Theor Appl Genet (1998) 97: 413–421

S. B. Altenbach

Quantification of individual low-molecular-weight glutenin subunit transcripts in developing wheat grains by competitive RT-PCR

Received: 22 August 1997 / Accepted: 10 February 1998

Abstract Oligonucleotide primers and competitive templates were developed so that competitive reverse transcriptase-polymerase chain reaction (RT-PCR) techniques could be used to measure the accumulation of transcripts corresponding to individual low-molecular-weight glutenin subunit (LMW-GS) genes in developing wheat (*Triticum aestivum*) grains. Primers were based on LMW-GS genes previously cloned and sequenced from the cultivar 'Cheyenne' or on the N-terminal amino acid sequences of the major LMW-GSs accumulated in the cultivar 'Yecoro Rojo'. Competitive templates varied from the LMW-GS target genes by either a restriction site or a small deletion between the primer binding sites. Amounts of individual LMW-GS transcripts were quantified by titrating the amplification products produced from RT-PCR of RNA against amplification products produced from known amounts of the corresponding competitive template. Competitive RT-PCR analysis revealed that transcripts for the different LMW-GS genes varied only 4.5-fold in their levels of accumulation in developing 'Cheyenne' grains 15 days post-anthesis (DPA). Nucleotide sequencing of two of the amplification products revealed LMW-GS genes not previously described in 'Cheyenne'. Both of these genes encode proteins with a single cysteine residue in the repetitive region. Clones corresponding to these new sequences were not represented in a cDNA library prepared from 'Cheyenne' endosperm RNA even though the sequences were abundant in developing seeds and present in the RNA population from which the library was prepared. Competitive RT-PCR was shown to be a sensitive method for quantifying the expression of closely related members of complex gene families that cannot be readily distinguished by hybridization analysis. Such techniques should provide insight into the regulation of LMW-GS genes that are critical for wheat flour quality.

Key words LMW glutenin subunit transcripts · competitive RT-PCR · wheat seed proteins · prolamines

Communicated by G. E. Hart

S. B. Altenbach (✉)
USDA-ARS Western Regional Research Center,
800 Buchanan Street, Albany, CA 94710, USA
e-mail: altnbach@pw.usda.gov
fax: +1 (510) 559–5777

Introduction

Wheat flour forms the basis of a variety of food products largely because of the unique viscoelastic properties conferred by the major storage proteins accumulated in the endosperm. The wheat gluten proteins consist of a complex array of proteins that are classified as either glutenin subunits, proteins that form large polymers by virtue of their ability to form intermolecular disulfide bonds, or gliadins, monomeric proteins having exclusively intramolecular disulfide bonds or no disulfide bonds. The glutenin subunits are further divided into high-molecular-weight glutenin subunits (HMW-GS) and low-molecular-weight glutenin subunits (LMW-GS) on the basis of their mobilities in SDS-PAGE, while the gliadins are subdivided into α, γ, and ω subclasses. Although the specific complement of gluten proteins is important in determining the technological properties of the flour, the molecular interactions of the various gluten proteins and the manner in which individual protein species influence dough properties is not well understood.

The gluten proteins are encoded by large families of closely related genes. Clones representative of the HMW-GS and LMW-GS and the α- and γ-gliadin gene families have been isolated and sequenced. The

HMW-GS gene family consists of 6 genes in most bread wheat cultivars, while the LMW-GS gene family may contain 30–40 members (Sabelli and Shewry 1991, Cassidy et al. 1998), and the α-gliadin gene family may be composed of as many as 150 different genes, many of which are pseudogenes that are not expressed (Anderson et al. 1997). Both glutenin subunits and gliadins are rich in glutamine (about 30–50 mol%) and proline (about 10–30 mol%) (Shewry et al. 1994) as determined from amino acid compositions of the different proteins as well as amino acid sequences derived from the nucleotide sequences of both cDNA and genomic clones. Proteins from both classes contain regions in which multiple glutamine residues are interspersed with short repeating sequences. Differences among members of the same gene family often result from deletions or additions of nucleotides in the repetitive portions of the coding regions.

Because of the complexity of the gluten gene families and the interrelatedness of the members, hybridization probes used to assess the levels of transcripts generally measure the expression of multiple genes. Consequently, changes in the expression of specific genes within a gluten gene family during seed development that might impact the quality of the resulting flour may be masked. Methodologies based on the polymerase chain reaction (PCR) provide a way to study the expression of individual genes of known sequence within complex gene families. Such techniques also facilitate expression studies of genes that have not yet been cloned and provide a means for identifying genes with expression patterns of interest from cDNA and genomic libraries. Moreover, such techniques can be made quantitative (Wang et al. 1989) and thus provide a means to compare the expression of individual gluten genes throughout seed development in plants that have been grown under defined environmental conditions. In this paper, a set of oligonucleotide primer pairs that specifically amplify a number of individual LMW-GS genes in the bread wheat cultivar 'Cheyenne' is presented, and a competitive RT-PCR (reverse-transcriptase polymerase chain reaction) assay is described that allows the levels of specific transcripts from this important gene family to be monitored in developing grains.

Materials and methods

Growth and collection of plant material

Wheat plants (*Triticum aestivum* cvs 'Cheyenne' and 'Arapahoe') were grown in a climate-controlled greenhouse following an 8-week period of vernalization at 4°C. The average daytime temperature was 21°C and the average nighttime temperature was 17°C. Natural light was supplemented with 100-W high-pressure sodium lights to maintain a daylength of 16 h. Flowering spikes were tagged at the beginning of anthesis. Whole grain or endosperm tissue was harvested at defined times after anthesis, frozen in liquid nitrogen, and stored at − 80°C.

