Plant Molecular Biology **18**: 515–525, 1992.

The lectin gene family of *Ricinus communis*: Cloning of a functional ricin gene and three lectin pseudogenes

James W. Tregear[1] and Lynne M. Roberts*
*Department of Biological Sciences, University of Warwick, Coventry, CV4 7AL, UK; (*author for correspondence); [1]Present address: AFRC Institute of Arable Crops Research, Long Ashton Research Station, Bristol, BS18 9AF, UK*

Received 26 March 1991; accepted in revised form 10 October 1991

Key words: lectin genes, pseudogenes, ricin, *Ricinus communis*

Abstract

Molecular hybridisation using a ricin cDNA probe has revealed that the ricin/*Ricinus communis* agglutinin (RCA) multigene family is composed of approximately eight members. Several genomic clones containing preproricin and preproricin-like sequences have been isolated. Partial analysis of three different genomic clones by DNA sequencing and ribonuclease protection has indicated that at least three members of the lectin gene family are non-functional. None of the original seventeen positive clones isolated appears to contain a *Ricinus communis* agglutinin (RCA) gene. One gene member analysed (pCBG3H1) represents a functional ricin gene similar in coding sequence to the published cDNA sequence and possesses typical eukaryotic consensus sequences and seed-specific elements within the flanking sequences. Investigation at the transcriptional level of the expression pattern of this gene revealed that mRNA accumulates during the post-testa stages of seed development. The pattern of accumulation of steady-state transcripts correlates closely with that previously observed at the protein and translatable RNA levels.

Introduction

Lectins are proteins which specifically bind carbohydrates and frequently agglutinate cells. In certain plant species the sugar binding polypeptide is synthesized as part of a larger precursor which also contains a sequence possessing ribosomal inactivating ability [21]. The proteolytic processing of such precursors during synthesis and vesicular transport in the plant cell results in the maturation and deposition of potent eukaryotic cytotoxins [27]. Ricin (from *Ricinus communis*) is perhaps the best known example of such toxic lectins.

Both ricin and its homologue, *Ricinus communis* agglutinin (RCA), are found within protein body organelles in the seeds of the castor oil plant (*R. communis*). Mature ricin is a dimeric glycoprotein (M_r 65000) composed of a toxic A chain (M_r 31000) disulphide linked to a sugar-binding

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X52908.

B chain (M_r 34000). When presented, it can bind and enter most types of eukaryotic cell by virtue of the B chain's ability to interact with cell surface galactosides. Upon endocytosis of the toxin, a proportion of the A chain is able to translocate an internal membrane and kill the cell by enzymatically depurinating a specific base in the 28S rRNA of eukaryotic ribosomes [5]. Structurally, mature RCA is composed of two non-covalently associated ricin-like heterodimers to give a tetramer which, on comparison with ricin, has different *in vivo* toxicity and agglutination characteristics and slightly differing sugar-binding specificities and *in vitro* toxicities [25]. Despite these differences, cDNA clones encoding preproricin [18] and preproagglutinin [28] have revealed extensive homologies, the A chains being 93% identical at the amino acid level whilst the corresponding B chains are 84% homologous.

It appears from protein analysis that ricin and RCA are coordinately synthesized in the castor bean endosperm tissue [29] (along with other storage protein types) and reach a peak representing 5% of the total particulate seed protein after seed testa formation. Expression also appears to be strictly tissue-specific, there being no evidence for the occurrence of ricin in the other tissues of the castor oil plant. Why these proteins are produced at such high levels within the endosperm remains a mystery, although it is generally believed that they may have a defence role against predators.

The availability of cDNA clones has allowed the isolation of genomic ricin sequences as a prelude to investigating further the developmental and tissue-specific control of ricin accumulation. We report here the nucleotide sequence of a functional member of the ricin gene family and partial analysis of other non-functional members.

Materials and methods

Plant material

Ricinus communis (castor oil) seeds of Sudanese origin were kindly provided by Croda Premier Oils, Hull.

Isolation of nuclear DNA from Ricinus communis

R. communis nuclear DNA was prepared in part by methods previously described [8, 15]. The cotyledons of germinating seeds were homogenised in EB + extraction buffer (0.44 M sucrose, 2.5% (w/v) Ficoll 400, 0.4% (w/v) Dextran T40, 25 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 0.5% Triton X-100, 10 mM 2-mercaptoethanol, 0.04% (w/v) ethidium bromide), the homogenate filtered and nuclei collected by low-speed centrifugation. The resulting pellets were resuspended in EB – (EB + buffer lacking ethidium bromide) and loaded onto step gradients containing 85% (w/v) sucrose, 80% (w/v) Percoll, 60% (w/v) Percoll and 40% (w/v) Percoll. Percoll solutions were prepared in a buffer containing 0.44 M sucrose, 25 mM Tris-HCl pH 7.6 and 10 mM $MgCl_2$. Nuclei were collected from the 85% sucrose/80% Percoll interface and washed in EB – . Lysis was achieved by resuspending the nuclei in 2 ml of 0.2 M sucrose, 10 mM Tris-HCl pH 7.4, 2 mM $CaCl_2$ and adding NaCl to 2.4 M ethidium bromide to 50 μM, EDTA to 2 mM, SDS to 0.1% (w/v), proteinase K to 0.2 mg/ml and 2-mercaptoethanol to 2.5% (v/v). The preparation was incubated for 2 h at 50 °C, after which it was dialysed against 20 mM Tris-HCl pH 8.5, 0.1 M NaCl, 1 mM EDTA. The preparation was then centrifuged at 4000 × *g* for 5 min to remove any insoluble polysaccharide, extracted four times with phenol/chloroform and purified by CsCl equilibrium gradient centrifugation [22].

