

S. Van Campenhout · L. Sági
J. Vander Stappen · G. Volckaert

Characterisation of type-I thionin loci from the A, B, D and R genomes of wheat and rye

Received: 30 January 1997 / Accepted: 23 June 1997

Abstract DNA sequences encoding type-I thionins were isolated from *Triticum aestivum* L. cv ‘Chinese Spring’ using PCR with consensus primers. Blunt-end cloning, sequencing and PCR-based chromosome assignment of these fragments uncovered the three orthologous sequences corresponding to the single-copy genes at the *Pur-1* loci on each of the group-1 chromosomes. Comparison with two previously published cDNA sequences revealed the presence of two introns that contain most of the polymorphic nucleotide sites. The observed orthologous DNA sequence variation among *Pur-1* loci, encoded by each of the A, B and D genomes, enabled us to establish interlocus relationships and to construct locus-specific primer sets. Analogously, the *Pur-R1* sequence from rye was isolated, and a locus-specific primer pair was constructed as well. Hence, four locus-specific primer sets are now available as molecular markers for the homoeologous 1AL, 1BL, 1DL and 1RL chromosome arms. Amplification from several diploid and tetraploid wheat species showed that the primers can be used as molecular tools for studying wheat phylogeny.

Key words Thionin · Orthologous DNA sequence variation · Locus-specific PCR markers · Wheat phylogeny · Wheat-rye translocation

Introduction

Thionins were first extracted as protein-lipid complexes from wheat endosperm by Balls et al. (1942) and further resolved into two fractions, α and β purothionins, by Redman and Fisher (1968). Electrophoretic analysis of the purothionin mixture from compensating nullitetrasonic and ditelosomic lines of ‘Chinese Spring’ by Fernández de Caleyá et al. (1976) led to the conclusion that these proteins are encoded by a set of triplicated (orthologous) loci on the long arms of the group-1 chromosomes, designated *Pur-A1* (encoding the β component), *Pur-B1* and *Pur-D1* (both contributing to the α fraction). Mak and Jones (1976) determined the amino acid sequence of β -purothionin. The α fraction could be split up into an $\alpha 1$ and an $\alpha 2$ component which differ at six amino acid positions from each other (Jones and Mak 1977). Other thionins have been discovered in several cereals (Carbonero and Garcia-Olmedo 1969; Redman and Fisher 1969), including rye (Hernández-Lucas et al. 1978). The latter one was shown to be located on chromosome arm 1RL by Sánchez-Monge et al. (1979). In addition, several homologous cysteine-rich polypeptides have been isolated from a wide range of plant taxa (reviewed by Garcia-Olmedo et al. 1989). Most of these were found to have antimicrobial properties (Stuart and Harris 1942; Fernández de Caleyá et al. 1972; Hernández-Lucas et al. 1974; Bohlmann et al. 1988; Ebrahim-Nesbat et al. 1989; Garcia-Olmedo et al. 1989).

Molecular cloning of barley cDNAs corresponding to α - and β -hordothionins and leaf-specific thionins (Hernández-Lucas et al. 1986; Ponz et al. 1986; Bohlmann and Apel 1987; Gausing 1987) and of a gDNA clone containing a barley α -hordothionin (Rodríguez-Palenzuela et al. 1988) showed that the structural gene includes regions encoding a typical signal peptide, a thionin domain and a C-terminal acidic extension which is interrupted by two introns. Thionins are

Communicated by G. E. Hart

S. Van Campenhout · L. Sági¹ · J. Vander Stappen
G. Volckaert (✉)
Laboratory of Gene Technology, Katholieke Universiteit Leuven,
Willem de Croylaan 42, B-3001 Leuven, Belgium

Present address:

¹Laboratory of Tropical Crop Improvement,
Katholieke Universiteit Leuven, Kardinaal Mercierlaan 92,
B-3001 Leuven, Belgium

synthesised as larger precursors of which the C-terminal acidic extension is posttranslationally excised (Ponz et al. 1983). Recently, Castagnaro et al. (1994) reported the cDNA sequence encoding the precursors of $\alpha 1$ - and $\alpha 2$ -purothionins. Besides the original purothionins (now known as type-I thionins), another type of thionin (type V) was found to be present in wheat endosperm and was further characterised by Castagnaro et al. (1992). Whereas type-I thionins are characterised by four disulfide bridges and a highly basic constitution, type-V thionins have three disulfide bridges and are neutral. The latter were shown to be encoded by a set of triplicated single-copy genes in the close proximity (less than 10 kb) of the *Pur-1* loci corresponding to the type-I thionins. The type-V sequences in wheat and *Aegilops* have shown to be highly conserved (Castagnaro et al. 1995).

We are aiming at constructing locus-specific markers spanning the whole group-1 chromosomes of wheat and rye and, therefore, the availability of DNA sequence data for purothionins together with their known chromosomal location makes the *Pur-1* loci good candidates for the development of polymerase chain reaction (PCR) markers specific to chromosome arms 1AL, 1BL, 1DL and 1RL of wheat and rye, presented here, respectively. Hence, in this study we report the PCR-based isolation and chromosome assignment of DNA sequences encoding the precursor of type-I thionins in wheat and rye. The observed inter-locus (orthologous) variation and the use of the locus-specific primer sets that were developed is discussed in relation to the relationships among A, B and D genomes and their phylogeny.