Extraction of RNA and DNA

For RNA extraction, 0.1–1 g of tissue from developing grains was ground to a fine powder in liquid nitrogen using a precooled mortar and pestle. The powder was transferred to a tube containing 10 ml of extraction buffer (10 m*M* NaCl, 10 m*M* TRIS-Cl, pH 9.0, 1 m*M* EDTA, 1% SDS) and extracted twice with 10 ml of phenol : $CHCl_3$: isoamyl alcohol (25 : 24 : 1), pH 7.9, and once with 10 ml of $CHCl_3$. Cold 8 *M* LiCl was added to the aqueous phase to a final concentration of 2 *M*. Following incubation at − 20°C overnight, the precipitated nucleic acid was collected by centrifugation at 8000 rpm for 20 min in a Sorvall SS-34 rotor, and the resulting pellet was washed with 5 ml of cold 2 *M* LiCl. After centrifugation, the final pellet was air-dried and resuspended in 200 µl of sterile water. DNA was removed from the sample by treatment with 10 units RQ1 RNase-free DNase (Promega, Madison, Wis.) in 40 m*M* TRIS, pH 8.0, 10 m*M* NaCl, 6 m*M* $MgCl_2$, 10 m*M* $CaCl_2$ at 37°C for 10–15 min. RNA was further purified using a RNeasy column (QIAGEN, Santa Clarita, Calif.) according to the manufacturer's instructions.

For genomic DNA extraction, a 4-g sample of young leaf tissue was ground to a fine powder in liquid nitrogen using a precooled mortar and pestle. The resulting powder was added to 20 ml of lysis buffer (100 m*M* TRIS-Cl, pH 9.5, 2% hexadecyltrimethyl ammonium bromide, 1.4 *M* NaCl, 1% polyethylene glycol 6000, and 20 m*M* EDTA) prewarmed to 74°C. Following the addition of 50 µl β-mercaptoethanol and 4 mg RNase A (QIAGEN, Santa Clarita, Calif.), the sample was incubated at 74°C for 20 min with intermittent shaking. The sample was then cooled to room temperature and extracted with 25 ml $CHCl_3$: isoamyl alcohol (24 : 1). After centrifugation, nucleic acids were precipitated from the aqueous fraction by the addition of 1 volume of isopropanol and incubation at room temperature for 30 min. The DNA was collected by centrifugation for 20 min at 6000 rpm in a Sorvall SS-34 rotor. The resulting pellet was dissolved in 5 ml 1 *M* NaCl at 62°C for 30 min. The DNA was purified further using a QIAGEN tip-100 column (QIAGEN, Santa Clarita, Calif.) according to the protocols of the manufacturer.

RT-PCR

RT-PCR was performed according to the basic protocols accompanying the reagents and enzymes supplied by Perkin Elmer (Foster City, Calif.). Between 50 and 200 ng of total RNA was reverse- transcribed in a reaction containing 50 m*M* KCl, 10 m*M* TRIS-Cl, pH 8.3, 5 m*M* $MgCl_2$, 1 m*M* of each dNTP, 2.5 µ*M* random hexamers, 1 unit/µl RNase inhibitor, and 2.5 units/µl MuLV reverse transcriptase in a final volume of 20 µl. The sample was incubated at room temperature for 10 min, followed by 15 min at 42°C, 5 min at 99°C, and 5 min at 5°C in a Perkin Elmer Cetus DNA Thermal Cycler 480.

Amplifications were performed in 100-µl reaction volumes containing 20 µl of the reverse transcription mix, 2.5 units AmpliTaq DNA polymerase, and 20 pmoles of each oligonucleotide primer. The concentrations of TRIS-Cl, pH 8.3, and KCl in the final reaction were adjusted to 10 m*M* and 50 m*M*, respectively. Primers were synthesized by National Biosciences Incorporated/Genovus (Plymouth, Minn.). For competitive RT-PCR assays, 10-µl volumes containing serial dilutions of the competitive templates were added to the amplification reactions. Amplifications were carried out at 95°C for 90 s, followed by 25 cycles of 95°C for 1 min, 60°C for 1 min and 72°C for 2 min. A final extension was carried out at 72°C for 7 min, and the samples were incubated at 4°C until analysis.

Amplification of each RNA sample without prior reverse transcription confirmed the absence of contaminating DNA.

Construction of competitive templates

Genomic clones for the F23A, F15A, and F24B LMW-GS genes (GenBank Accession Numbers U86027, U86028 and U86026,

Table 1 LMW-GS genes that have been distinguished by PCR analysis, primers and primer sequences designed to amplify each gene, sizes of amplification products obtained from RT-PCR of RNA from developing 'Cheyenne' grains, methods used to confirm identities of amplification products, and competitive templates corresponding to each LMW-GS gene.

Gene	Forward primer	Reverse primer	Product size	Confirmation	Competitive template
F23A	F2 AGACCTTCCTC GTCTTTGCC	R10 ATTGGATGGAAC CCTGAACC	659 bp	Restriction digestion with *Hph*I, *Msp*A1I, *Dde*I, *Spe*I	F23-C *Nco*I site removed from genomic clone
F15A	F2 AGACCTTCCTCGTC TTTGCC	R13 ACCTTGACCTGACT GTTGGG	640 bp	Restriction digestion with *Hph*I, *Mun*I, *Sfc*1, *Alu*I, *Spe*I	F15-C *Nco*I site removed from genomic clone
F24B	F2 AGACCTTCCTCGTC TTTGCC	R17 GTTGGTAGAGACC TTGAACCG	641 bp	Restriction digestion with *Alw*NI, *Mun*I, *Aci*I, *Hin*fI, *Spe*I	F24-C *Nco*I site removed from genomic clone
METSHIP	F5 ATGGAGACTAGCCA CATCCC	R27 GCTGCTGGAGGAA TACTTTGC	500, 525, 570 bp	Sequence analysis of 500-bp fragment, Restriction digestion with *Dde*I, *Rsa*I	F5/R27-525-6 F5/R27 amplification product cloned, internal deletion in clone
GLERPS	F7 TGGTTTGGAGAG ACCATCG	R27 GCTGCTGGAGGAA TACTTTGC	575 bp	Sequence analysis, restriction dgestion with *Rsa*I	F7/R27-5-2-14 F7/R27 amplification product cloned, internal deletion in clone

