *Eco*RI + *Hin*dIII (Fig. 2C). Blots of these gels hybridized with an α-gliadin probe identified which fragments contain α-gliadin sequences (Figs. 2B and D). Such restriction patterns were used to tentatively associate clones into subfamilies. For example, CNN16 and CNN35 have similar *Eco*RI and *Eco*RI + *Hin*dIII patterns, and sequencing subsequently showed that they belong to a closely related subfamily (Anderson 1991). Similarly, CNN5, CNN52, CNN54 have similar restriction patterns and belong to another subfamily (Anderson and Greene 1997). However, similarity in restriction patterns suggests, but does not prove, relationships. CNN10 and CNN18 contain α-gliadin genes within 7.6-kb *Eco*RI fragments and 3.5-kb *Eco*RI + *Hin*dIII fragments, but the double digest with *Eco*RI + *Hin*dIII shows different adjacent genomic fragments, and sequence analyses show these to be relatively distantly related α-gliadin genes (Anderson and Greene 1997).

It was anticipated that this study of the α-gliadin gene family would reveal information about the *Gli-2* loci organization: instances of multiple genes on single

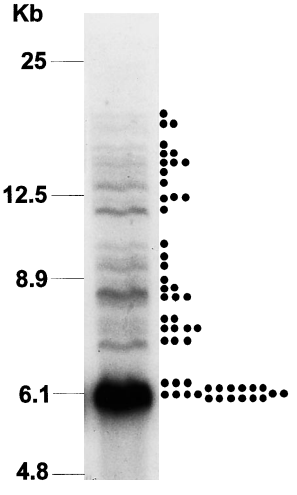

Fig. 1 α-Gliadin genomic restriction pattern and cloned *Eco*RI fragments. Total wheat cv Cheyenne DNA was restricted with *Eco*RI, blotted to nitrocellulose membrane and probed with an α-gliadin coding sequence. DNA marker sizes are shown on the left of the figure. *Dots* on the right side indicate individual α-gliadin-clone *Eco*RI inserts. All *dots* to the immediate right of the intense 6.1-kb band represent 6.1-kb inserts

Table 1 Wheat cv Cheyenne genomic lambda libraries. Listed are the complete set of Cheyenne genomic lambda clones constructed for study of the prolamine gene families

Library[a]	Insert sizes[b]	No. of clones[c]
Sep6-CNN	11.0–17.0	1.5×10^6
GT15/16	1.8–2.3	4.7×10^5
GT17	2–2.8	2.5×10^5
R18	3–4.5	1.2×10^6
GT18	3–4.5	4.1×10^5
GT19	4–6	2.9×10^5
R19	4–6	5.6×10^5
R20	5–7.5	2.8×10^6
GT20	5–7.5	5.3×10^5
R21	6–10	4.4×10^5
R22	8–12	2.4×10^5
R23	10–15	1.5×10^6
R24	10–17	1.4×10^6
R25	12–20	1.1×10^6
R26	15–20 +	8.8×10^5
R27	17–20 +	2.1×10^5
Total		15.9×10^6

[a] Sep6-CNN library was constructed from partial *Eco*RI fragments. All other libraries were constructed from fractionated complete *Eco*RI digests

[b] DNA size range used in library construction and size range of majority of inserts