Materials and methods

Plant material

Aneuploid 'Chinese Spring' (*Triticum aestivum*, L., AABBDD) stocks were developed by Sears (1966) and Sears and Sears (1978); in particular the group-1 nulli-tetrasomic stocks (nulli means absence of the indicated chromosome and tetra means that an extra homoeologous chromosome pair is present to compensate for this absence) and the ditelosomic stocks were employed to ascertain the chromosomal arm locations of the isolated DNA sequences. G. De Wever (Veredelingsstation van Heverlee, Linter, Belgium) provided *Triticum monococcum* (Einkorn, AA); *T. dicoccum* (ssp. *macratherrum* and ssp. *majus*, AABB); *T. durum* (AABB) cvs 'GK Basa', 'Desfontaines', 'Mexicali'; *Triticum timopheevii* (AAGG) and *T. spelta* (Ongebaarde witte, AABBDD). Rye (*Secale cereale* L., RR) var 'Petkus', a wheat 1R(1B) substitution line, a 1R addition to 'Chinese Spring' line (CS + 1R) and 1RS.1BL and 1BS.1RL wheat-rye translocation lines (CS/Imperial) were received from Dr. P. Langridge (Waite Agricultural Research Institute, Glen Osmond, Australia). Dr. S. Reader (John Innes Centre, Norwich, UK) provided *Triticum urartu* (AA) accession number 1010003, *T. monococcum* (AA) accession number 1040006, *T. tauschii* (DD) accession number 2220001 and *T. dicoccoides* (AABB) accession number 1060026.

DNA isolation and polymerase chain reactions

Genomic DNA was isolated from young leaves on a small scale as described by Van Campenhout et al. (1995). For the design of consensus and sequence-specific primers, DNA sequences were compared using the multiple alignment programme of the GENMON (version 4.3) software package (GBF, Braunschweig, Germany). Oligonucleotides were purchased from Pharmacia Biotech (Uppsala, Sweden). PCR reactions contained 10 mM TRIS-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl₂, 200 μ M of each dNTP, 1 μ M of each primer, 2 units *AmpliTaq* DNA polymerase (Perkin Elmer Cetus, Norwalk, Connecticut, USA) and 50 ng genomic DNA in a total volume of 100 μ l. The reaction was overlaid with mineral oil and subjected to incubation for 1 min at 94°C, followed by 30 cycles of 1 min at 94°C, 1 min at 58°C, 90 s at 72°C on a TRIO-thermoblock (Biometra, Göttingen, Germany). To check the specificity of the PCR, we analysed 10 μ l of the reaction mixture electrophoretically on 1.2% agarose gels. Following staining with ethidium bromide, DNA was visualised under UV light. The 123-bp DNA ladder (Gibco BRL, Gaithersburg, Md.) was used as size marker.

Cloning of PCR products

The PCR products were reamplified with *UltmaTaq* DNA polymerase (Perkin Elmer Cetus) that has a 3'-5' exonuclease proofreading activity and generates blunt-end PCR products (rather than fragments with an extra A at the 3' end). Reactions were performed in 100 μ l using commercially supplied buffer, 1.5 mM MgCl₂, 40 μ M of each dNTP, 0.5 μ M of each primer, 2.5 units *UltmaTaq* DNA polymerase and 8 μ l of the original PCR mixture as template. After cycling 20 times at 94°C for 1 min, 55°C for 1 min and 72°C for 1.5 min the reaction products were checked by ethidium bromide agarose gel electrophoresis. The reamplified PCR products were then selectively precipitated by adding an equal volume of polyethylene glycol mix (26.2% PEG 8000, 6.6 mM MgCl₂, 0.6 M NaOAc, pH 5.2). The precipitated samples were then left for 15 min on ice, spun at 13 000 rpm for 15 min and the supernatant (containing residual primers, nucleotides and truncated PCR products) carefully removed. Pellets were washed twice with 70% ethanol and dissolved in 20 μ l water. Subsequently, PCR products were phosphorylated by T4 polynucleotide kinase (New England Biolabs, Beverly, Mass.) in the presence of 1 mM rATP according to the specifications of the supplier. Ligation was carried out as follows: a reaction volume of 10 μ l contained 100 ng PCR product (treated as described above), 100 ng of *Sma*I-cut and dephosphorylated pUC19 vector, 1 mM rATP and 2 units of T4 DNA ligase in the appropriate buffer (Promega, Madison, Wis.). This ligation mixture was incubated at 22°C for 4 h and subsequently used to transform *E. coli* strain DH5 α , allowing the blue-white selection of colonies. The presence of inserted DNA of the expected length in recombinant vectors was determined by PCR analysis directly on white colonies according to Verhasselt et al. (1993).

DNA sequence analysis

DNA sequencing of the cloned fragments was done using Qiagen tip-20 (Qiagen, Hilden, Germany) purified plasmids, and cycle sequencing with the dye-labeled ABI primers -21M13 and M13reverse (Applied Biosystems, Foster City, USA). Sequencing gels were run on a 373A DNA sequencer (Applied Biosystems). The GENMON package was used for pairwise sequence comparisons, calculations of similarity scores, multiple alignments and the conversion of DNA sequences to protein sequences.

Results

Isolation and characterisation of *Pur-1* DNA sequences

To isolate *Pur-1* sequences on each of the group-1 chromosomes of wheat and chromosome 1R from rye, we selected a consensus primer set on the basis of sequenced regions showing minimal variation among thionin genes. This strategy was chosen to maximise the probability of amplifying DNA sequences representing the coding region of the precursor at each locus. After a comparison of two wheat cDNA sequences that encode precursors of the type-I thionins (pTTH14 and pTTH1, Castagnaro et al. 1994) and two type-V thionin sequences (one cDNA and one PCR product, Castagnaro et al. 1992) we found that the sequences immediately downstream of the stop triplet were quite conserved, thereby providing us with the opportunity to design primer PCO3 (Table 1). On the contrary, regions in the proximity of the initiation codon showed only limited homology between pTTH14 and the type-V thionins from wheat (pTTH1 thionin is only a partial cDNA clone and does not contain the initiation triplet). Nevertheless, primer PCO5 (Table 1) could be designed around the initiation codon of the pTTH14 sequence which shows homology to barley leaf-specific thionins (Gausling 1987).