Preparation of ^{32}P-radiolabelled DNA probes

Southern blots, plaque lifts and northern blots were probed with a 1.6 kb *Pst* I fragment from the ricin cDNA clone pRCL617 [18] labelled either by nick translation [26] or by the random priming method [6].

Nucleic acid hybridisation

20 μg aliquots of nuclear DNA were cleaved with excess restriction enzyme and the resulting fragments were separated on 0.6% (w/v) agarose gels

[22]. Southern transfer and hybridisation were carried out using Hybond-N nylon membrane according to the manufacturers' instructions. Following hybridisation, the membranes were washed for 15 min at room temperature in 2 × SSC, 0.1% SDS and for 1 h at 60 °C in 1 × SSC. Electrophoresis of total RNA was performed as previously described [23]. Northern transfer and hybridisation was carried out using Hybond-N nylon membrane according to the manufacturers' instructions. Following hybridisation, the membranes were washed at the same stringency as the genomic Southern blots.

Construction and screening of genomic libraries

Two *R. communis* genomic libraries were prepared in the bacteriophage lambda vector charon 35 [19]; one from partial *Sau* 3A digests and one from partial *Eco* RI digests. Standard construction and screening methods were followed in both cases [22]. Bacteriophage lambda *in vitro* packaging extracts were obtained from Promega Biotech and used according to the manufacturers' recommendations. A total of 10^6 *in vitro* packaged bacteriophage were screened directly without prior amplification using the *Escherichia coli* strain K803 (*sup*E *hsd*R gal *met*B) as host. Following plaque hybridisation, the filters were washed at the same stringency as the genomic Southern blots. Plaque purification of positive clones was carried out using *E. coli* strain DH1 (*sup*E44 *hsd*R17 *rec*A1 *end*A1 *gyr*A96 *thi*-1 *rel*A1).

Subcloning and sequence analysis of putative lectin clones

DNA fragments were subcloned from lectin-positive lambda clones into the plasmid pUC8 [31] and thence into M13 vectors [24] using conventional cloning techniques [22]. DNA sequencing was carried out with the chain termination method [30], as modified in Biggin *et al.* [2].

Isolation of total RNA

Total RNA was isolated from the endosperm of developing castor oil seeds as described previously [18]. An alternative method [20] was used for the extraction of total RNA from leaf and root tissues.

Ribonuclease protection analysis of transcripts

[α^{32}P]-radiolabelled RNA probes were prepared by *in vitro* transcription from constructs containing DNA fragments of interest cloned into the plasmids pGEM-blue3 and pGEM-blue4 (Promega Biotech). The transcription and ribonuclease protection methods used were as described previously [16].

Results and discussion

Estimated size of the Ricinus communis *lectin gene family*

Figure 1. shows an autoradiograph of a genomic Southern blot probed with a ricin cDNA fragment and washed at low stringency. Since only one of the lectin genomic clones characterised in this study contains a *Hin*d III site, it is assumed that the number of hybridizing fragments in the *Hin*d III digest approximately represents the size of the *R. communis* lectin gene family. Thus, the gene family appears to consist of about 8 members. The varying intensities of the bands seen on the autoradiograph are probably due to the occurrence of differential hybridisation, despite the low stringency used. A somewhat different pattern of relative intensities is seen if an RCA I cDNA fragment is used as the probe (data not shown).

Restriction site diversity of lectin-positive genomic clones

A total of 17 lectin-positive bacteriophage lambda clones were identified and plaque-purified from

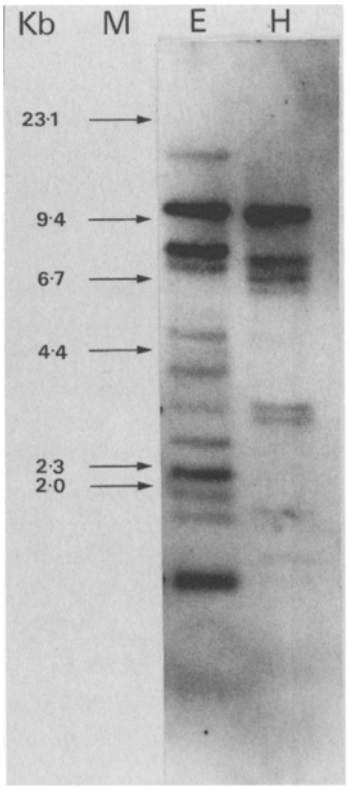


Fig. 1. Estimation of the castor bean lectin gene family size. 10 μg samples of high-molecular-weight castor bean nuclear DNA were digested with either *Eco* RI or *Hin*d III and the resulting fragments electrophoresed on a 0.6% agarose gel and Southern-blotted onto a nylon membrane. The membrane was probed with a radiolabelled ricin cDNA fragment and washed at low stringency. In track M are shown the positions of the lambda DNA size markers. Tracks E and H show the patterns of hybridisation obtained with the *Eco* RI and *Hin*d III digests respectively.