[c] Number of independent clones upon initial plating of packaged ligations

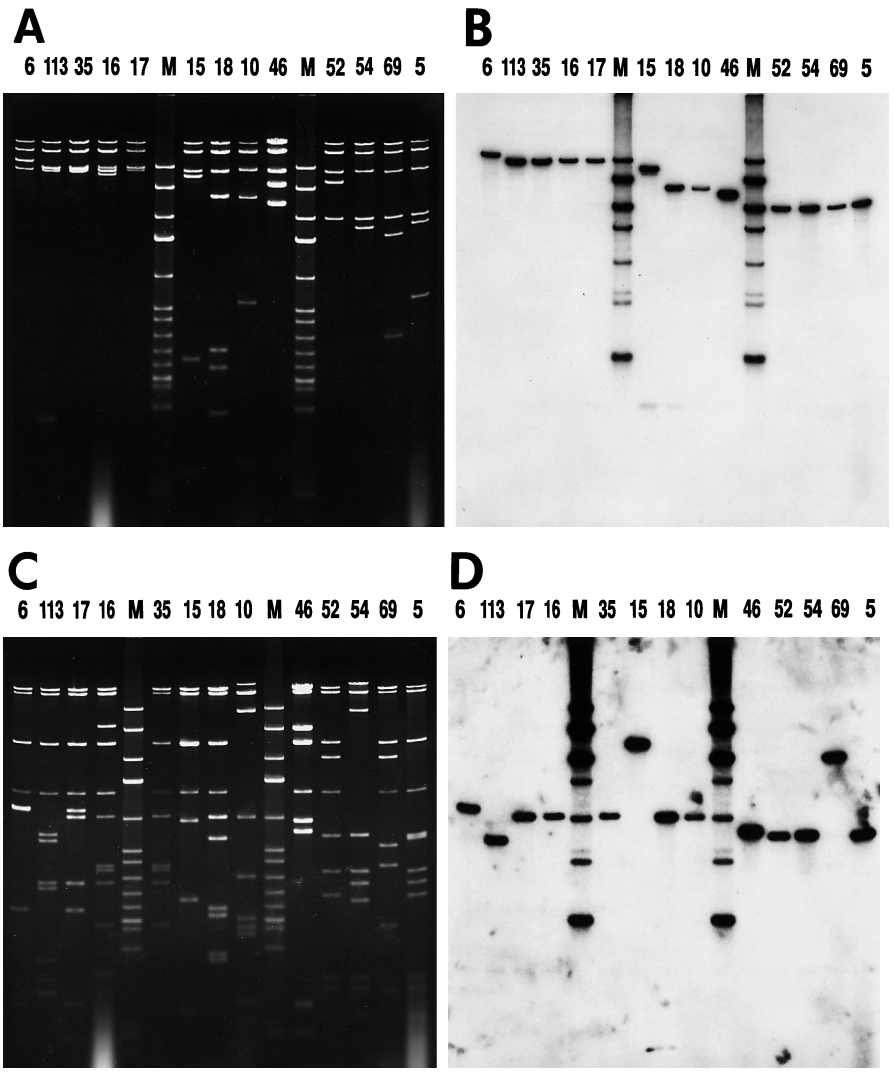

Fig. 2A–D α-Gliadin clone characterization. α-Gliadin lambda clones were restricted with *Eco*RI (**A** and **B**) or *Eco*RI + *Hin*dIII (**C** and **D**), separated by agarose gels, then blotted, and probed with an α-gliadin sequence. Clone numbers are given above each lane and are short versions of the complete name; i.e., clone number 5 is CNN5. **A** and **C**: ethidium bromide staining. **B** and **D**: autoradiography. Markers (*M lanes*) are restrictions of plasmid α-gliadin sequence containing *Eco*RI subclones of CNN5 and CNN10. The three larger fragments common to all clones are the separated and annealed lambda vector arms

large lambda inserts would have allowed estimates of gene spacing. However, examination of a large number of α-gliadin lambda clones, including all clones from a partial *Eco*RI library of up to 17-kb inserts, identified only one (CNNE3A, data not shown) with apparently two α-gliadin sequences. This suggests a minimal spacing of 17 kb between α-gliadin genes. Anderson and Greene (unpublished) have also identified a 48-kb cosmid containing a single α-gliadin sequence. Since most of the Cheyenne α-gliadin genes are found at the *Gli-2A* locus on chromosome 6A, the minimum size of this locus is greater than 2×10^6 bp (assuming 120 genes of a minimal 17 kb spacing), and probably significantly larger.

DNA sequence analysis

DNA sequence analysis was carried out on 17 distinct clones selected to represent different restriction patterns (and presumably different gene subtypes) and also on some clones with similar restriction patterns (to study similarity within a subtype). Sequences of ten previously unreported clones (Table 2) extend from 400 or more bp upstream of the start codon through and beyond the polyadenylation site. Complete sequences can be obtained using the GenBank accession numbers given in Table 2. Extensive flanking sequence was determined for many of the clones, particularly the entire 6.1-kb *Eco*RI fragment of CNN5 and the 7.6-kb *Eco*RI fragment of CNN10.

The Cheyenne genomic clones of Table 2 exhibit both flanking DNA conservation and divergence among individual genes. It has already been reported that CNN16, CNN35, and CNNE17A are members of a closely related α-gliadin subfamily with a similar DNA sequence over 3500 bp of the *Hin*dIII fragment containing the genes (Anderson 1991). Another subfamily is represented by CNN5, CNN52, CNN54, Cosα I and CNN328A, which are homologous within a 3000-bp *Hin*dIII fragment.