This PURCO primer set amplified a PCR product of about 900 bp in both 'Chinese Spring' nulli-tetrasomic lines and in rye cv 'Petkus'. Because products of approximately 457 bp were expected to be amplified based on the pTTH14 sequence, larger sizes would indicate the presence of introns. PCR products from three nulli-tetrasomic lines (N1AT1B, N1BT1D and N1DT1A) as well as from 'Petkus' were cloned as blunt-end fragments. In each case, 15 white bacterial colonies were subjected to PCR-based clone analysis. On average, 67% of the colonies showed plasmid inserts of about 900 bp. Five positive clones arising from

each genotype were sequenced and altogether, 4 *Pur-1* variants were revealed (named PURA1, PURB1, PURD1 and PURR1; Fig. 1). Comparison of these sequences with the pTTH14 and pTTH1 cDNA sequences corroborated the presence of two introns: the first varies in length from 340 to 353 bp and the second varies from 87 to 91 bp (Fig. 1). There is a perfect match between PURB1 and pTTH1 on one hand, and PURD1 and pTTH14 on the other hand, as far as their common cDNA parts are concerned. The intron sequences are typically A + T-rich and are delimited by 5' and 3' boundaries CAG/GTAA and CAG/, respectively, which are part of the plant splice junction consensus sequence (Hanley and Schuler 1988).

The respective regions of the precursor encoding the signal peptide, the mature thionin and the posttranslationally excised C-terminal peptide were identified by alignment to wheat type-V thionin gene *TthV* (Castagnaro et al. 1992) and to barley type-I thionin gene, *Hth α* (Rodriguez-Palenzuela et al. 1988). Both introns interrupt the region encoding the C-terminal acidic peptide at exactly the same positions as in *TthV* and *Hth α* .

Chromosome assignment of the *Pur-1* sequences

Sequence PURA1 was found in 3 clones arising from N1DT1A and in 1 clone from N1BT1D, sequence PURB1 was found in 4 clones arising from N1AT1B and in 2 clones from N1DT1A and sequence PURD1 was found in 4 clones arising from N1BT1D and in 1 clone from N1AT1B. Thus, the PURA1, PURB1 and PURD1 sequences predominated among clones with inserts originating from N1DT1A, N1AT1B, and N1BT1D, respectively. Hence, it was assumed that PURA1, PURB1 and PURD1 sequences represent the product from the group-1 chromosome in the tetrasomic status. A PCR-based analysis of nulli-tetrasomic lines (Van Campenhout et al. 1995) was used to obtain the chromosomal location of the wheat sequences. For this purpose, sequence-specific primers P1A5, P1B5,

Table 1 Primers designed during this study, target sequences from which they were derived and chromosomal locations determined

Primer set	Primer sequences	Target sequence	Length of PCR product (bp)	Chromosomal location (locus)
PURCO	PCO5: CTGCCAGCCATGGGAAGCAA PCO3: CTTGCCCTGTGAAATCTCAGAC	pTTH14: Castagnaro et al. (1994)	888–901	Group 1 (<i>Pur-1</i>)
PUR1A	P1A5: GCCTAAGCTGCCCTAAGGA PCO3: CTTGCCCTGTGAAATCTCAGAC	PURA1 PCR product	696	1AL (<i>Pur-A1</i>)
PUR1B	P1B5: GCTCAGAAGTTATGCGCAGG P1B3: TCAGTACGCATCAAGGGATGG	PURB1 PCR product and pTTH1: Castagnaro et al. (1994)	704	1BL ^a (<i>Pur-B1</i>)
PUR1D	P1D5: GCTCAGAAGTTATGCTCAAC PCO3: CTTGCCCTGTGAAATCTCAGAC	PURD1 PCR product and pTTH14: Castagnaro et al. (1994)	751	1DL ^a (<i>Pur-D1</i>)
PUR1R	P1R5: ACTGCTACAACCTTTGCCGT P1R3: ACACAAGCATCACACAGCG	PURR1 PCR product from Petkus (rye)	693	1RL (<i>Pur-R1</i>)

^a The chromosomal location of the PUR1B and PUR1D targets is interchanged compared to the report of Castagnaro et al. (1994) (see text)