the castor bean genomic libraries constructed; one from the partial *Eco* RI library (lambda clone 1) and 16 from the partial *Sau* 3A library (lambda clones 2–17). Southern blots were performed on *Hin*d III and *Eco* RI digests of DNA from each clone in order to classify the clones into subgroups. Four different positive clones were chosen for further analysis, namely clones 1, 3, 7 and 10. DNA fragments containing sequences homologous to the ricin cDNA were subcloned into pUC8 and partial restriction maps constructed. The data obtained are summarised in Fig. 2. The putative lectin gene in clone 1 was subcloned as two *Eco* RI fragments. It was later shown that a short region of 344 bp, represented in Fig. 2 as a dotted line, is not present in either of the two fragments subcloned, although it presumably links the two in clone 1 and therefore in the *R. communis* genome. The restriction maps of clones 1, 3 and 7 (the latter subcloned to give 7A and 7B) do not closely resemble any of the cDNA or genomic clones previously described, (clones 7A and 7B represent different but closely linked genes or pseudogenes). Only clone 10 appears to match that of the ricin cDNA exactly [18], and none match or resemble the RCA cDNA [28]. Little more could be deduced as to which clones represent ricin-like genes and which clones are RCA-I-like.

Features of the lectin genes

DNA sequencing was performed on four of the five putative lectin genes subcloned, namely those from clones 1, 3, 7A and 10. Although the lectin gene from clone 10 (subsequently known as pCBG3H1) was sequenced in its entirety, only partial sequences were obtained for the other clones.

The sequence data obtained from clones 1 and 3 revealed the presence of three and two frameshift mutations respectively, indicating that both must be pseudogenes (data not shown). Similarly, no *in vivo* transcripts homologous to the putative lectin gene in clone 7A could be identified by RNase protection. Thus, at least three members of the castor bean lectin gene family appear to be non-functional. Other lectin gene families are known to contain structural pseudogenes of this type [13, 32].

In contrast however, the nucleotide sequence of pCBG3H1 (Fig. 3) revealed an apparently functional ricin gene. The lectin sequence shares a high degree of homology with the pAKG ricin genomic clone described by Halling *et al.* [10]. The two genes encode an identical toxic lectin precursor consisting of a 35 amino acid residue presequence, a 267 amino acid residue A chain, a 12 amino acid linker peptide and a 262 amino