Table 2 Wheat α-gliadin clones. All reported α-gliadin clones are listed. The cv Cheyenne genomic clones also list the *Eco*RI and *Eco*RI + *Hin*dIII fragment sizes from Fig. 2 and data not shown

Clone	Clone type[a]	Cultivar	*Eco*RI[b]	*Eco*RI + *Hin*dIII[b]	Access. no.[c]	Reference
CNN5	G	Cheyenne	6.1	3.0	U08287	This paper
CNN10	G	Cheyenne	7.6	3.5	U51303	This paper
CNN16	G	Cheyenne	12.0	3.5	X54517	Anderson 1991
CNN18	G	Cheyenne	7.6	3.5	U51305	This paper
CNN35	G	Cheyenne	12.0	3.5	X54688	Anderson 1991
CNN52	G	Cheyenne	6.1	3.0	U51306	This paper
CNN54	G	Cheyenne	6.1	3.0	U51307	This paper
CNN113	G	Cheyenne	12.0	2.9	U51304	This paper
CNN318A	G	Cheyenne	7.6	5.3	U51302	This paper
CNN328A	G	Cheyenne	6.2	3.0	U50984	This paper
CNNE17A	G	Cheyenne	16.0	4.0	U54689	Anderson 1991
CNNE18C	G	Cheyenne	9.5	2.3	U51309	This paper
CNNE24A	G	Cheyenne	14.5	6.0	U51308	This paper
CNNCosαII	G	Cheyenne	6.1	3.0	U51310	Anderson and Greene 1997
pA10	C	Cheyenne	–	–	K02069	Kasarda et al. 1984
pA26	C	Cheyenne	–	–	M11074	Okita et al. 1985
pA42	C	Cheyenne	–	–	M11073	Okita et al. 1985
pA212	C	Cheyenne	–	–	M10092	Okita et al. 1985
pA735	C	Cheyenne	–	–	M11075	Okita et al. 1985
pA1235	C	Cheyenne	–	–	M11076	Okita et al. 1985
YAM2	G	Yamhill	6.1	3.0	X01130	Anderson et al. 1984
W8233	G	Yamhill	6.0	3.0	X00627	Rafalski et al. 1984
W1215	G	Yamhill	9.8	3.0	K03074	Sumner-Smith et al. 1985
W8142	G	Yamhill	7.7	3.2	K03075	Sumner-Smith et al. 1985
OKYAM[d]	G	Yamhill	6.2	3.0[e]	–[f]	Reeves and Okita 1987
OKURARTU[d]	G	*T. urartu*	6.2	3.0[e]	–[f]	Reeves and Okita 1987
pMM1	C	Chin. Sp.	–	–	X17361	García-Maroto et al. 1990

[a] G = genomic, C = cDNA
[b] α-Gliadin homologous fragments from *Eco*RI and *Eco*RI + *Hin*dIII digests
[c] GENBANK accession numbers
[d] No clone name given in reference
[e] Reported as *Hin*dIII fragments
[f] Not submitted to GENBANK

A pairwise comparison of all known α-gliadin flanking sequences indicates that complete sequence divergence occurs between subfamilies whose members show high conservation within the subfamily. Figure 3 shows sample dot matrices of the results. Most α-gliadin genomic-characterized sequences are bounded at the 5′ end by a conserved *Hin*dIII site at about −600 bp (or −400 bp in the CNN16 subfamily where a base change created another *Hin*dIII site). The only further upstream sequence obtained thus far are from clones CNN5, CNN10, CNNE18C, and OKYAM. These four sequences fall into three sequence classes (CNN10 and OKYAM are similar through the extent of the CNN10 sequence), each class diverging completely from the other two at approximately −630 bp from the start codons.

