Theor Appl Genet (2004) 108:1098–1102
DOI 10.1007/s00122-003-1520-9

ORIGINAL PAPER

I. Galasso · L. Lioi · C. Lanave · R. Bollini · F. Sparvoli

Identification and isolation of lectin nucleotide sequences and species relationships in the genus *Lens* (Miller)

Received: 14 January 2003 / Accepted: 23 October 2003 / Published online: 20 November 2003

Abstract Genes for lectin, a component of legume storage proteins, were identified and characterised in two lentil cultivars (*Lens culinaris* ssp. *culinaris*) and six wild relatives. In each taxon no differences were found among the two or three lectin clones sequenced, while differences were observed among lectin genes isolated from the different taxa. All of the clones analysed contained an insert of 828 bp and showed a high similarity with the nucleotide sequence of *Pisum sativum* seed lectin, PSL1. The deduced amino acid lectin sequences in all taxa were 275 amino acids long, and their multiple alignment showed that most of the variation among them occurred in regions which are not important for metal- and sugar-binding. The data from Southern blot analysis indicated the presence of only one lectin gene in all *Lens* taxa except *L. tomentosus*. Phylogenetic analyses carried out on the lectin sequences showed the existence of two main clusters and clearly indicated that *L. nigricans* falls outside the two groups.

Introduction

Lens culinaris Medik. has two subspecies—the lentil crop ssp. *culinaris* and its wild progenitor ssp. *orientalis* (Boiss.) Ponert. The other species included in the genus *Lens* Mill., *L. ervoides* (Bring.) Grande, *L. nigricans* (Bieb.) Godr., *L. odemensis* Ladiz. and the two recently recognised species *L. lamottei* Czefr. and *L. tomentosus* Ladiz., only occur as wild species and are unused sources of desirable agronomic and quality traits (Bayaa et al. 1994). However, for the breeder to fully exploit this potential, a comprehensive understanding of both the genetic variation and the taxonomic relationships of this crop and its wild relatives is required.

Legume seeds contain three major classes of storage proteins—vicilin, legumin and lectin—which together constitute more than 80% of the total seed protein and are major contributors to the nutritional quality of the seeds (Vitale and Bollini 1995). One of these, the lectins, is a family of highly homologous proteins with binding sites for divalent cations, such as calcium and manganese, which are essential for the carbohydrate-binding activity (Loris et al. 1998; van Damme et al. 1998). Lectins have been found mainly in the seeds of most legume plants but are also quite abundant in several vegetative organs (van Damme et al. 1998). It has been suggested that lectins play a role in seed maturation, cell-wall assembly, defence mechanisms and rhizobial nodulation of legume roots (Liener 1997).

Depending on the post-translational processing of the protomer, legume lectins can be separated into two major classes: single-chain lectins, such as that found in the common bean (*Phaseolus vulgaris* L.), phytohemagglutinin (PHA), and two-chain lectins, as the one found in pea (*Pisum sativum* L.) (Vitale and Bollini 1995). In cultivated lentil the seed lectin is a homodimeric two-chain protein, the monomers of which contain an α- and a β-chain of about 7 kDa and 17 kDa, respectively (Foriers et al. 1981; Loris et al. 1993). Although lentil lectin has been well-characterised, data on its DNA sequence and genomic organisation are still lacking. In the study reported here, we isolated lectin nucleotide sequences present in the genome of cultivated and wild *Lens* species and explored their usefulness for phylogenetic analysis.

Communicated by J.S. Heslop-Harrison

I. Galasso (✉) · L. Lioi
Istituto di Genetica Vegetale,
CNR, via G. Amendola 165/A, 70126 Bari, Italy
e-mail: galasso@ibba.cnr.it
Tel.: +39-02-23699421
Fax: +39-02-23699411

C. Lanave
Sezione di Bioinformatica e Genomica ITB,
CNR, via G. Amendola 168/5, 70126 Bari, Italy

I. Galasso · R. Bollini · F. Sparvoli
Istituto di Biologia e Biotecnologia Agraria,
CNR, via Bassini 15, 20133 Milan, Italy

Materials and methods

Plant materials

Two cultivated lentil cultivars [*Lens culinaris* ssp. *culinaris*], and one accession from each wild *Lens* taxa together with one sample of *Pisum sativum* and *Phaseolus vulgaris* were used in this study (Table 1).