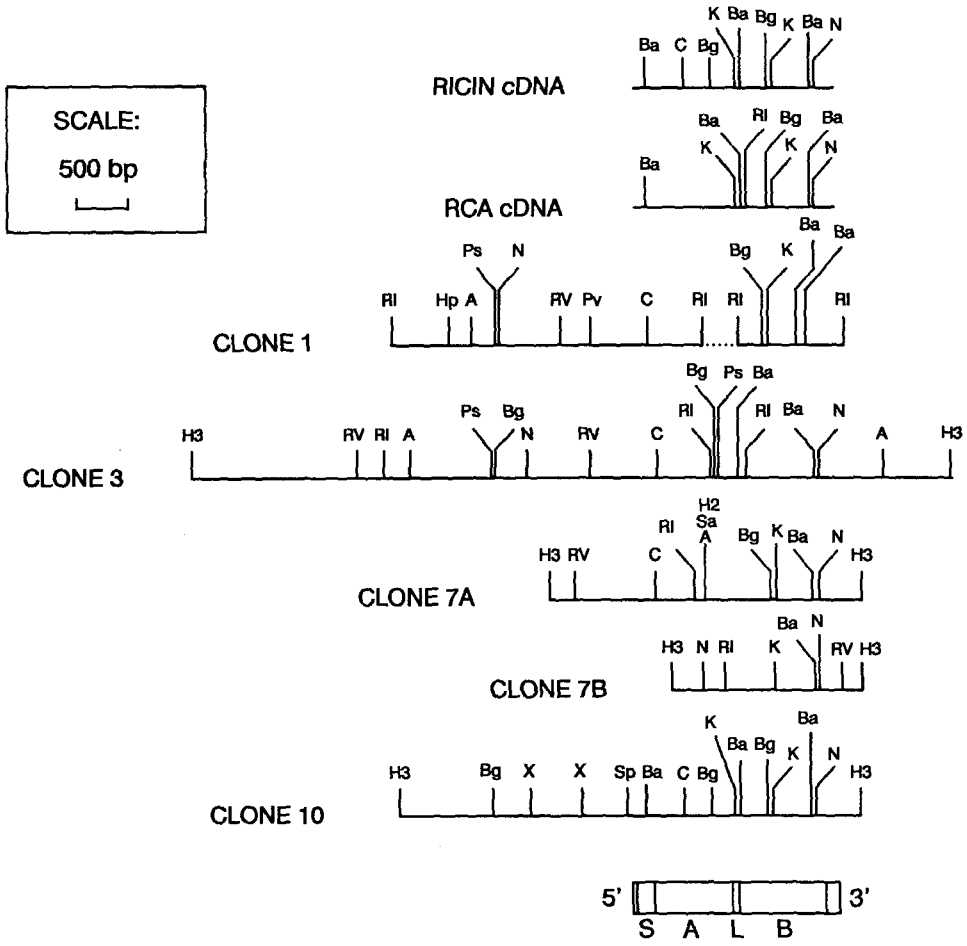
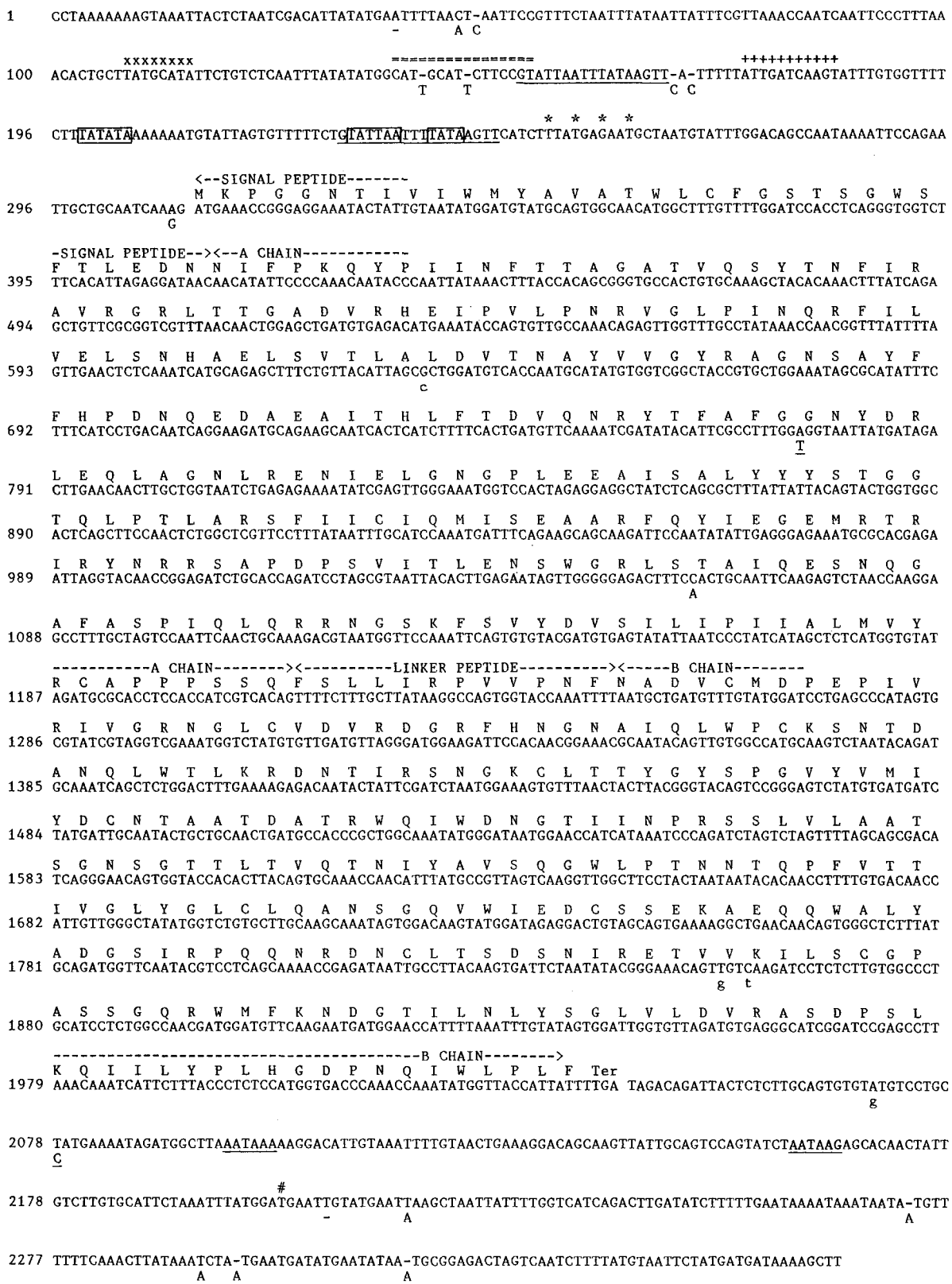


Fig. 2. Restriction maps of lectin subclones. Restriction sites are abbreviated as follows: A, *Acc* I; Ba, *Bam* HI; Bg, *Bgl* II; C, *Cla* I; RI, *Eco* RI; RV, *Eco* RV; H2, *Hinc* II; H3, *Hind* III; Hp, Hpa I; K, *Kpn* I; N, *Nco* I; Ps, *Pst* I; Pv, *Pvu* II; Sa, *Sal* I; Sp, *Sph* I; X, *Xba* I. Beneath the restriction maps is a scale diagram showing the positions of the signal sequence (S), A chain sequence (A), linker sequence (L), and B chain sequence (B). The extent of the 5 and 3′ untranslated regions of the clone 10 (pCBG3H1) ricin gene is also shown.

acid residue B chain. Within the coding region, the pCBG3H1 sequence differs at only two nucleotide positions from that of pAKG and at only 4 positions from that of the ricin cDNA clone described previously from our own laboratory [18]. Thus the polypeptide encoded must be of the ricin D type [7]. As with the pAKG genomic clone, no introns are present. This appears to be a typical feature of plant lectin genes [e.g. 32, 13]. The nucleotide sequence shown in Fig. 3 extends 310 bp upstream from the translation initiation codon and 321 bp downstream from the translation termination codon.