More sequence information is available for 3′ flanks. Most known sequences are homologous at least through the polyadenylation site (Fig. 4), except for the CNN16 pseudogene subfamily which diverges 28 bp downstream from the stop codon (Anderson 1991). Among other genes, sequences downstream from the polyadenylation site can be separated into four sequence classes with CNN5, CNN10, CNNE18C, and CNN113 as representative class members. Three of these are the same groups as identified by 5′ flanking-sequence analyses. The CNN10 group includes CNN10, CNN18, W1215, and OKYAM. In this group, CNN18 diverges from the others at 400 bp downstream from the polyadenylation site. CNN10, W1215, and OKYAM continue homology through the available sequence: approximately 500 bp for OKYAM and 1500 bp for W1215. In all cases, there were no further detectable homologies between sequence classes past the first point of sequence divergence except for CNN113 which contained a short homology to the CNN5 subfamily 900 bp further downstream. More sequences are needed at and beyond these junction points, but we speculate, based on sequence conservation, that the borders of the functional α-gliadin DNA sequence are from about −630 to the polyadenylation site. This possibility can now be tested in vivo using wheat transformation (Weeks et al.1993) with reporter genes flanked by the regions to be tested.

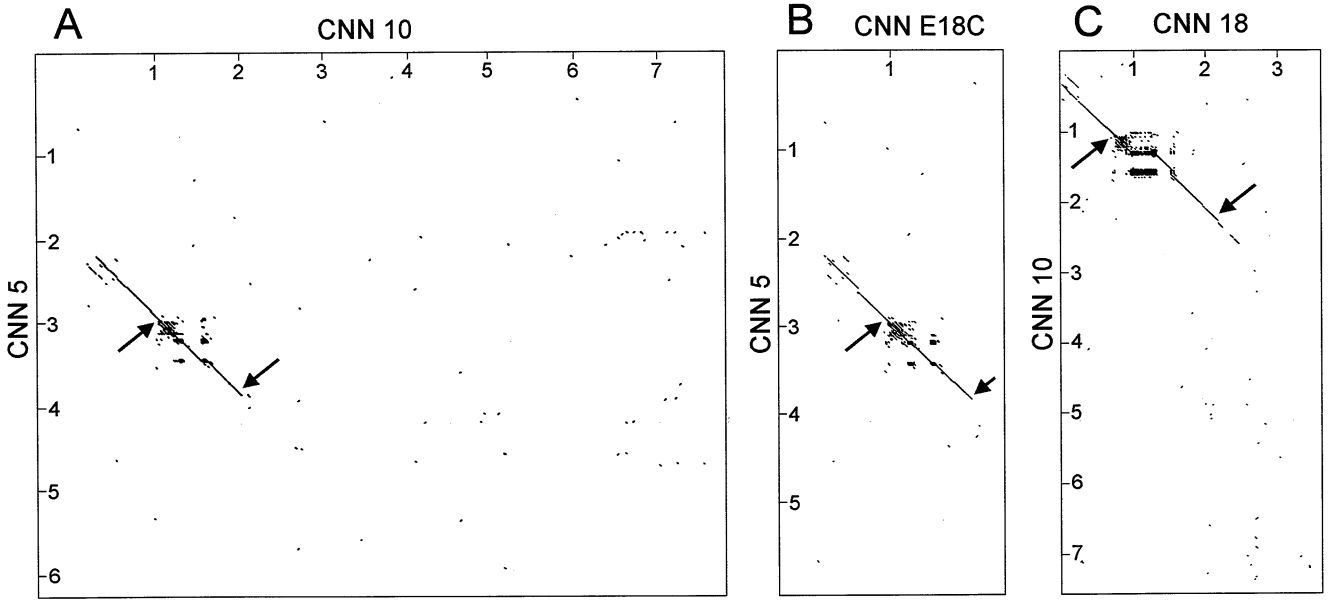

Fig. 3A–C Homology matrices of selected α-gliadin DNA sequences. The α-gliadin sequences were compared pairwise by the Dotplot option of the Lasergene software. The criterion of homology was an 80% match over a 20-bp window. Initiation codons (ATG) and polyadenylation sites are indicated by *arrows below* (to the left) *and above* (to the right), respectively, on the diagonals. Different scales for each comparison are marked in kilobases on the axes. **A** CNN5 vs CNN10. **B** CNN5 vs CNNE18C. **C** CNN10 vs CNN18