		→ SIGNAL PEPTIDE	
PUR1A	1	<u>CTGCCAGCCATGGGAAGCAAGGGCC</u> TC AAGGGTGTGATGGTGTCTTACTCATACT	
PUR1B	1	<u>CTGCCAGCCATGGGAAGCAAGGGCC</u> TC AAGGGTGTGATGGTGTCTTACTCATACT	
PUR1D	1	<u>CTGCCAGCCATGGGAAGCAAGGGCC</u> TC AAGGGTGTGATGGTGTCTTACTCATACT	
PUR1R	1	<u>CTGCCAGCCATGGGAAGCAAGGGCC</u> TC AAGGGTGTGATGGTGTCTTACTCATACT	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
		→ THIONIN	
PUR1A	57	GGGGTGGTTCGGAAACAGTGCAGTAGAAGGCAAGAGTGTCTGCAAGAGCACC	
PUR1B	57	GGGGTGGTTCGGAAACAGTGCAGTAGAAGGCAAGAGTGTCTGCAAGAGCACC	
PUR1D	57	GGGGTGGTTCGGAAACAGTGCAGTAGAAGGCAAGAGTGTCTGCAAGAGCACC	
PUR1R	57	GGGGTGGTTCGGAAACAGTGCAGTAGAAGGCAAGAGTGTCTGCAAGAGCACC	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
PUR1A	113	TaGGAAARACTGCTCAACCTTTGCCCGcCCGTGGTCTCAGAAGTTATGCGCA	
PUR1B	113	TgGAAARACTGCTCAACCTTTGCCCGcCCGTGGTCTCAGAAGTTATGCGCA	
PUR1D	113	TgGAAARACTGCTCAACCTTTGCCCGcCCGTGGTCTCAGAAGTTATGCGCA	
PUR1R	113	TaGGAARACTGCTCAACCTTTGCCCGcCCGTGGTCTCAGAAGTTATGCGCA	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
PUR1A	169	AaCGTtTgAGGGTAAACTCAAGTGGCTTAAGCTGCCCTAAGGaeTTCCTCAA	
PUR1B	169	aaCGTCTGAGGGTAAACTCAAGTGGCTTAAGCTGCCCTAAGGgTTCCTCAA	
PUR1D	169	AaCGTCTGAGGGTAAACTCAAGTGGCTTAAGCTGCCCTAAGGgTTCCTCAA	
PUR1R	169	AaCTCTGTcGGTAAACTCAAGTGGCTTAAGTAgCaAGCTGCCCTAAGGaaTTCCTCAA	
Consensus		+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
		→ ACIDIC PEPTIDE	
PUR1A	225	ATTGGTCTTGGTCCAACTCAGGTAAGAGAAATGAACCTCTTCTATTTTAT	
PUR1B	225	ATTGGCCCTTGGTCCAACTCAGGTAAGAGAAATGAACCTCTTCTATTTTAT	
PUR1D	225	ATTGGCCCTTGGTCCAACTCAGGTAAGAGAAATGAACCTCTTCTATTTTAT	
PUR1R	225	ATTGGCCCTTGGTCCAACTCAGGTAAGAGAAATGAACCTCTTCTATTTTAT	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
		< intron 1	
PUR1A	277	ATACTACCTCTTTTTT...AAAATAGAacAGATGTCATTCTGATAAGCATATAG	
PUR1B	277	ATACTACCTCTTTTTT...AAAATAGAacAGATGTCATTCTGATAAGCATATAG	
PUR1D	277	ATACTACCTCTTTTTT...AAAATAGAacAGATGTCATTCTGATAAGCATATAG	
PUR1R	277	ATACcAACCTCTTTTTT...AAAATAG...ATGTCATTCTAAAGCATATAG	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
PUR1A	329	TCAAGTAAAAATaTCCtGtTTCTTACT...AATATGATgAAGTATTTGAATA	
PUR1B	329	TCAAGTAAAA...TTCTAGTTTCTTACT...AATATGATAacGATTTTGAATg	
PUR1D	332	TCAAGTAAAAATTTCTAGTTTCTTACT...AATATGATAAAGTATTTcaAATA	
PUR1R	325	TCAAGTAAAAATTTCTAGTTTCTTACTctGATATGATAAAGTATTTGAATA	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
PUR1A	382	TGTATAGTGGCCTTAGATGATCAGATTGtCAt.GAATAA.TTTTCCAGgAG.TTT	
PUR1B	381	TGTATAGTGGCCTTAGATGATCAGATTGtCCAt.GgATAA.tTTTCCACAAG.TTT	
PUR1D	385	TGTATAGTGGCCTTAGATGATCAGATTGtCCAt.GAATAA.TTTTCCACAAG.TTT	
PUR1R	381	TGTATAGTGGCCTTAGATGATCAGATTGtCCAtGAtAAATTTTCCACAAGaa.T	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
PUR1A	435	TTTTTTCATAATTTTCTACCAACAAGTAATTG.GAAAATcACTAACAgAACGA	
PUR1B	434	TtCTTCCAcgATTTTCTACCAACAAGTAA.TagGAAAATAACTAACAgAACGA	
PUR1D	439	TTTTTTCATAATTTTCTACCAACAAGTAATTG:GAAAATAACTAACAgAAATGA	
PUR1R	436	TaTTTTTCATAATTTTCTaTCAACAAGTAATTG:GAAAATAACTAACAgAACcA	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
PUR1A	490	ATGATAGAGCCATGTGAATTCAAAACATCAAGCATATTGTAACcGGGGAAATTA	
PUR1B	489	ATGATAGAGCCATGTGAATTCAAAACATCAAGCATATTGTAAGTgAGGAATTA	
PUR1D	495	ATGATAGAGCCATGTGAATTCAAAACATCAAGCATATTGTAACCTGGGAATTA	
PUR1R	492	AaGATAGAGCCAcGTGAATTCAAAACATCAACATATTGTAACCTGGGGAAATTA	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
PUR1A	546	TTTGTTCGAcACAGGATTTTATTACAGCTAGATATTTTAAATTGACAGATGAAC	
PUR1B	538	TTTGTTCGACAGGATTTTATTACAGCTAGATATTTTAAATTGACAGATGAAC	
PUR1D	551	TTTGTTCGAAAGGATTTTATTACAGCTAGATATTTTAAATTGACAGATGAAC	
PUR1R	548	TTTGTTCGACAGGATTTaAtaTAcTcAcTAgATATTTTAAATTGACAGATgAgC	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
		intron 1 >	
PUR1A	602	CAGACCCATGGAGTACGCAACTGGAAATGAGTCTTCCcTATGTGACPAcATG	
PUR1B	594	CAGACCCATGGAGTACGCAACTGGAAATGAGTCTTCCcTATGTGACPAcATG	
PUR1D	607	CAGACCCATGGAGTACGCAACTGGAAATGAGTCTTCCcTATGTGACPAcATG	
PUR1R	604	CAGACCCgTTGAGTATGCAACTGGAAATGAGTCTTCCcTATGTGACPAcATG	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
PUR1A	658	GTCACCGCAGTAAATAAATTCATCTTGATTTGCATTCCTGTAAC...TaAT	
PUR1B	650	GTCACCGCAGTAAAGAAATTCATCTTGATTTGCATTCCTGTAACCTGTAATgAT	
PUR1D	663	GTCACCGCAGTAAAGAAATTCATCTTGATTTGCATTCCTGTAACCTGTAATgAT	
PUR1R	660	GTCAGCGCAGTAAAGAAATTCATCTTGATTTGCATTCCTGTAACCTAAT...*	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
		< intron 2	
PUR1A	710	tgAAAAGGcGATTCcTGACACA...TAaATGcTATTTGGGcACcAGCTGCTGAC	
PUR1B	706	gaAAAAGGcTAGTtTcTGACACA...TAATATGcTATTTGGGcACcAGCTGCTGAC	
PUR1D	719	gaAAAAGGcTAGTTCcTGACACA...TAATATGcTATTTGGGcACcAGCTGCTGAC	
PUR1R	712	tgAAAAGcATAGTTCcTGACAcTcTATATGcTATTTGGGcACcAGCTGCTGAC	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
		intron 2 >	
PUR1A	763	GATGAAGAGATGAAACTCTATGTGAAAcgTGTGGTATGCTTGTGTCAaTTTCTG	
PUR1B	759	GATGAAGAGATGAAACTCTATGTGAAaAtTGTGTGATGCTTGTGTCAaTTTCTG	
PUR1D	772	GATGAAGAGATGAAACTCTATGTGAAaAtTGTGgCgATGCTTGTGTCAaTTTCTG	
PUR1R	768	GAgGAAGAGATGAAACTCTATGTGAAAcgTGTGGTATGCTTGTGTCAgTTTCTG	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
PUR1A	819	TAACGcTGATGCTGGCCtCaCtCCCTTGATGCGTAATGATGCGTATCCATGCTCT	
PUR1B	815	TAACGGTgATGCTGGCCtCcATCCCTTGATGCGTAATGATGCGTgTCCATGCTCT	
PUR1D	828	TAACGGTgATGCTGGCCtCaCtCCCTTGATGCGTAATGATGCGTATCCATGCTCT	
PUR1R	824	TAACGGTgATGCTGGCCtCcGtCCCTTAgTGCcTAATGATGCGTATCCATGCTCT	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	
PUR1A	875	<u>GAGATTTACAGGGCAG</u>	
PUR1B	871	<u>GAGATTTACAGGGCAG</u>	
PUR1D	884	<u>GAGATTTACAGGGCAG</u>	
PUR1R	880	<u>GAGATTTACAGGGCAG</u>	
Consensus		*****+*****+*****+*****+*****+*****+*****+*****+*****+*****+*****	