It is likely that the ricin gene cloned here uses more than one transcription start site; the most 5′ position determined (see below) occurring 60 bp upstream from the ATG codon. As with the coding region, the pCBG3H1 promoter and 5′ untranslated region sequences (Fig. 3) share a high degree of homology with the corresponding regions in the pAKG clone [10]. Within the region for which sequence data of both clones are available, the pCBG3H1 sequence shows one insertion of 1 bp, five 1 bp deletions and two 1 bp substitutions with respect to the pAKG sequence.

A number of sequence elements within the 5′



flanking region are potentially involved in the transcription of the gene. The location of the TATA box associated with the most 5′ transcription start point is not entirely clear since the two TATA sequences which might perform a TATA box function occur further upstream at 13 bp and 52 bp respectively. Both are apparently outside the typical range of distances normally seen. There is however a more appropriately situated TATTAA sequence at −22 (with respect to the most 5′ putative transcription start point), which might function as a TATA box. The major transcription start point of the pCBG3HI ricin gene occurs 9 bp downstream from the start point mentioned above and is presumably associated with the TATAAG sequence 22 bp upstream from it. The latter sequence is also found at a similar position in the promoter of the *Ricinus communis* 2S seed albumin gene [12]. Although the pCBG3H1 gene does not contain a canonical CAAT or AGGA box sequence, it does contain a CAAGT element also present in the pAKG clone. It was suggested that the pAKG clone CAAGT element may represent a CAAT box analogue, although there is no functional evidence to support this as yet. Immediately upstream from this sequence is an ATTGA motif which might represent an AGGA box analogue, since it consists of a trinucleotide which matches the consensus sequence (G/T)NG and is flanked by adenine residues. However, the homology is limited and, once again, there is no evidence to support its functional role as an expression signal.

A further element associated with a number of seed-specific genes can also be identified. Two separate motifs resembling the CATGCATG RY repeat described previously [4] are located upstream, 113 bp and 143 bp respectively from the putative transcription start site. The more 5′ motif contains two mismatches whilst the more 3′ motif contains only one. Interestingly, the more 3′ motif occurs within a 14 bp region which closely matches a sequence element in the promoter of the *R. communis* 2S albumin gene [12], which is expressed in an endosperm-specific pattern similar to that of the lectin genes. The RY repeat motif in this region is not present at the corresponding position in the promoter of the pAKG ricin gene [10].

In common with the 5′ untranslated and flanking regions, the 3′ untranslated and flanking regions of the pCBG3H1 clone share a high degree of homology with their counterparts in the pAKG genomic clone. Within the 158 bp of 3′ untranslated region and 163 bp of 3′ flanking region sequenced, a total of three 1 bp deletions, one 1 bp insertion and three 1 bp substitutions were found. The polyadenylation site position shown in Fig. 3 was predicted by RNase protection analysis (data not shown). Although several other potential sites were identified, it is not clear whether these are attributable to other members of the gene family, or are artefactual. Two putative polyadenylation signals, with sequences AATAAA and AATAAG, have been previously identified in the 3′ untranslated region of a ricin cDNA [18] and a ricin genomic clone [10]. These sequences are also conserved in the pCBG3H1 clone where they occur 102 and 49 nucleotides upstream from the putative poly(A) site respectively. From the

Fig. 3. Nucleotide and deduced primary sequence of pCBG3H1 (lambda clone 10) ricin gene. The pCBG3H1 nucleotide sequence is compared with the corresponding sequences of the pRCL617 ricin cDNA [18] and pAKG ricin genomic [10] clones previously described. Capital and small letters underneath the pCBG3H1 sequence indicate differing nucleotide residues in the pAKG and pRCL617 sequences respectively, whilst underlined capitals indicate differing nucleotide residues common to both the pAKG and pRCL617 clones. Nucleotides absent from either the pCBG3H1 or the pAKG sequences are indicated by dashes on the appropriate line. No insertions or deletions occur between the pCBG3H1 and pRCL617 sequences. The putative transcription start and polyadenylation sites of the pCBG3H1 gene are represented by the symbols* and # respectively. In the 5′ flanking region, two TATA sequences plus a TATTAA sequence which might be involved in transcription initiation are shown within boxes. The position of a putative CAAT/AGGA box analogue is indicated by the symbol +. The underlined regions show the respective positions of two 18 bp sequences which match almost perfectly. A 14 bp region which is highly homologous to a sequence in the promoter region of a castor bean 2S albumin gene is indicated by the symbol =. This region contains an RY repeat-like element, another example of which occurs further upstream, shown by the symbol ×. In the 3′ untranslated region, the two putative polyadenylation signals are underlined.

sequences of various ricin cDNA clones [18 and unpublished observations], it has been shown that different polyadenylation sites are used by closely related ricin genes. Moreover, it is possible that multiple polyadenylation sites might be used by the individual genes themselves, as has already been shown to occur with a number of plant genes [3]. There may be some significance in that fact that the pCBG3H1 ricin gene contains two polyadenylation signals.