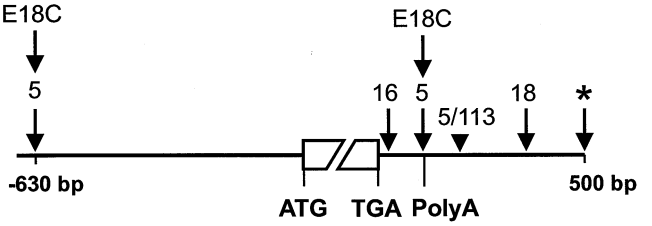

Fig. 4 Divergence points in α-gliadin flanking DNA sequences. All known α-gliadin genes with significant flanking sequences were compared pairwise to determine points of sequence divergence. The *broken box* represents the coding sequence and *lines* represent the flanking sequences. *Arrows* show points of pairwise divergence. The reference sequence is CNN10. The upstream divergence site is 630 bp from the initiation codon. The downstream divergence sites are counted as from the stop codon. *ATG* = initiation codon. *TGA* = termination codon. *PolyA* = polyadenylation site. Clone numbers above the *arrows* indicate where that clone class' sequence diverges from that of CNN10. The *arrowhead* indicates the position at which the sequences of CNN5 and CNN113 diverge from one another. The *asterisk* indicates the end of the available sequence for OKYAM; CNN10 and W1215 continue homologous further downstream

An examination of the extensive flanking sequences of CNN5 and CNN10 found open reading frames of up to 450 bp, but no standard control elements were associated with them; nor did a search of DNA and protein databases find any similar sequences. Use of flanking regions to probe total DNA found extensive hybridization data (not shown) indicating the presence of repetitive DNA. However, internal comparison did not reveal significant repeated DNA within an individual clone or between CNN5 and CNN10 flanks (Fig. 3A). Neither gene's flanking DNA contained simple sequences, but the CNN10 3′ flank had several regions with as high as 75% AT, and one 62% GC region.

Southern analysis and gene copy numbers

Since the α-gliadin genes are closely related to the γ-gliadins, ω-gliadins, and LMW-glutenins, a careful Southern analysis needs to exclude cross-hybridization among the genes of the various gliadin types. No genes from the ω-gliadin family have yet been reported. However, our laboratory has cloned a large number of genes of the γ-gliadin and LMW-glutenin families (Anderson, unpublished). Using representatives of each family as probes, Southern analysis shows that the α-gliadin, γ-gliadin, and LMW-gutenin genes hybridize most strongly to bands from their respective families, although a low degree of cross-hybridization does occur (Fig. 5). Densitometry of these gels and autoradiograms to quantify both relative amounts of DNA and hybridization intensities indicate that each probe results in approximately 20-times more signal intensity when hybridized to other members of its respective family than to the other two families under the conditions used in this study. Southern analysis of the γ-gliadin and LMW-glutenin gene families reveals no signal strong enough to approach even a single-copy density on an α-gliadin-probed Southern. On the other hand, a study

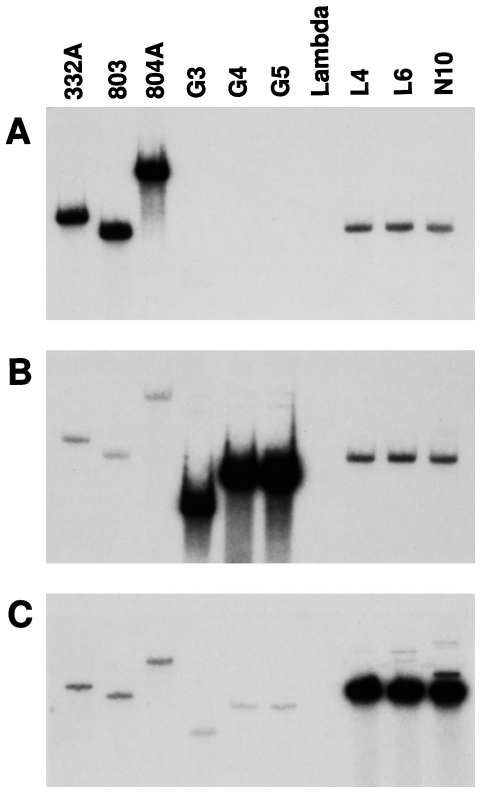

Fig. 5 Cross-hybridization among gliadin families. Three individual *Eco*RI clones of each of the α-gliadin (*332A*, *803*, *804A*), γ-gliadin (*G3*, *G4*, *G5*), and LMW-glutenin (*L4*, *L6*, *N10*) gene families were restricted with *Eco*RI. The restriction fragments were separated by agarose-gel electrophoresis, blotted onto nitrocellulose membranes, and probed for the following sequences: **A** α-gliadin, **B** γ-gliadin, **C** LMW-glutenin

of the γ-gliadin and the LMW-glutenin gene families must consider additive effects of weak hybridizations to the large α-gliadin gene family; i.e., the 6.1-kb *Eco*RI band of Fig. 1 may appear in γ-gliadin- and LMW-glutenin-probed *Eco*RI Southerns.