P1B3 and P1D5 (Table 1), characterized by unique positions at their 3' end (Fig. 1), were constructed. This analysis showed that the primer sets PUR1A, PUR1B and PUR1D (Table 1) amplified respectively 1A, 1B and 1D chromosome-specific products of the predicted size (Fig. 2). Furthermore, ditelosomic analysis confirmed the long arm location of the amplification products (not shown). Hence, it can be concluded that PUR1A, PUR1B and PUR1D lie within the loci *Pur-A1* (on 1AL), *Pur-B1* (on 1BL) and *Pur-D1* (on 1DL) loci, respectively.

These primers were also applied to diploid (*T. monococcum*, *T. urartu*, *T. tauschii*), tetraploid (*T. dicoccoides*, *T. dicoccum*, *T. durum*, *T. timopheevii*) and hexaploid (*T. spelta*) wheat species. As illustrated in Fig. 2, amplification was genome-specific: primer sets PUR1A, PUR1B and PUR1D specifically amplified a product in species with the A, B and D genomes respectively. Since the G genome is known to be related to the B genome (Shands and Kimber 1973), it was not unexpected that the *Pur-B1*-specific primer set amplifies a PCR product in *T. timopheevii* (genomic substitution AAGG).

The specificity (in a wheat background) of the *Pur-R1*-specific primer set (PUR1R, Table 1) was confirmed by an analysis of wheat-rye translocation, substitution and addition lines. A 693-bp product is only formed in genotypes bearing 1R or 1RL (results not shown).

Orthologous variation and interlocus relationships

In Fig. 3, the deduced amino acid sequences of the isolated orthologous sequences comprising the complete thionin precursor are aligned. The PUR1A, PUR1B and PUR1D sequences are identical to previously published sequences for β purothionin (Mak and Jones 1976), α 1-purothionin (Ohtani et al. 1977) and α 2-purothionin (Hase et al. 1978), respectively. PUR1B has a deviant signal peptide and an additional tyrosine residue at the carboxyl terminus. The positions of the eight cysteine residues (forming four disulfide bridges) within the mature protein are completely

Fig. 1 Multiple alignments of *Pur-1* DNA sequences. Isolation was done by PCR using primer set PURCO (see Table 1) on nullitetrasomic 'Chinese Spring' lines and on 'Petkus rye'. These sequences, which encode thionin precursors, consist of three regions: one encoding a signal peptide; the second a mature thionin; and the third, a C-terminal acidic protein. The latter region is interrupted by two introns, as indicated. Primers designed during this study (see Table 1) as well as intron borders are *underlined*. Conserved positions are denoted by *, positions marked by + are at least 50% conserved, and dots represent aligning gaps. The termination triplets are indicated by \blacklozenge . Nucleotide sequences are deposited under EMBL, Genbank and DDBJ accession numbers X96445, X96446, X96448, X96449

conserved, as are the other thionins of type-I found in barley (Lecomte et al. 1982; Rodriguez-Palenzuela et al. 1988). Within the aligned consensus sequence, 27 positions of difference can be found. The mature thionin region is more divergent (24.4%:11 non-conserved positions out of 45) than the contiguous sequences (signal peptide 11.1%; acidic peptide 20.0%). Further inspection of the 27 different positions reveals that the rye sequence is quite distant from the wheat sequences, as it exhibits 11 unique amino acids. The better relationship exists between the sequences from 1B and 1D, which have 13 residues in common (compared to 8 between PURA1-PURB1 and 9 between PURA1-PURD1). This is in accordance with the fact that both these purothionins were originally found to be present in a single α fraction (Jones and Mak 1977).

To analyse the inter-locus relationships in more detail, we calculated percentages of DNA sequence similarities for pair-wise comparisons (Table 2). Analysis of the exon DNA sequences led to the same conclusions as the comparison of the protein sequences: *Pur-R1* is the most distant sequence and *Pur-B1* and *Pur-D1* show the highest similarity percentage score (93.2%) among the pairwise comparisons. When the entire sequences of

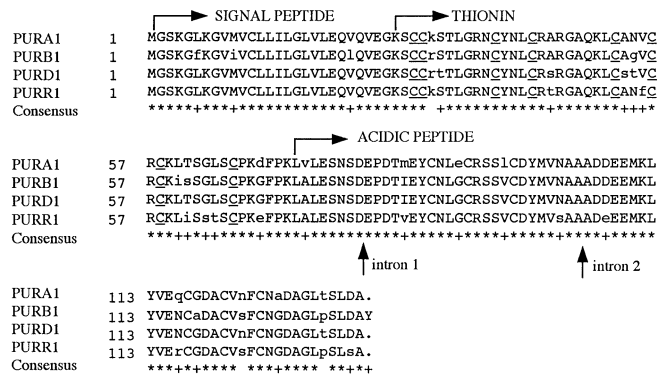


Fig. 3 Comparisons of amino acid sequences deduced from the corresponding DNA sequences (see Fig. 1). The signal peptide, mature thionin and acidic peptide, together constituting the thionin precursor, as well as the positions of the introns are indicated. The eight conserved cysteine residues within the thionins are underlined. Positions marked * are identical, + are at least 50% conserved, · is beyond the coding sequence