Developmental expression of lectin genes

For the purpose of assessing the pattern of lectin transcript accumulation during seed development, developing seeds were subdivided into 6 arbitrary groups; stages A to C (pre-testa formation) and stages D to E (post-testa formation), as described earlier [29]. Total RNA was extracted from seed endosperm of each stage, as well as from dry seeds, germinating seeds, roots and leaves. Figure 4 shows an autoradiograph of blotted RNA samples extracted from the various tissues and developmental stages, probed with a ricin cDNA and washed at low stringency so as to identify all lectin gene transcripts. The result clearly demonstrates that lectin mRNAs accumulate in a seed-specific fashion. No lectin mRNA is detected in the pre-testa stages of seed development, but by stage D, a band corresponding to the lectin transcripts is evident. The level of lectin transcripts increases gradually up to and including stage F, but no lectin mRNA is detected in the dry seed tissue. This is in accordance with the developmental pattern observed at the level of translatable mRNA [29]. Lectin transcripts are not detected in either leaf or root tissue or in the endosperm of germinating seeds.

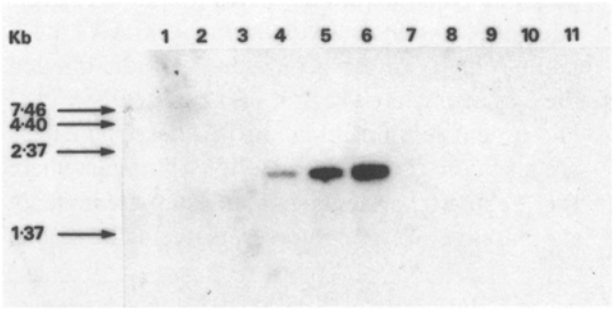


Fig. 4. Northern analysis of lectin transcript levels in different tissues and stages of seed development. 10 μg total RNA samples were electrophoresed through a 1.2% agarose formamide gel [23] and northern-blotted onto a nylon membrane. The membrane was probed with a radiolabelled ricin cDNA fragment and washed at low stringency. The positions of the RNA size markers are indicated by the arrows. The sources of tissue used are as follows: track 1, maturing endosperm stage A; track 2, stage B; track 3, stage C; track 4, stage D; track 5, stage E; track 6, stage F; track 7, dry seed endosperm; tracks 8 and 9, endosperm tissue from 3- and 7-day germinating seeds respectively; track 10, mature leaf tissue; track 11, root tissue.

RNase protection analysis

In order to identify the transcription start sites of the pCBG3H1 ricin gene and verify its seed-specific pattern of expression, RNase protection analysis was carried out using aliquots of the same total RNA samples used above in the northern blot analyses. The [α^{32}P]-radiolabelled RNA probe was obtained by *in vitro* transcription of pGEM-blue (Promega Biotech) containing a *Sph* I-*Bam* HI DNA fragment spanning the 5′ end of the pCBG3H1 ricin gene (position 137 to 374 in Fig. 3). Figure 5 shows an autoradiograph of the RNase digestion products. The protected fragments occur as a series of doublets or triplets which represent artefacts caused by 'breathing' of the complementary strands giving rise to a 'nibbling' effect [33]. This effect is commonly observed with A/T-rich sequences. If the 'nibbling' effect is taken into account, as many as 4 possible transcription start site positions can still be identified. It seems unlikely that all these positions could be used, but the possibility of multiple transcription start sites or of heterogeneous mRNA 5′-processing cannot be ruled out. Alternatively, the heterogeneity could be due to differences in the transcripts of different members of the gene family. At least two potential TATA boxes are present in the pCBG3H1 gene 5′flanking region (Fig. 3), a phenomenon seen in the promoters of several plant genes which initiate