Protein extraction, SDS-PAGE and immunoblot analysis

Total seed protein was extracted from dry seeds by homogenisation and separated by electrophoresis on an 18% sodium dodecyl sulfate-polyacrylamide gel (SDS-PAGE) (Bollini and Chrispeels 1978). Immunoblot analysis was performed according to Burnette (1981) using goat antibodies against *L. culinaris* lectin (LCA) (Vector, AS-2044) at 1:500 dilutions. Peroxidase-linked donkey anti-goat IgG (Promega, Madison, Wis.; V805A) was used as the secondary antibody.

Isolation of lectin nucleotide sequences and Southern hybridisation

Genomic DNA was extracted from leaf tissue (Dellaporta et al. 1983). Lectin sequences were amplified by PCR using specific primers designed on the sequence of pea seed lectin gene, PSL1 (EMBL, M18160). The two primers, MAS-For (5′-ATG-GCTTCTCTTCAAACCCAAATGACT-3′) and SSK-Rev (5′-CTATGCATCTGCAGCTTGCTTAGAAC-3′), flank the entire seed lectin gene of pea. The amplification programme consisted of one cycle at 95°C for 5 min, 40 cycles of 94°C for 1 min, 56°C for 1 min and 72°C for 1 min, and a final elongation at 72°C for 7 min. In all taxa a single band of about 820 bp was amplified. The PCR product of each taxon was then purified and cloned in pGEM-T plasmid. Two or three clones for each sample were selected and the inserts sequenced in both directions with an automated sequencer. The nucleotide sequences obtained are available in the EMBL Sequence Database (for accession numbers see Table 1). For Southern hybridisation, 5 µg of genomic DNA was digested with the restriction enzyme *Rsa*I, size-separated and transferred onto a nylon membrane. The insert of the Est1-17 clone was labelled with α-[^{32}P]-dATP and used for filter hybridisation as described in Sparvoli et al. (1996). The filter was washed in 1.5 mM sodium citrate pH 7.0, 15 mM NaCl, 0.5% SDS (0.1× SSC), at 65°C.

Sequence comparison and phylogenetic analysis

Deduced amino acid sequences were multialigned using the CLUSTALW programme. Evolutionary analysis was carried out using the programmes implemented in the PAUP* package (Swofford 1998). Genetic distances were calculated by using the General Time Reversible (GTR) model (Lanave et al. 1984; Saccone et al. 1990) at the synonymous nucleotide substitutions. The phylogenetic trees were reconstructed using the ML/GTR/NJ-UPGMA methods, then TreeView was used to draw the unrooted and rooted tree.

Results and discussion

SDS-PAGE and immunoblot analysis

Anti-*L. culinaris* LCA antibodies detected two major lectin polypeptides in all *Lens* extracts and in *P. sativum* (Fig. 1). The most abundant one was about 22 kDa (Fig. 1B) and therefore must correspond to the β-chain. This polypeptide is of a similar size in the two lentil cultivars (lanes 1, 2) and in *L. ervoides* (lane 7) and is slightly larger than those found in the remaining species. The electrophoretic mobility of these β-chains was higher than that observed by Fories et al. (1981). When we compared our results with the amino acid sequence of the *L. c*. ssp *culinaris* β-chain, which has a size of 19.89 kDa, we found more agreement. Such anomalous behaviour on SDS-PAGE has also been reported for other lectins (Sparvoli et al. 1998). The second major cross-reacting component that was present in the *Lens* taxa and the pea extracts was about 31 kDa and most likely represents residual amounts of still uncleaved lectin precursor.

Putative α-polypeptides were detected only in *L. odemensis*, *L. lamottei* and *P. sativum* (lanes 5, 6 and 9, respectively) as a faint band in the range of 5–6 kDa, which is where the α-chain would be expected to migrate. This polypeptide was never detected in the other samples even if the SDS-PAGE was overloaded. Given the high level of sequence identity determined for all of the α-chains, this lack of detection in the other lentil protein extracts suggests that this polypeptide is present at a level too low to be detected. In addition, α-chains may be split into different bands since isoforms are generated as the result of multiple C-terminal cleavages (Young et al. 1996). No cross-reacting polypeptides were observed in the *P. vulgaris* extract (lane 10).