We have shown that a large number of α-gliadin *Eco*RI bands exist in cv Cheyenne (Anderson et al. 1984; Fig. 1). Based on such data we had originally estimated that there were 50–60 genes in Cheyenne (Anderson et al. 1984). In addition, we estimated 20 restriction bands, although close band proximity limited detailed analysis. Therefore, additional analyses were undertaken to obtain multiple, higher-resolution Southerns using a number of restriction enzymes and DNA from cultivars and aneuploids. Figure 6A shows two such Southerns with copy number reconstructions of 1, 3, and 9 copies per haploid genome of α-gliadin DNA in adjacent lanes. The banding pattern is complex, up to 28 distinct bands of at least one copy could be distinguished in the original blots. Many of the bands were composed of multiple copies, from 2 up to 30. With such complex patterns, it is likely that different bands contain both distinctly different α-gliadin genes/subfamilies and multiple members of specific α-gliadin sub-families (Anderson 1991; see below).

Estimates were made of the number of α-gliadin DNA sequences using densitometry and the copy number reconstructions (such as the *Xba*I example in Fig. 6A) and then averaging the results over nine individual experiments. The calculation yields 99 copies for Cheyenne, 40 for Chinese Spring, and 59 for the durum line. However, there is a fundamental technical problem in estimating the size of a complex multigene family using Southern analysis. Unless the family is nearly homogeneous, the differences in homology with the probe (usually a sequence from a single member of the gene family) will result in different extents of hybridization with different family members. Note in Fig. 6A that the 3.0-kb and 6.1-kb copy number markers in lane 9 (both from clone CNN5) give a more intense signal than the 3.5-kb and 7.6 kb markers (both from clone CNN10) using the CNN5 coding region as a probe, even though the same amount of DNA was used. Further analysis of the signals from a set of clones suggests that there can be a 2-fold difference in the relative signals under the hybridization conditions of this study. Sequencing and comparison of sequence show that the two genes used in the copy number estimates (Fig. 6; clones CNN5 and CNN10) are among the most distantly related of the α-gliadin (Anderson and Greene 1997). Based on these observations, the copy number estimates were increased by 50% to give approximately 150 α-gliadin gene copies for cv Cheyenne, 60 for Chinese Spring, and 90 for the durum. While these are only estimates, we believe the numbers are as accurate as can be obtained with available techniques.

Chromosome assignments

Several reports have identified specific wheat gliadin gene restriction fragments with one or more of the three genomes of hexaploid wheat (Anderson et al.1984; Harberd et al.1985; Reeves and Okita 1987; Sabelli and Shewry 1991). The use of cv Chinese Spring and its aneuploids, especially the set of group-6 nullisomic-tetrasomic lines, has been used to make α-gliadin-fragment chromosome assignments (Harberd et al.1985). However, while Chinese Spring is an important cultivar for genetic research, it is of poor flour-utilization quality. Our work has focused on cv Cheyenne, a high-quality wheat. No nullisomic-tetrasomic Cheyenne lines are available, but we have explored the use of Cheyenne chromosome substitutions into Chinese Spring and the appropriate Chinese Spring nullisomic-tetrasomic lines to make chromosome assignments of restriction fragments. Figure 6B shows the chromosome assignments for Cheyenne and Chinese Spring from a *Hin*dIII restriction experiment. Twelve Cheyenne fragments were assigned to the A-genome, and eight fragments are tentatively assigned to either the B- and/or D-genomes. One of these, a 5.0-kb *Hin*dIII fragment, is either from the B- or D-genome, and