respectively) were kindly provided by Dr. O. Anderson (USDA-ARS, Albany, Calif.). *Eco*RI fragments containing the complete coding regions of these LMW-GS genes were first subcloned into the plasmid pBlueScript SK(+) (Stratagene Cloning Systems, La Jolla, Calif.). A unique *Nco*I site contained within the coding region of each gene was then removed by digesting the plasmids with *Nco*I, making the DNA ends blunt with the Klenow fragment of DNA polymerase I and religating the ends of the plasmids. The resulting plasmids are called F23-C, F15-C, and F24-C (Table 1). DNA fragments resulting from amplification of these plasmids with gene-specific primers are not cleaved by *Nco*I and therefore can be distinguished from amplification products obtained from RT-PCR of RNA.

For LMW-GS genes that had not been cloned previously, a portion of each gene was amplified by RT-PCR from 'Cheyenne' RNA using the appropriate primer pairs. After purification of the amplification products on a 2% agarose gel and elution of specific fragments using the GENECLEAN II kit (BIO101, La Jolla, Calif.), the resulting DNA was reamplified using *Pfu* DNA polymerase (Stratagene Cloning Systems, La Jolla, Calif.) and inserted into the plasmid pCR-Script SK(+) (Stratagene Cloning Systems, La Jolla, Calif.) according to the manufacturer's instructions. The resulting DNA was used to transform Epicurian Coli XL1-BLUE MRF'KAN competent cells (Stratagene Cloning Systems, La Jolla, Calif.). Many of the resulting clones had the correct primer binding sites but contained small internal deletions. These deleted plasmids were used as competitive templates without further modification.

Analysis of amplification products

Ten-microliter aliquots of amplification reactions were analyzed in 2% agarose gels in TBE buffer following standard procedures. For competitive assays to quantify levels of transcripts corresponding to F23A, F15A or F24B LMW-GS genes, the amplification products were digested with *Nco*I prior to electrophoresis. Gels were scanned, and the amount of material in each band was quantified using an IS1000 bioimaging system (Alpha Innotech Corp, San Leandro, Calif.).

DNA sequencing

DNA sequencing was performed by the University of California Berkeley DNA Sequencing Facility. Nucleotide sequence analyses were performed using GENEWORKS 2.5 software (Oxford Molecular Group, Campbell, Calif.).

Amplification of cDNA library

A wheat endosperm cDNA library containing 1.0×10^6 primary plaques was constructed in the vector lambda ZAP II by Stratagene Cloning Systems (La Jolla, Calif.). Equal amounts of endosperm RNA from 5, 10, 15, 20, 25 and 30 days post-anthesis (DPA) developing 'Cheyenne' grains were pooled and used as the starting material to ensure that a representative endosperm library was obtained. Aliquots of the library were heated to 70°C for 5 min prior to amplification with primer pairs specific for LMW-GS genes.

Results

Primer selection and validation

The nucleotide sequences of LMW-GS genes F23A, F15A and F24B were aligned, and oligonucleotide

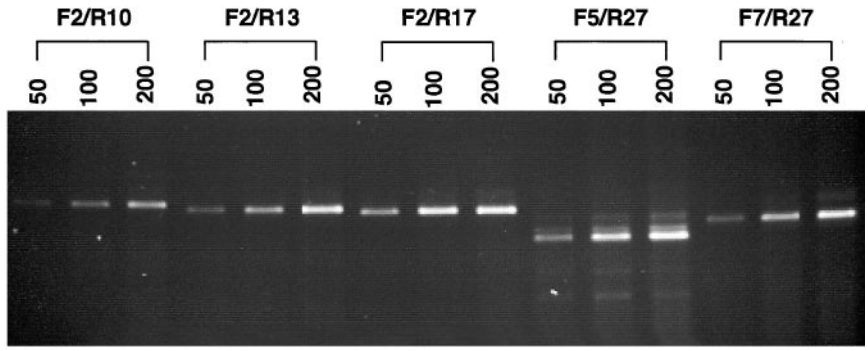

Fig. 1 RT-PCR products produced from increasing amounts of total RNA from 15 DPA grains with primer pairs specific for LMW-GS genes. The amount of RNA (ng) and the primer pairs used in each RT-PCR reaction are indicated *above* each *lane*

primers specific for each gene were designed. Selected primer pairs F2/R10, F2/R13, and F2/R17 span the 5′ coding regions of the genes where the repetitive sequences vary among gene family members. These primers range in size from 19 to 21 nucleotides with GC contents between 47% and 55% and Tms between 59°C and 61°C. The primer sequences are shown in Table 1. When RNA from developing 'Cheyenne' grains was used as the template for RT-PCR, all three primer pairs amplified single fragments of the expected sizes (Fig. 1). The resulting amplification products were further characterized by digestion with a variety of restriction enzymes that distinguish products resulting from the desired gene from those obtained from related sequences (Table 1).