Table 2 Pairwise DNA sequence similarities of the orthologous *Pur-1* loci

	<i>Pur-A1</i>	<i>Pur-B1</i>	<i>Pur-D1</i>	<i>Pur-R1</i>
<i>Pur-A1</i>		89.88	90.50	86.36
<i>Pur-B1</i>	91.30		90.80	84.51
<i>Pur-D1</i>	92.59	93.24		86.15
<i>Pur-R1</i>	89.05	87.76	89.37	

Similarity percentages were calculated using the pairwise alignment programme of the GENMON software package. Results of the comparison of the protein-coding regions are presented below the diagonal, and comparisons of the whole PCR products are presented above the diagonal

the PCR products are compared, all the similarity scores show somewhat smaller values, which is due to the fact that there is less sequence homology between the introns, especially the larger one (Fig. 1). The inter-locus relationships stay broadly the same, although the degree of similarity between *Pur-B1* and *Pur-D1* is less pronounced, 90.8%, which is barely higher than that between *Pur-A1* and *Pur-D1*, 90.5%.

Discussion

Strategy

This work is instrumental to the development of chromosome-specific PCR markers for the homoeologous group-1 chromosomes of wheat and rye in our laboratory. One approach is the development of sequence-tagged-sites, i.e. PCR markers from known DNA sequences that map specifically onto the chromosomes

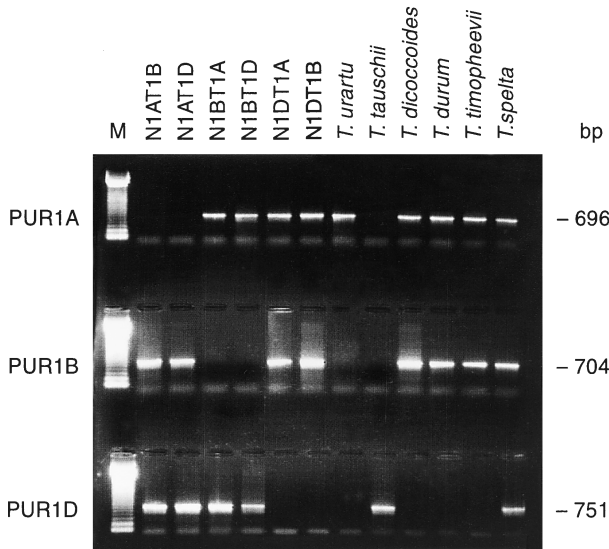


Fig. 2 PCR-based chromosome assignment of wheat thionin sequences PURA1, PURB1 and PURD1. Amplification was done with *Pur-A1*-, *Pur-B1*- and *Pur-D1*-specific primer sets PUR1A (top), PUR1B (middle) and PUR1D (bottom) on nulli-tetrasomic stocks. The specific absence of a PCR product in a line in which a particular chromosome-pair is absent indicates the chromosomal location of the product. The size of the amplification products is shown. DNAs of hexaploid wheat cv ‘Chinese Spring’ (CS) and its derived group-1 aneuploid stocks were used as indicated. In addition, the genome specificity of the primer sets is illustrated by their amplification pattern in *T. urartu* (AA), *T. tauschii* (DD), *T. dicoccoides* (AABB), *T. durum* (AABB), *T. timopheevii* (AAGG) and *T. spelta* (AABBDD). M 123-bp size marker

of interest. This strategy is hampered in bread wheat by its hexaploid character, consisting of three related genomes, A, B and D, that have originated from wild diploid species that have given rise to tetraploid and hexaploid wheats through natural hybridisation. In spite of this, we have previously developed locus-specific primer sets for genes belonging to the low-molecular-weight (LMW) glutenin gene family on each of the group-1 chromosomes of bread wheat (Van Campenhout et al. 1995). In this study we showed that this strategy is also appropriate for type-I thionins, an orthologous set of single-copy genes on the group-1 chromosomes. Sufficient orthologous variation was found to allow for the construction of locus-specific primer sets for each of the *Pur-1* loci of interest. The value of these primer sets as chromosome-specific markers for genotype selection and characterisation was further investigated by the analysis of wheat-rye recombinants (Van Campenhout et al. manuscript in preparation).

Genomic DNA sequences for each of the *Pur-1* loci were efficiently isolated using consensus PCR and subsequent blunt-end cloning, a straightforward method requiring minimal workload. This strategy could be used to search for new thionin variants in cereals. Chromosome assignment in wheat by PCR has been reported by D'Ovidio et al. (1995), Röder et al. (1995) and Van Campenhout et al. (1995) and is rapid, convenient and easy. Castagnaro et al. (1994) analysed two cDNA clones for type-I thionins and assigned pTTH14 as $\alpha 1$ on 1B and pTTH1 as $\alpha 2$ on 1D. When this is compared with our results, this clearly must be an error. First, the specificity of the 1B-specific sequence was tested on aneuploids (Fig. 2) as well as on 1BL.1RS and 1BS.1RL translocation lines (results not shown). Secondly, the PURB1 and PURD1 sequences encode for the $\alpha 1$ - and $\alpha 2$ -purothionins, respectively, the genes of which are known to be located on 1B and 1D, respectively (Jones and Mak 1977). This means that pTTH14 is in fact $\alpha 2$ -purothionin and pTTH1, $\alpha 1$ -purothionin.

Genome relationships

Analogous to what is shown in this work on *Pur-1*, interlocus relationships for a broad set of loci could be determined. In this way, we could gain more insight into the relationships between the A, B and D genomes within hexaploid wheat. The interlocus relationships, derived from the comparisons of the different *Pur-1* DNA sequences within 'Chinese Spring' (Table 2), showed that the locus on 1D is most closely related to the locus on 1B (*Pur-B1*). The 1A-specific sequence is more related to the 1D-specific sequence than to the 1B sequence.