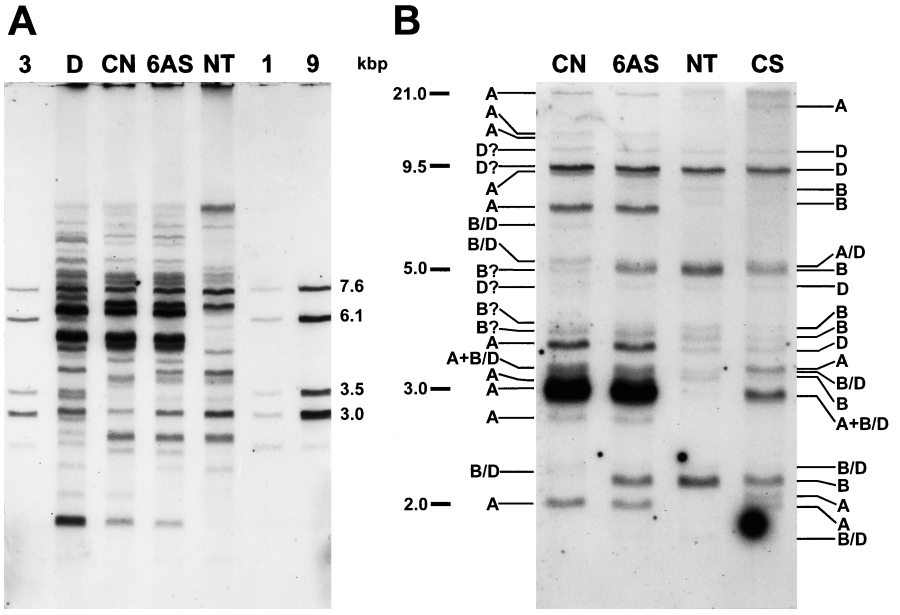

Fig. 6A, B Southern analysis of the α-gliadin gene family. Genomic DNA was restricted and separated on 0.7% agarose gels, then transferred to nylon membranes and probed with an α-gliadin coding-sequence DNA. Copy number reconstructions used cloned genes (CNN5 and CNN10) added to a background of sea urchin DNA. **A** *Xba*I. *Lanes labeled 1*, *3*, and *9* indicate those numbers of copies of the bands per haploid genomic DNA. *D* = Durum wheat (A and B genomes). *CN* = cv Cheyenne. *6AS* = Cheyenne 6A chromosome substituted into cv Chinese Spring. *NT* = Chinese Spring Nulli-6A-Tetra-6B. **B** *Hin*dIII. Genome assignments for individual bands are indicated on the left for cv Cheyenne and on the right for cv Chinese Spring. Tentative assignments are indicated by *question marks*. Bands with more than one component are indicated as the sums of two genomes. Bands assigned to one of two possible genomes are indicated by a *slash* between the genome designations. DNA marker sizes are given in kilobase pairs (kbp)

may be the same as the B-genome fragment assigned in Chinese Spring (Harberd et al. 1985) but reported missing in Cheyenne by Reeves and Okita (1987). This 5.0-kb *Hin*dIII fragment in cv Cheyenne genomic DNA matches the α-gliadin clone CNN10, and we tentatively assign CNN10 to chromosome 6B. Several sequenced clones are assigned to the 6A chromosome by comparing *Eco*RI, *Eco*RI + *Hin*dIII (Table 1), and *Hin*dIII fragments: CNN5, CNN16, CNN18, CNN35, CNN52, CNN 54, CNN113, CNN328A, CNNCosαII, CNNE17A.

The *Hin*dIII fragment assignments in cv Chinese Spring agree with most of the assignments of 12 bands by Harberd et al. (1985). The differences can be attributed to the higher resolution of the analysis in Fig. 6B. For example, the faint band they see at about 2 kb and assign to 6A, we resolve into two 6A bands. Harberd et al. (1985) also assign the 3-kb band to 6A, but a single-copy band of this size also appears in the NT lane (Fig. 6B; the Chinese Spring N6AT6B DNA). This indicates that the 3.0-kb band in Chinese Spring is a composite of multiple genes on 6A and 1–2 genes in 6D and/or 6B. A major point of difference with their assignments is the 3.2-kb fragment they assign to the 6D chromosome. We see a compound band composed mainly of a 6A component superimposed on a lesser component of either 6B or 6D, or both. Figure 6B also resolves and assigns eight bands in addition to those seen by Harberd et al. (1985). Thus, a total of 20 *Hin*dIII bands were assigned in Chinese Spring: 16 correlated to one chromosome, while four bands are either from one of two possible chromosomes or else the band could be a composite of fragments from two chromosomes. Although this analysis concentrates on the 6A chromosome, it demonstrates that it is possible to make significant numbers of chromosome assignments if the Southern-analysis technique is sufficiently sensitive and multiple restriction enzymes are employed. Additional analyses with aneuploids and substitutions should be able to complete and confirm assignments of most restriction fragments and help assign some of the other characterized clones.