Non-degenerate primers were also designed that correspond to the amino acid sequences METSHIP and GLERPS which are found in the N-terminal regions of some of the most predominant LMW-GSs incorporated into the glutenin polymer of the cultivar 'Yecoro Rojo' (Lew et al. 1992). These primers were tested in combination with a number of primers homologous to conserved regions of the cloned LMW-GS genes. Primer pair F7/R27 amplified a single 575-bp band, while primer pair F5/R27 amplified three bands of 500, 525 and 570 bp when 'Cheyenne' RNA was used as the template for RT-PCR (Fig. 1). The 500-bp fragment amplified by F5/R27 and the F7/R27 amplification product were confirmed by nucleotide sequence analysis to be encoding LMW-GS (Fig. 2).

Reaction conditions for RT-PCR were optimized for all primer pairs. Figure 1 demonstrates that increasing the amounts of RNA template from 50 to 200 ng resulted in corresponding increases in the amounts of amplification products for all primer pairs.

Cloning of DNA fragments amplified by F5/R27 and F7/R27

For precise quantification of transcript levels, plasmids were constructed that could be used as competitors in RT-PCR reactions. The starting plasmids had the same sequences as the target genes to ensure that competitive templates and target genes amplified with equal efficiencies. However, the plasmids were modified so that they either missed restriction sites or contained small deletions between the primer binding sites. Thus, amplification products resulting from the competitive templates could be distinguished from products resulting from RT-PCR of RNA in the same reaction.

When the 500-bp fragment amplified by the F5/R27 primer pair was cloned for preparation of a competitive template, the resulting clones contained the correct primer binding sites (Fig. 3a, lanes 7–11). However, most clones also contained small internal deletions (Fig. 3a, lanes 1–5), suggesting that the F5/R27 amplification product was unstable in the plasmid vector. Only plasmid F5/R27-4-4-5 contained an insert of the expected size (lanes 5 and 7). All plasmids also contained inserts with internal deletions when the F7/R27 amplification product was cloned. Since the plasmids F5/R27-525-6 containing a 430-bp insert and F7/R27-5-2-14 containing a 425-bp insert were suitable templates for competitive RT-PCR, no further attempts were made to isolate clones containing inserts the size of the original amplification products. However, an attempt was made to identify cDNA clones corresponding to the F5/R27 and F7/R27 amplification products from a cDNA library prepared from 'Cheyenne' endosperm RNA since these genes had not been characterized previously from this cultivar. Primer pairs F5/R27 and F7/R27 amplified distinct fragments from either genomic DNA or the RNA preparation from which the cDNA library was constructed. However, the same primers did not amplify the corresponding sequences from an aliquot of the cDNA library, indicating that these genes were not represented in the library (Fig. 3b). In contrast, sequences corresponding to the F15A, F23A, and F24B LMW-GS genes were amplified from genomic DNA, RNA, and the cDNA library with the appropriate primer pairs.

Competitive RT-PCR of RNA from developing wheat grains

Competitive RT-PCR assays were carried out using 100 ng of RNA from 15 DPA developing wheat grains (cv 'Cheyenne') and twofold dilutions of the competi-

tive templates within the range of 0.05 pg to 2 pg, depending on the plasmid (Fig. 4). The levels of individual LMW-GS transcripts were determined by comparison with the amount of amplification products obtained from known amounts of the competitive template, with the amount of transcripts being equivalent to the concentration at which the products from the competitive template equal the products produced from the target RNA (Gilliland et al. 1990). This value is expressed in aMoles/μg RNA, thereby taking into account differences in the sizes of amplification products and competitive templates and facilitating comparisons of transcript levels. Transcripts amplified by the primer pairs F7/R27 (Fig. 4e) and F5/R27 (500-bp fragment) (Fig. 4d) were present in the greatest amounts (2.25 aMoles/μg), while F15A transcripts were the least abundant (0.5 aMoles/μg) (Fig. 4b). Transcripts corresponding to the F5/R27 570-bp fragment (Fig 4d), the F24B gene (Fig. 4c), the F5/R27 525-bp fragment (Fig 4d), and the F23A gene (Fig 4a) were present in intermediate amounts (0.6, 0.8, 1.2, and 1.7 aMoles/μg, respectively).

Sample to sample variation was assessed by subjecting seven different preparations of endosperm RNA from 20 DPA developing grains (cv 'Arapahoe') to RT-PCR with each of the five primer pairs. The amounts of amplification products produced from 100 ng of each RNA sample with each of the five primer pairs were compared. The most variability was detected among PCR products obtained using the F2/R17 primer pair (data not shown). Consequently, all seven RNA preparations were evaluated by competitive RT-PCR using the F2/R17 primer pair (Fig. 5). In the seven samples, the amount of transcripts corresponding to the F24B LMW-GS ranged from 0.85 aMoles/μg to 1.45 aMoles/μg with an average value of 1.2 aMoles/μg.

Discussion

Gene-specific primers were developed so that the levels of individual LMW-GS transcripts could be quantified in developing wheat grains. For studying the expression of genes within complicated multigene families, such as those encoding the wheat gluten proteins, competitive RT-PCR offers distinct advantages over blotting techniques. First of all, hybridization probes take