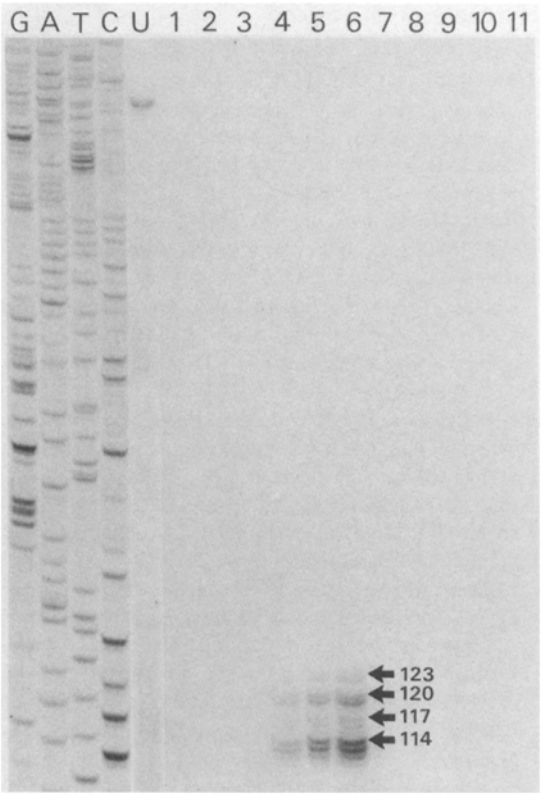


Fig. 5. RNase protection analysis of transcripts using a 237 bp *Sph* I/*Bam* HI fragment probe spanning the 5′ end of the pCBG3H1 (lambda clone 10) ricin gene. In each case 10 μg total RNA were analysed by RNase protection. On the left are sequencing ladder markers (G, A, T and C). Track U shows the undigested RNA probe. Tracks 1 to 11 show the digestion products of the following samples: track 1, maturing endosperm stage A; track 2, stage B; track 3, stage C; track 4, stage D; track 5, stage E; track 6, stage F; track 7, dry seed endosperm; tracks 8 and 9, endosperm tissue from 3- and 7-day germinating seeds respectively; track 10, mature leaf tissue; track 11, root tissue. The sizes of the RNase-protected fragments in nucleotides are indicated.

transcription from more than one point [11]. All possible transcription start sites are shown in Fig. 3. The putative transcription start site corresponding to the most abundant RNase-protected species is that which occurs 51 bp upstream from the translation initiation codon. Nevertheless it is assumed that the putative transcription start site at −60 (with respect to the translation initiation codon) is used *in vivo*, albeit at a low level, since this represents the mRNA species which shares the longest region of homology with the pCBG3HI probe.

The pattern of lectin gene transcript accumulation evident from the autoradiograph in Fig. 5 is in good agreement with the pattern seen in the northern blot shown in Fig. 4. Overall, the developmental profile of lectin gene transcript levels observed here closely resembles the profiles observed at the protein and translatable mRNA levels [9, 29]. This indicates a close correlation between the accumulation of the lectin proteins and the amounts of steady-state transcripts. It remains to be determined by run-on transcription experiments whether the accumulation of lectin gene transcripts is a result of transcriptional activation or decreased degradation of transcripts or both.

Other R. communis *lectin genes*

In this study, five different members of the *Ricinus communis* lectin gene family have been cloned. Three of the clones have been shown to contain lectin pseudogenes, whilst the fourth clone characterised has been shown to contain a functional gene which encodes a ricin D type polypeptide. It remains to be determined whether or not the lectin gene in the fifth clone (which contains an incomplete 5′ terminus) contains a functional lectin gene. Southern blotting has shown that the castor bean lectin gene family consists of approximately eight members, so the number of genes which remain unaccounted for amongst the five clones characterised here is probably about three. One of these genes must presumably encode an RCA I polypeptide, since none of the genomic clones resembles the previously described RCA I cDNA clone [28]. At least one of the other genes must presumably encode a ricin polypeptide since we believe that the pAKG ricin clone [10], despite its high degree of homology with the pCBG3H1 clone, represents a different member of the lectin gene family. In the absence of restriction fragment length polymorphism (RFLP) data, this conclusion is not yet unequivocal. However, a number of observations support the

contention. Firstly, the RNase protection data obtained with the pCBG3H1 clone are strikingly different from the S1 nuclease mapping data obtained with the pAKG clone, which initiates transcription further downstream and apparently from only one point. Secondly, whilst it is possible that the nucleotide sequence differences between the two clones could be caused by inter-varietal divergence, it is notable that the pCBG3H1 promoter contains a typical seed-specific RY repeat element [4] at −113 (with respect to the most 5′ transcription start point) which the pAKG clone lacks. Thirdly, a total of 10 out of the 17 lectin-positive lambda clones isolated were found by high-stringency hybridisation to contain DNA fragments highly homologous to the ricin cDNA. Although this could be caused by a cloning bias towards ricin genes, it seems likely that this observation reflects the presence of multiple-copy ricin genes in the *R. communis* genome. Finally, the postulated existence of at least two ricin genes in the genome of the castor bean strain used is consistent with the pattern of hybridisation observed with high-stringency Southern blots (data not shown). A ricin E gene of the type described in Ladin *et al.* [17] may also be present in the genome, although to date such a gene has only been observed to occur in small grain castor bean seeds [1].

The work described here will allow further studies into the mechanism of developing endosperm-specific lectin gene expression in *Ricinus communis*. The sequence data obtained has allowed the identification of possible *cis*-acting DNA elements in the pCBG3H1 ricin gene promoter and will now allow us to begin to investigate the molecular basis of the regulation of ricin gene expression.

Acknowledgements

The authors thank the Science and Engineering Research Council for a research studentship to J.W.T. Thanks are also extended to Martin Hartley for critical reading of the manuscript and to David Griffiths for advice on the DNA preparation method.