References

Anderson OD (1991) Characterization of members of a pseudogene subfamily of the wheat α-gliadin storage protein genes. Plant Mol Biol 16: 335–337

Anderson OD, Greene (1997) The α-gliadin gene family. II. Microsatellites within the coding sequence are responsible for protein size variation and contribute to pseudogene generation. Theor Appl Genet 95: 59–65

Anderson OD, Litts JC, Gautier M-F, and Greene FC (1994) Nucleic acid sequence and chromosome assignment of a wheat storage-protein gene. Nucleic Acids Res 12: 8129–8144

Anderson OD, Halford NG, Forde J, Yip R, Shewry PR, Greene FC (1988) Structure and analysis of the high-molecular-weight glutenin genes from *Triticum aestivum* cv Cheyenne. Proc 7th Int Wheat Genet Symp, pp 699–704. Institute of Plant Science Research (Pub.), Cambridge. Bable Press, Avon

Anderson OD, Greene FC, Litts JC (1991) Structure of the α-gliadin gene family from the bread wheat cultivar Cheyenne. In: Bushuk, Tkachuk (eds) Gluten proteins 1990. pp 640–645

D'Ovidio R, Lafiandra D, Tanzarella OA, Anderson OD, Greene FC (1991) Molecular characterization of bread wheat mutants lacking the entire cluster of chromosome 6A-controlled gliadin components. J Cereal Sci 14 : 125–129

D'Ovidio R, Tanzarella O, Masci S, Lafiandra D, Porceddu E (1992) RFLP and PCR analyses at *Gli-1*, *Gli-2*, *Glu-1* and *Glu-3* loci in cultivated and wild wheats. Hereditas 116 : 79–85

Dale RMK, McClure BA, Houchins JP (1985) A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18s rDNA. Plasmid 13 :31–40

Garcia-Maroto F, Marana C, Garcio-Olmedo F, Carbonero P (1990) Nucleotide sequence of a cDNA encoding an α/β-type gliadin from hexaploid wheat (*Triticum aestivum*). Plant Mol Bio 14 : 867–868

Harberd NP, Bartels D, Thompson RD (1985) Analysis of the gliadin multigene loci in bread wheat using nullisomic-tetrasomic lines. Mol Gen Genet 198 : 234–242

Kasarda DD, Okita TW, Bernardin JE, Baecker PA, Nimmo CC, Lew EJ-L, Dietler MD, Greene FC (1984) Nucleic acid (cDNA) and amino-acid sequences of α-type gliadins from wheat (*Triticum aestivum*). Proc Natl Acad Sci USA 81 : 4712–4716

Lassner M, Dvorak J (1986) Preferential homogenization between adjacent and alternate subrepeats in wheat rDNA. Nucleic Acids Res 5499–5512

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Okita TW, Cheesbrough V, Reeves CD (1985) Evolution and heterogeneity of the α-/β-type and γ-type gliadin DNA sequences. J Biol Chem 260 : 8203–8213

Payne PI (1987) Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. Annu Rev Genet 38 : 141–153

Pomeranz Y (1988) Composition and functionality of wheat flour components. In: Pomeranz Y (ed) Wheat chemistry and technology, vol II. Am Assoc Cereal Chem Inc, St Paul, USA, pp 97–158

Rafalski JA, Scheets K, Metzler M, Peterson DM, Hedgcoth C, Söll D (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 α/β-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

Sanchez-Pescador R, Urdea MS (1984) Use of unpurified primers in sequencing. DNA 3 : 339–343

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USa 74 : 5463–5467

Shewry PR, Tatham AS, Kasarda DD (1992) Cereal proteins and coeliac disease. In: Marsh MN (ed) Coeliac disease. Blackwell Scientific Publications, London, pp 305–348

Shewry PR, Tatham As (1990) The prolamin storage proteins of cereal seeds: structure and evolution. Biochem J 267 : 1–12

Sumner-Smith M, Rafalski JA, Sugiyama T, Stoll M, Söll D (1985) Conservation and variability of wheat α/β-gliadin genes. Nucleic Acids Res 13 : 3905–3916

Weeks JT, Anderson OD, Blechl AE (1993) Rapid production of multiple independent lines of fertile transgenic wheat (*Triticum aestivum*). Plant Physiol 102 : 1077–1084