a

```
atg gag act agc cac atc ccT AGC TTG GAG AAA CCA TTG CAA CAA
 M   E   T   S   H   I   P   S   L   E   K   P   L   Q   Q

CAA CCA TTA CCA CTA CAA CAA ATA TTA TGG TAC CAA CAA CAA CAA
 Q   P   L   P   L   Q   Q   I   L   W   Y   Q   Q   Q   Q

CCC ATC CAA CAA CAA CCA CAA CCA TTT CCA CAA CAG CCA CCA TGT
 P   I   Q   Q   Q   P   Q   P   F   P   Q   Q   P   P   C

CCA CAG CAA CAA CAA CCA CCA TTA TCG CAG CAA CAA CAA CCA CCA
 P   Q   Q   Q   Q   P   P   L   S   Q   Q   Q   Q   P   P

TTT TCA CAA CAA CAA CCA CCA TTT TTG CAG CAA CAA CAA CCC GTT
 F   S   Q   Q   Q   P   P   F   L   Q   Q   Q   Q   P   V

CTA CCG CAA CAA CCA CCA TTT TCG CAG CAA CAA CAA CAA TTT CCG
 L   P   Q   Q   P   P   F   S   Q   Q   Q   Q   Q   F   P

CAG CAA CAA CAA CCA CTT TTA CCG CAA CAA CCA CCA TTT TCG CAG
 Q   Q   Q   Q   P   L   L   P   Q   Q   P   P   F   S   Q

CAA CAA CCA CCA TTT TCT CAG CAG CAG CAA CAA CCA CCA TTT TCG
 Q   Q   P   P   F   S   Q   Q   Q   Q   Q   P   P   F   S

CAG CAA CAA CAA CAA CCA ATT CTA CCG CAA CAA CCA CCA TTT TCA
 Q   Q   Q   Q   Q   P   I   L   P   Q   Q   P   P   F   S

CAA CAC CAA CAA CCA GTT CTA CCG CAA CAA CAA ATA CCA TCT GTT
 Q   H   Q   Q   P   V   L   P   Q   Q   Q   I   P   S   V

CAG CCA TCT ATC TTG CAG CAG CTA AAC CCA Tgc aaa gta ttc ctc
 Q   P   S   I   L   Q   Q   L   N   P   C   K   V   F   L

cag cag c
 Q   Q
```

b

```
  t ggt ttg gag aga cca tcg CAG CAA CAA CCA TTA CCA CCA CAA
    G   L   E   R   P   S   Q   Q   Q   P   L   P   P   Q

CAA ACA TTA TCG CAC CAC CAA CAA CAA CAA CCC ATC CAA CAA CAA
 Q   T   L   S   H   H   Q   Q   Q   Q   P   I   Q   Q   Q

CCA CAA CCA TTT TCA CAA CAG CAA CCA TGT TCA CAG CAA CAA CAA
 P   Q   P   F   S   Q   Q   Q   Q   P   C   S   Q   Q   Q

CAA CCA TTA TCG CAG CAA CAA CAA CCA CCA TTT TCA CAA CAA CAA
 Q   P   L   S   Q   Q   Q   Q   P   P   F   S   Q   Q   Q

CCA CCA TTT TCG CAG CAA CAA CAA CAA CCA TTA TCG CAG CAA CAA
 P   P   F   S   Q   Q   Q   Q   Q   P   L   S   Q   Q   Q

CAA CCA CCA TTT TCA CAA CAA CAA CCA CCA TTT TCG CAG CAA CAA
 Q   P   P   F   S   Q   Q   Q   P   P   F   S   Q   Q   Q

CAA CCA CCA TTT TCA CAA CAA CAA CCA CCA TTT TCG CAG CAA CAA
 Q   P   P   F   S   Q   Q   Q   P   P   F   S   Q   Q   Q

CAA CCA GTT CTA CCG CAA CAA CCA TCA TTT TCA CAG CAA CAA CTA
 Q   P   V   L   P   Q   Q   P   S   F   S   Q   Q   Q   L

CCA CCA TTT TCA CAG CAA CAA TCA CCG TTT TCG CAA CAA CAA CAA
 P   P   F   S   Q   Q   Q   S   P   F   S   Q   Q   Q   Q

ATA GTA CTA CAG CAA CAA CCA CCA TTT TTG CAA CAA CAA CAA CCA
 I   V   L   Q   Q   Q   P   P   F   L   Q   Q   Q   Q   P

AGT CTA CCG CAA CAA CCA CCA TTT TCG CAG CAA CAA CAA CAA CTA
 S   L   P   Q   Q   P   P   F   S   Q   Q   Q   Q   Q   L

GTT CTA CCG CAA CAA CAA ATA CCA TTT GTT CAT CCA TCT ATC TTG
 V   L   P   Q   Q   Q   I   P   F   V   H   P   S   I   L

CAG CAG CTA AAC CCA Tgc aaa gta ttc ctc cag cag c
 Q   Q   L   N   P   C   K   V   F   L   Q   Q
```

Fig. 2a,b Partial nucleotide sequences and deduced amino acid sequences of LMW-GS genes amplified from 'Cheyenne' RNA by F5/R27 and F7/R27 primer pairs. **a** Nucleotide sequence of F5/R27 500-bp amplification product determined from sequencing of F5/R27-4-4-5 clone. The locations of the 23-bp and the 29-bp perfect direct repeats are *boxed* and *underlined*, respectively. Sequences corresponding to the F5 and R27 primers are indicated in *lowercase*. The 72-bp region *between* the *arrowheads* is missing in the clone F5/R27-525-6 that was used for the competitive template. **b** Nucleotide sequence of F7/R27 amplification product determined by direct sequencing. The 42-bp perfect repeats are *underlined*. Sequences corresponding to F7 and R27 primers are indicated in *lowercase*. Modifications in clone F7/R27-5-2-14 used for the competitive template include a 24-bp insertion at the position of the *open arrowhead* and a 174-bp deletion *between* the *closed arrowheads*. The deduced amino acid sequence is shown *below* each DNA sequence in *single letter abbreviation*, and the cysteine residue within the repetitive region of each protein is *enclosed* in a *box*