References

1. Araki T, Funatsu G: The complete amino acid sequence of the B chain of ricin E isolated from small grain castor bean seeds. Ricin E is a gene recombination product of ricin D and *Ricinus communis* agglutinin. Biochim Biophys Acta 911: 191–200 (1987).
2. Biggin MD, Gibson TJ, Hong GF: Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc Natl Acad Sci USA 80: 3963–3965 (1983).
3. Dean C, Tamaki S, Dunsmuir P, Faureau M, Katayama C, Dooner H, Bedbrook J: mRNA transcripts of several plant genes are polyadenylated at multiple sites *in vivo*. Nucl Acids Res 14: 2229–2240 (1986).
4. Dickinson CD, Evans RP, Nielson NC: RY repeats are conserved on the 5′-flanking regions of legume seed-protein genes. Nucl Acids Res 16: 371 (1988).
5. Endo Y, Mitsui K, Motozuki M, Tsurugi K: The mechanism of action of ricin and related toxic lectins on eukaryotic ribosomes. J Biol Chem 262: 5908–5912 (1987).
6. Feinberg AP, Vogelstein B: A technique of radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132: 6–13 (1983).
7. Funatsu G, Yoshitake S, Funatsu M: Primary structure of the Ile chain of ricin D. Agric Biol Chem 42: 501–503 (1978).
8. Gallagher TF, Ellis RJ: Light-stimulated transcription of gene for two chloroplast polypeptides in isolated pea leaf nuclei. EMBO J 1: 1493–1498 (1982).
9. Gifford DJ, Greenwood JS, Bewley JD: Deposition of matrix and crystalloid storage proteins during protein body development in the endosperm of *Ricinus communis* L. cv. Hale seeds. Plant Physiol 69: 1471–1478 (1982).
10. Halling KC, Halling AC, Murray EE, Ladin BF, Houston LL, Weaver RF: Genomic cloning and characterisation of a ricin gene from *Ricinus communis*. Nucl Acids Res 13: 8019–8033 (1985).
11. Herrmann A, Schultz W, Hahlbrock K: Two alleles of the single-copy chalcone synthase gene in parsley differ by a transposon-like element. Mol Gen Genet 212: 93–98 (1988).
12. Irwin SD, Keen JN, Findlay JBC, Lord JM: The *Ricinus communis* 2S albumin precursor: A single preproprotein may be processed into two different heterodimeric storage protein. Mol Gen Genet 222: 400–408 (1990).
13. Kaminski PA, Buffard D, Strosberg AD: The pea lectin gene family contains only one functional gene. Plant Mol Biol 9: 497–507 (1987).
14. Keller GH, Taylor JM: Effect of hypophysectomy on the synthesis of rat liver albumin. J Biol Chem 251: 3768–3772 (1976).
15. Kislev N, Rubenstein I: Utility of ethidium bromide in the extraction from whole plants of high molecular weight maize DNA. Plant Physiol 66: 1140–1143 (1980).
16. Kreig PA, Melton DA: *In vitro* RNA synthesis with SP6 RNA polymerase. Meth Enzymol 155: 397–415 (1987).

17. Ladin BF, Murray EE, Halling KC, Tilakaratne N, Long GL, Houston LL, Weaver RF: Characterisation of a cDNA encoding ricin E, a hybrid ricin-*Ricinus communis* agglutinin gene from the castor oil plant *Ricinus communis*. Plant Mol Biol 9: 287–295 (1987).
18. Lamb FI, Roberts LM, Lord JM: Nucleotide sequence of cDNA coding for preproricin. Eur J Biochem 148: 265–270 (1985).
19. Loenen WAM, Blattner FR: Lambda charon vectors (Ch32, 33, 34 and 35) adapted for DNA cloning in recombination-deficient hosts. Gene 26: 171–179 (1983).
20. Logemann J, Schell J, Willmitzer L: Improved method for the isolation of RNA from plant tissues. Anal Biochem 163: 16–20 (1987).
21. Lord JM, Roberts LM: Plant toxins. Microbiol Sci 4: 376–378 (1987).
22. Maniatis T, Fritsch EF, Sambrook J: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
23. May MJ, Hartley MR, Roberts LM, Kreig PA, Osborn RW, Lord JM: Ribosome inactivation by ricin A chain: a sensitive method to assess the activity of wild-type and mutant polypeptides. EMBO J 8: 301–308 (1989).
24. Messing J, Vieira J: A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene 19: 269–276 (1982).
25. Olsnes S, Saltvedt E, Pihl A: Isolation and comparison of galactose-binding lectins from *Abrus precatorius* and *Ricinus communis*. J Biol Chem 249: 803–810 (1974).
26. Rigby PWJ, Dieckmann M, Rhodes C, Berg P: Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA Polymerase I. J Mol Biol 113: 237–251 (1977).
27. Roberts LM, Lamb FI, Lord JM: Biosynthesis and molecular cloning of ricin and RCA. 24. In: Membrane Mediated Cytotoxicity, pp. 73–82. Alan R. Liss, New York (1987).
28. Roberts LM, Lamb FI, Pappin DJC, Lord JM: The primary sequence of *Ricinus communis* agglutinin. Comparison with ricin. J Biol Chem 260: 15682–15686 (1985).
29. Roberts LM, Lord JM: Protein biosynthetic capacity in the endosperm tissue of ripening castor bean seeds. Planta 152: 420–427 (1981).
30. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463–5467 (1977).
31. Vieira J, Messing J: The pUC plasmids, and M13 mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19: 259–268 (1982).
32. Vodkin LO, Rhodes PR, Goldberg RB: A lectin insertion has the structural features of a transposable element. Cell 34: 1023–1031 (1983).
33. Williams JG, Mason PJ: Hybridisation in the analysis of RNA. In: Glover DM (ed) DNA Cloning, vol. 2. IRL Press (1985).