Also, relationships between genomes within the genus *Triticum* can be investigated by applying locus-

specific primer sets to the species of interest. Since in this work the *Pur-1*-specific primer sets amplified a PCR product in the relevant diploid and tetraploid wheats, they can be considered to be genome-specific. The fact that the *Pur-B1*-specific primer set also amplifies a PCR product in *T. timopheevii* gives support to earlier reports about the close relationship of the B and G genomes (Shands and Kimber 1973; Jaaska 1978; Tsunewaki and Ogihara 1983; Talbert et al. 1991).

The amino acid sequences of purothionins from some diploid and tetraploid *Triticum* species have been reported, and these data have been used to gain more insight in wheat phylogeny. *T. monococcum* and *T. urartu* have a purothionin identical to the β -purothionin form specified by the *Pur-A1* locus of *T. durum* and *T. aestivum* (Jones et al. 1982; Jones and Mak 1983; Kerby et al. 1988). Therefore, Kerby et al. (1988) concluded that both species possess the A genome and are possible donors of this genome to polyploid wheat. In this work, primer set PUR1A directed to the *Pur-A1* locus of *T. aestivum* amplified a product in *T. monococcum*, *T. urartu* and other species having the A genome in common as well. Jones et al. (1982) found that the gene coding for $\alpha 2$ -purothionin of *T. aestivum* is contained in the D genome. Hence, these researchers concluded that the purothionin of *T. tauschii*, the donor of the D genome to *T. aestivum*, must be very homologous to the $\alpha 2$ -purothionin of *T. aestivum*. DNA sequencing of the PCR product amplified with primer set PUR1D from *T. tauschii* could reveal the extent of this homology. Kerby et al. (1990) determined the amino acid sequences of the purothionins from five putative B genome donors from the section *Sitopsis*. Although three of them had sequences very similar to the $\alpha 1$ -purothionin form specified by the *Pur-B1* locus of *T. durum* and *T. aestivum* (Jones et al. 1982), none of them was identical to it. As a result of the close taxonomic relationship of the species of the *Sitopsis* section, other studies have also implicated more than one species in this section as being the likely donor of the B set of chromosomes (Talbert et al. 1991). So far, however, the progenitor of the B genome has remained unidentified. A PCR-based approach, using genome-specific primer sets would make it possible to collectively analyse and thereby possibly identify all of the putative B-genome donors. Alleles from the putative B-genome donors could then be compared to those at *Pur-B1* of hexaploid wheat. This may shed new light on the origin of the B genome in polyploid wheat.

Acknowledgements We thank Geert De Wever, Peter Langridge and Stephen Reader for gifts of plant material and Robert Koebner for helpful comments on the manuscript. S. Van Campenhout holds a predoctoral fellowship from the Flemish IWT (Vlaams Instituut voor de Bevordering van het Wetenschappelijk-Technologisch Onderzoek in de Industrie). This work was financed in part by the EU Biotechnology programme Area III.1 "Conservation of Genetic Resources".