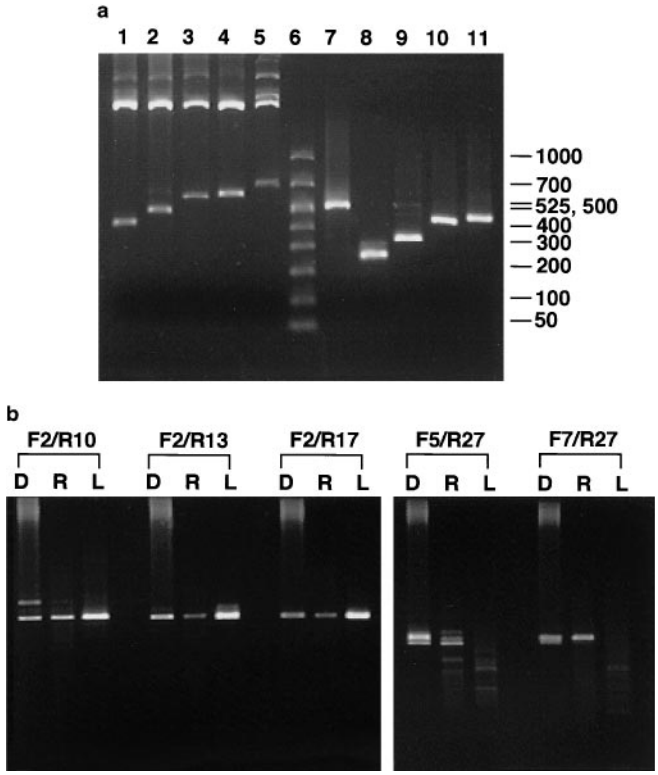

Fig. 3a,b Cloning of LMW-GS sequences. **a** Direct cloning of F5/R27 amplification products. The 500-bp F5/R27 amplification product was ligated to the vector pCR-Script SK(+), and the resulting clones were characterized by restriction digestion and PCR analysis. *Lanes 1–5* Digestion of five representative plasmids with *Bss*HII which cleaves out the insert plus 171 bp of polylinker. DNA from clone F5/R27-4-4-5 which contains an insert of the expected size is shown in *lane 5*. *Lanes 7–11* Amplification of plasmids from *lanes 1–5* with F5/R27 primer pairs. The amplification product from clone F5/R27-4-4-5 (*lane 7*) comigrates with the 500-bp amplification product obtained from RT-PCR of seed RNA (not shown). *Lanes 8–11* show amplification products resulting from clones in *lanes 1–4*, respectively. A 50- to 1000-bp molecular-weight ladder (FMC Bioproducts, Rockland, Me.) used as molecular-weight standards is shown in *lane 6*, and the sizes of the fragments in basepairs are indicated on the *right*. **b** Identification of LMW-GS gene sequences in an endosperm cDNA library. Genomic DNA, endosperm RNA, and a cDNA library were amplified with primer pairs specific for individual LMW-GS genes. The primer pairs used for the amplifications are indicated *above* the *lanes*. The *first lane* in each set contains amplification products from 1 μg genomic DNA (*D*), the *middle lane* contains amplification products from 500 ng of the RNA sample from which the library was constructed (*R*), and the *third lane* contains amplification products from a 10-μl aliquot of the amplified cDNA library (*L*)

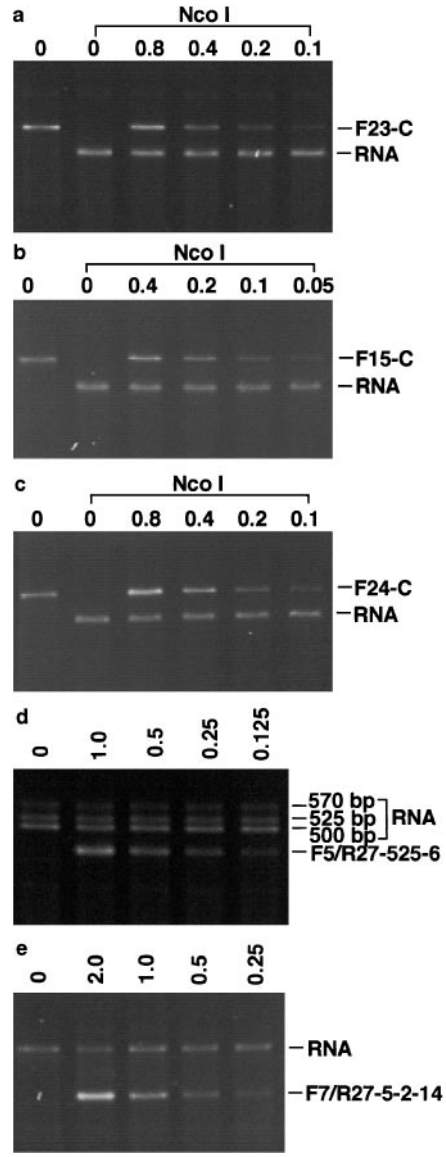

Fig. 4a–e Competitive RT-PCR of LMW-GS genes from cv 'Cheyenne'. RNA (100 ng) from 15 DPA grain was reverse-transcribed and amplified with gene-specific primers along with known amounts of competitive templates specific for each LMW-GS gene. **a** F2/R10 primer pairs plus F23-C template, **b** F2/R13 primer pairs plus F15-C template, **c** F2/R17 primer pairs plus F24-C template, **d** F5/R27 primer pairs plus F5/R27-525-6 template; **e** F7/R27 primer pairs plus F7/R27-5-2-14 template. The amount of competitive template in picograms added to each reaction is shown *above* each *lane*. The *first lane* in each panel contains amplification products resulting from RT-PCR of RNA alone. In panels **a, b** and **c**, amplification products digested with *Nco*I prior to electrophoresis are indicated with a *bracket*. The positions of amplification products produced from the competitive templates (F23-C, F15-C, F24-C, F5/R27-525-6, and F7/R27-5-2-14) and from RT-PCR of the RNA are noted on the *right* of each panel

into account only the relatedness of gene sequences. While hybridization probes have been used to divide cloned members of the gluten gene families into subgroups of closely related sequences (Okita et al. 1985), it is impossible to determine how many family members are detected by a given probe in the absence of cloned sequences from the entire set of genes within a cultivar. In addition, there is not only high homology among family members but also interrelatedness between members of the different gluten gene families, particularly among the LMW-GS and the α- and γ- gliadins (Colot et al. 1989). Thus, some hybridization probes

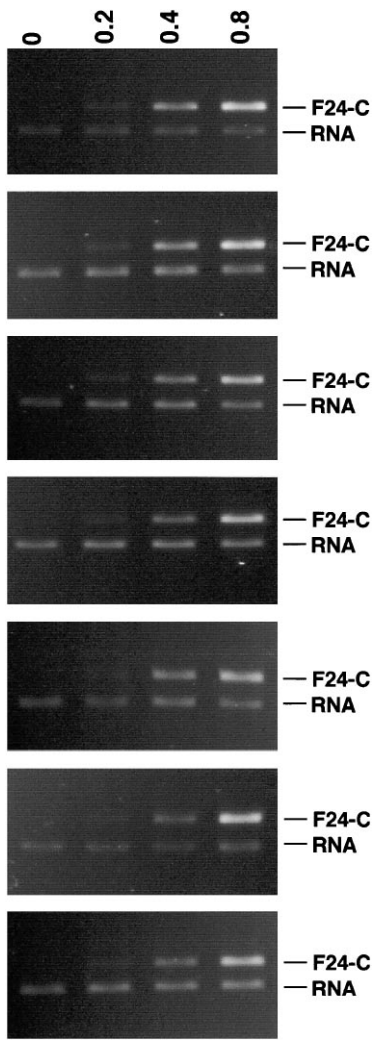

Fig. 5 Competitive RT-PCR analysis of transcripts corresponding to the F24B LMW-GS gene from seven different preparations of RNA from cv 'Arapahoe'. Endosperm RNA (100 ng) from 20 DPA developing grains was reverse-transcribed and amplified with the F2/R17 primer pair along with known amounts of the F24-C competitive template. Amplification products were digested with *Nco*I prior to electrophoresis. The amount of F24-C in picograms added to each reaction is shown *above* each *lane*. Each panel represents the analysis of a different RNA preparation. The positions of amplification products produced from the competitive template F24-C and from RT-PCR of the RNA are noted on the *right* of each panel

may cross-hybridize with members from different gluten gene families (Anderson et al. 1997; Bartels and Thompson 1983). In contrast, PCR techniques can distinguish very closely related genes on the basis of small sequence changes. By using primers that span the repetitive regions, we can distinguish transcripts from genes that differ by only small changes in repeat length. Restriction digestion of the amplification products can further differentiate similar transcripts, and sequencing can be used to verify their identities. In this paper, RT-PCR was used to distinguish seven individual LMW-GS transcripts. Two of these genes, F15A and F24B, share some 87% homology in the coding region and 90% homology in the 3′ flanking sequence. RT-PCR assays using competitive templates also allow for the quantification of individual transcripts. Competitive RT-PCR revealed that transcripts for seven of the LMW-GS are accumulated to similar levels in developing 'Cheyenne' grains, showing only a 4.5-fold difference in amounts.

RT-PCR facilitated analysis of the expression of LMW-GS genes that had not yet been cloned in 'Cheyenne'. By designing primers based on the N-terminal amino acid sequences of the most abundant LMW-GSs in another cultivar, we were able to identify several new transcripts. The sequence of the F7/R27 amplification product does not correspond to any LMW-GS gene reported thus far, while a sequence very similar to that of the F5/R27 amplification product was amplified from genomic DNA of the cultivar 'Chinese Spring' by Van Campenhout et al. (1995) (GenBank Accession Number X84960). The F5/R27 sequence differs from this sequence in four nucleotides that result in four changes in the amino acid sequence of the resulting protein. D'Ovidio et al. (1997) also used PCR to identify a LMW-GS with a similar sequence from durum wheat. The proteins encoded by both the F5/R27 and F7/R27 amplification products have a single cysteine in the midst of the repetitive region that distinguish them from the other LMW-GSs characterized in 'Cheyenne' (Table 2). Proteins encoded by the F23A, F15A, and F24B genes and the B11-33 cDNA (GenBank Accession Number M11077) (Okita et al. 1985) all have a cysteine preceding the repetitive region 5 amino acids from the N-terminus, whereas the protein encoded by the L4 gene (GenBank Accession Number U86030) does not have any cysteine residues in the N-terminal half of the protein. The number and arrangement of cysteine residues in these proteins is likely to be important in the formation of glutenin polymers and thus may influence quality parameters. Although this type of LMW-GS is among the most abundant LMW-GSs incorporated into the glutenin polymer in the cultivar 'Yecoro Rojo' (Lew et al. 1992), the F5/R27 and F7/R27 transcripts were only slightly more abundant in developing 'Cheyenne' grain than the other LMW-GS genes sampled.