References

- Balls AK, Hale WS, Harris TH (1942) A crystalline protein obtained from a lipoprotein of wheat flour. *Cereal Chem* 19: 279–288
- Bohlmann H, Apel K (1987) Isolation and characterization of cDNAs coding for leaf-specific thionins closely related to the endosperm-specific hordothionin of barley (*Hordeum vulgare* L.). *Mol Gen Genet* 201: 446–454
- Bohlmann H, Clausen S, Behnke S, Giese H, Hiller C, Reimann-Philipp U, Schrader G, Barkholt V, Apel K (1988) Leaf-specific thionins of barley: a novel class of cell wall proteins toxic to plant pathogenic fungi and possibly involved in the defense mechanism of plants. *EMBO J* 7: 1559–1565
- Carbonero P, Garcia-Olmedo F (1969) Purothionins in *Aegilops-Triticum* spp. *Experientia* 25: 1110
- Castagnaro A, Marañá C, Carbonero P, Garcia-Olmedo F (1992) Extreme divergence of a novel wheat thionin generated by a mutational burst specifically affecting the mature protein domain of the precursor. *J Mol Biol* 224: 1003–1009
- Castagnaro A, Marañá C, Carbonero P, Garcia-Olmedo F (1994) cDNA cloning and nucleotide sequence of $\alpha 1$ and $\alpha 2$ thionins from hexaploid wheat endosperm. *Plant Physiol* 106: 1221–1222
- Castagnaro A, Segura A, Garcia-Olmedo F (1995) High conservation among sequences encoding type-V thionins in wheat and *Aegilops*. *Plant Physiol* 107: 1475–1476
- D'Ovidio R, Masci S, Porceddu E (1995) Development of a set of oligonucleotide primers specific for genes at the *Glu-1* complex loci of wheat. *Theor Appl Genet* 91: 189–194
- Ebrahim-Nesbat F, Behnke S, Kleinhofs A, Apel K (1989) Cultivar-related differences in the distribution of cell-wall-bound thionins in compatible and incompatible interactions between barley and powdery mildew. *Planta* 179: 203–210
- Fernández de Caleyá R, González-Pascual B, Garcia-Olmedo F, Carbonero P (1972) Susceptibility of phytopathogenic bacteria to wheat purothionins in vitro. *Appl Microbiol* 23: 998–1000
- Fernández de Caleyá R, Hernández-Lucas C, Carbonero P, Garcia-Olmedo F (1976) Gene expression in allopolyploids: genetic control of lipopurothionins in wheat. *Genetics* 83: 687–699
- García-Olmedo F, Rodríguez-Palenzuela P, Hernández-Lucas C, Ponz F, Marañá C, Carmona MJ, López-Fando J, Fernández JA, Carbonero P (1989) The thionins: a protein family that includes purothionins, viscotoxins and crambins. In: Milfin B (ed) *Oxford surveys of plant molecular and cell biology*, vol 6. Oxford University Press, London, pp 31–60
- Gausing K (1987) Thionin genes specifically expressed in barley leaves. *Planta* 171: 241–246
- Hanley BA, Schuler MA (1988) Plant intron sequences: evidence for distinct groups of introns. *Nucleic Acids Res* 16: 7159–7176
- Hase T, Matsubara H, Yoshizumi H (1978) Disulfide bonds of purothionin, a lethal toxin for yeasts. *J Biochem* 83: 1671–1678
- Hernández-Lucas C, Fernández de Caleyá R, Carbonero P (1974) Inhibition of brewer's yeasts by wheat purothionins. *Appl Microbiol* 28: 165–168
- Hernández-Lucas C, Carbonero P, Garcia-Olmedo F (1978) Identification and purification of a purothionin homologue from rye (*Secale cereale* L.). *J Agric Food Chem* 26: 794–798
- Hernández-Lucas C, Royo J, Paz-Ares J, Ponz F, Garcia-Olmedo F, Carbonero P (1986) Polyadenylation site heterogeneity in mRNA encoding the precursor of the barley toxin α -hordothionin. *FEBS Lett* 200: 103–106
- Jaaska V (1978) NADP-dependent aromatic alcohol dehydrogenase in polyploid wheats and their diploid relatives. On the origin and phylogeny of polyploid wheats. *Theor Appl Genet* 53: 209–217
- Jones BL, Mak AS (1977) Amino acid sequences of the two purothionins of hexaploid wheat. *Cereal Chem* 54: 511–523
- Jones BL, Mak AS (1983) The amino acid sequence of a purothionin from *Triticum monococcum*, a diploid wheat. *Eur J Biochem* 133: 621–625
- Jones BL, Lookhart GL, Mak A, Cooper DB (1982) Sequences of purothionins and their inheritance in diploid, tetraploid, and hexaploid wheats. *J Hered* 73: 143–144
- Kerby K, Kuspira J, Jones BL (1988) Biochemical data bearing on the relationship between the genome of *Triticum urartu* and the A and B genomes of the polyploid wheats. *Genome* 30: 576–581
- Kerby K, Kuspira J, Jones BL, Lookhart GL (1990) Biochemical data bearing on the origin of the B genome in the polyploid wheats. *Genome* 33: 360–368
- Lecomte JTL, Jones BL, Llinás M (1982) Proton magnetic resonance studies of barley and wheat thionins: structural homology with crambin. *Biochemistry* 21: 4843–4849
- Mak AS, Jones BL (1976) The amino acid sequence of wheat β -purothionin. *Can J Biochem* 54: 835–842
- Ohtani S, Okada T, Yoshizumi H, Kagamiyama H (1977) Complete primary structures of two subunits of purothionin A, a lethal protein for brewer's yeast from wheat flour. *J Biochem* 82: 753–767
- Ponz F, Paz-Ares J, Hernández-Lucas C, Carbonero P, Garcia-Olmedo F (1983) Synthesis and processing of thionin precursors in developing endosperm from barley (*Hordeum vulgare* L.). *EMBO J* 2: 1035–1040
- Ponz F, Paz-Ares J, Hernández-Lucas C, Garcia-Olmedo F, Carbonero P (1986) Cloning and nucleotide sequence of a cDNA encoding the precursor of the barley toxin α -hordothionin. *Eur J Biochem* 156: 131–135
- Redman DG, Fisher N (1968) Fractionation and comparison of purothionin and globulin components of wheat. *J Sci Food Agric* 19: 651–655
- Redman DG, Fisher N (1969) Purothionin analogues from barley flour. *J Sci Food Agric* 20: 427–432
- Röder MS, Plaschke J, König SU, Börner A, Sorrells ME, Tanksley SD, Ganai MW (1995) Abundance, variability and chromosomal location of microsatellites in wheat. *Mol Gen Genet* 246: 327–333
- Rodríguez-Palenzuela P, Pintor-Toro JA, Carbonero P, Garcia-Olmedo F (1988) Nucleotide sequence and endosperm-specific expression of the structural gene for the toxin α -hordothionin in barley (*Hordeum vulgare* L.). *Gene* 70: 271–281
- Sánchez-Monge R, Delibes A, Hernández-Lucas C, Carbonero P, Garcia-Olmedo F (1979) Homoeologous chromosomal location of the genes encoding thionins in wheat and rye. *Theor Appl Genet* 54: 61–63
- Sears ER (1966) Nullisomic-tetrasomic combinations in hexaploid wheat. In: Riley R, Lewis KR (eds) *Chromosome manipulation and plant genetics*. Oliver and Boyd, London, pp 29–45
- Sears ER, Sears LMS (1978) The telocentric chromosomes of common wheats. In: Ramanujam S (ed) *Proc 5th Int Wheat Genet Symp. Kapoor Art Press, New Delhi*, pp 389–407
- Shands H, Kimber G (1973) Reallocation of the genomes of *Triticum timopheevii* Zhuk. In: Sears ER, Sears LMS (eds) *Proc 4th Int Wheat Genet Symp. University of Missouri, Columbia*, pp 101–108
- Stuart LS, Harris TH (1942) Bactericidal and fungicidal properties of a crystalline protein isolated from unbleached wheat flour. *Cereal Chem* 19: 288–300
- Talbert LE, Magyar GM, Lavin M, Blake TK, Moylan SL (1991) Molecular evidence for the origin of the S-derived genomes of polyploid *Triticum* species. *Am J Bot* 78: 340–349
- Tsunewaki K, Ogihara Y (1983) The molecular basis of genetic diversity among cytoplasm of *Triticum* and *Aegilops* species. II. On the origin of polyploid wheat cytoplasm as suggested by chloroplast DNA restriction fragment patterns. *Genetics* 104: 155–171
- Van Campenhout S, Vander Stappen J, Sági 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
- Verhasselt P, Van Campenhout S, Voet M, Wellens B, Volckaert G (1993) A general method for routine sequencing of cloned DNA fragments using commercial dye primers and its application in sequence analysis of *Toxoplasma gondii* RH genome fragments. *J DNA Sequencing Mapping* 4: 71–77