The difficulty encountered in cloning the F5/R27 and F7/R27 amplification products coupled with their absence from the cDNA library helps to explain why the clones thus far identified from libraries do not correspond to the predominant LMW-GS sequences in the glutenin polymer fraction (Lew et al. 1992). In addition, it suggests that PCR techniques may be essential for characterizing sequences representative of the entire gene family. It is tempting to speculate that the regularity of repetitive sequences in the 5′ portion of the LMW-GS genes amplified by F5/R27 and F7/R27 are to blame for problems in cloning these genes. A 241-bp

Table 2 N-terminal amino acid sequences of LMW-GSs from the cultivar 'Cheyenne' deduced from nucleotide sequences of cDNA clones, genomic clones, or amplification products. Signal peptide sequences were omitted. Cysteine residues are underlined

LMW-GS	N-terminal sequence
F15A[a]	METSCISGLERPWQQQ PLPPQQSFSQQPPFSQQQQQPLPQQPSFSQQQPP
F23A[a]	METRCIPGLERPWQQQPLPPQQTFPQQPLFSQQQQQQLFPQQPSFSQQQP
F24B[a]	METSCIPGLERPWQQQPLQQKETFPQQPPSSQQQQPFPQQPPFLQQQPSF
B11-33[b]	METSCISGLERPWQQQPLPPQQSFSQQPPFSQQQQQPLPQQPSFSQQQPP
L4[a]	ISQQQQQPPFSQQQQPQFSQQPPFSQQQQPPFSQQQQQPPFAQQQQPPFS
F5/R27[c]	METSHIPSLEKPLQQQPLPLQQILWYQQQQPIQQQPQPFPQQPPCPQQQQ
F7/R27[c]	GLERPSQQQPLPPQQTLSHHQQQQPIQQQPQPFSQQQPCSQQQQQPLSQQ

[a] Genomic clone (Cassidy et al. 1998)
[b] cDNA clone (Okita et al. 1985)
[c] Amplification product; additional residues may precede sequence

region of the F5/R27 amplification product contains two 23-bp perfect repeats and three 29-bp perfect repeats (Fig. 2a). In the F7/R27 amplification product, three 42-bp perfect repeats are found in a 151-bp region (Fig. 2b). Interestingly, one of the 29-bp repeats was absent, and a 23-bp repeat was interrupted in the clone F5/R27-525-6 while two of the 42-bp repeats were absent from the F7/R27-5-2-14 clone. While the repetitive regions of the F15A, F23A, and F24B genes all have some perfect repeats, these are not nearly as long or as frequent as those found in the F5/R27 and F7/R27 amplification products. The repetitive regions of several other LMW-GS genes, in particular, L4 and LP1211 (GenBank Accession Number X07747) (Pitts et al. 1988) are about 100 bp longer than those in the genes amplified by F5/R27 and F7/R27. The repetitive regions from these clones are quite complicated and contain a number of long perfect repeats, some of which are overlapping. Genomic clones for these genes have been obtained, suggesting that the arrangement as well as the number of direct repeats may impact the successful cloning of LMW-GS genes.

The gluten proteins are extremely complex and the repetitive structures as well as the position and number of cysteine residues in the various protein species could have dramatic effects on the functional properties of the resulting flour. Initial RT-PCR studies focused on distinguishing LMW-GS transcripts in the cultivar 'Cheyenne' because a number of cloned sequences were available from this cultivar (Okita et al. 1985; Cassidy et al. 1998). However, the same primer pairs detect transcripts for homologous LMW-GS genes in a number of other cultivars, including 'Arapahoe', a hard red winter wheat, 'Butte', a hard red spring wheat, and 'Bobwhite', a hard white spring wheat (data not shown). Thus, primer pairs and competitive templates should be of use for evaluating levels of specific LMW-GS transcripts in a variety of wheat cultivars. Competitive RT-PCR techniques now offer a way to quantify the expression of individual genes that encode the major gluten proteins in developing grains and should provide new information on the importance of individual components in flour quality.

Acknowledgments Thanks to Kent Suyenaga and Sitsari Kitisakkul for excellent technical assistance, Dr. O. Anderson for making clones and sequences available prior to publication, Charlene Tanaka for preparing RNA for the cDNA library, Drs. W. Hurkman and F. DuPont for helpful discussions and review of manuscript, and Drs. D. Kasarda and R. D'Ovidio for critical review of manuscript. This work was supported by USDA-ARS CRIS 5325-43000-023-00D. Mention of a specific product name by the United States Department of Agriculture does not constitute an endorsement and does not imply a recommendation over other suitable products.

References

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

Bartels D, Thompson RD (1983) The characterization of cDNA clones coding for wheat storage proteins. Nucl Acids Res 11: 2961–2977

Cassidy B, 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 (in press)

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

D'Ovidio R, Simeone M, Masci S, Porceddu E (1997) Molecular characterization of a LMW-GS gene located on chromosome 1B and development of primers specific for the *Glu-B3* complex locus in durum wheat. Theor Appl Genet 95:1119–1126

Gilliland G, Perrin S, Blanchard K, Bunn HF (1990) Analysis of cytokine mRNA and DNA: Detection and quantitation by competitive polymerase chain reaction. Proc Natl Acad Sci USA 87: 2725–2729

Lew E J-L, Kuzmicky DD, Kasarda DD (1992) Characterization of low molecular weight glutenin subunits by reversed-phase high-performance liquid chromatography, sodium dodecyl sulfate-polyacrylamide gel electrophoresis, and N-terminal amino acid sequencing. Cereal Chem 69:508–515

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

Pitts EG, Rafalski JA, Hedgcoth C (1988) Nucleotide sequence and encoded amino acid sequence of a genomic gene region for a low molecular weight glutenin. Nucleic Acids Res 16:11376–11376

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

Shewry PR, Tatham AS, Halford NG, Barker JHA, Hannappel U, Gallois P, Thomas M, Kreis M (1994) Opportunities for manipulating the seed protein composition of wheat and barley in order to improve quality. Transgenic Res 3:3–12

Van Campenhout S, Vander Stappen J, Sagi L, Volckaert G (1995) Locus-specific primers for LMW glutenin genes on each of the group 1 chromosomes of hexaploid wheat. Theor Appl Genet 91:313–319

Wang AM, Doyle MV, Mark DF (1989) Quantitation of mRNA by the polymerase chain reaction. Proc Natl Acad Sci USA 86:9717–9721