Table 1 List of taxa used in this study with the respective donor code, EMBL accession number and clone name

Number	Species	Subspecies	Donor code[a] or name	EMBL no.	Clone name
1	*Lens culinaris*	*culinaris*	Cultivar Eston	AJ318218	Est1-17
2	*Lens culinaris*	*culinaris*	Cultivar Laird	AJ318217	Lai2-71
3	*Lens culinaris*	*orientalis*	ILWL 144	AJ318219	Ori4-31
4	*Lens tomentosus*		No. 133	AJ421799	Tom133
5	*Lens odemensis*		ILWL 223	AJ318222	Ode9-1
6	*Lens lamottei*		ILWL 430	AJ419573	Lam8-51
7	*Lens ervoides*		ILWL 387	AJ318220	Erv6-22
8	*Lens nigricans*		ILWL 112	AJ318221	Nig7-42
9	*Pisum sativum*		Local market		
10	*Phaseolus vulgaris*		Cultivar Tendergreen		

[a] ILWL, From ICARDA, Aleppo, Syria. *L. tomentosus* no. 133 was kindly provided by Prof. G. Ladizinsky

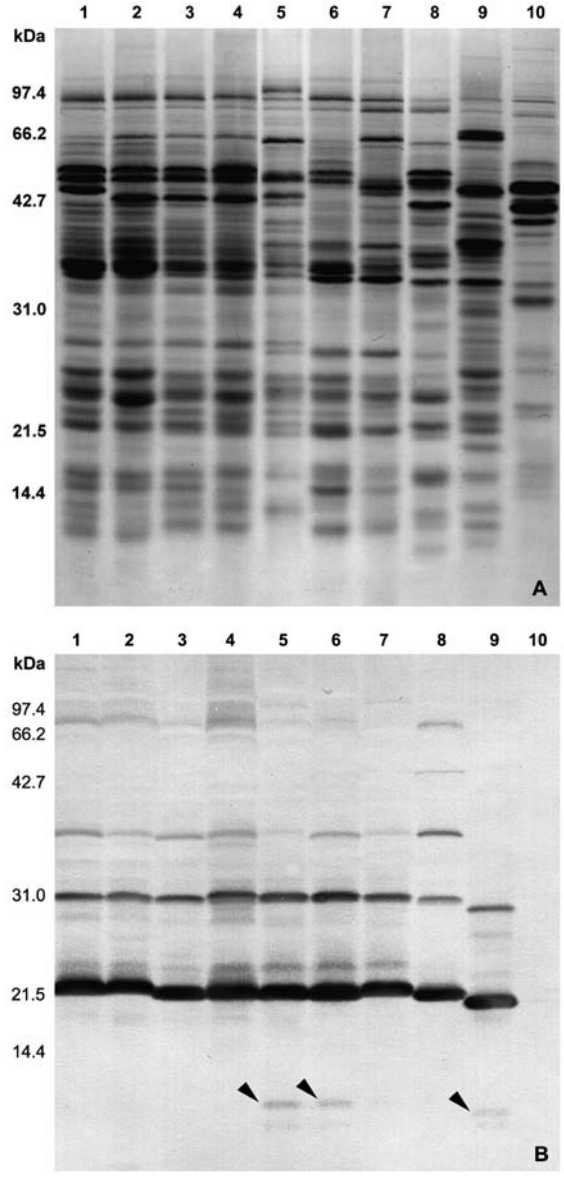

Fig. 1A, B SDS-PAGE and immunoblot analysis of total seed protein extracts. **A** Polypeptides stained with Coomassie brilliant blue. **B** Lentil lectin-related polypeptides detected with an antiserum against *Lens culinaris* lectin (LCA). *Arrows*: α-chain polypeptides. Numbers at the top of plates correspond to those listed in Table 1. Molecular weight is indicated on the left

Lectin gene isolation and genomic organisation

PCR amplification always yielded a single band of about 820 bp. An analysis of the amplified DNA band with several restriction endonucleases to check whether it contained a single type or different DNA fragments indicated that each lentil species contained a single type of DNA sequence. When two or three clones were sequenced for each taxon, this absence of heterogeneity within each taxon was confirmed, whereas when the nucleotide sequences of genes belonging to different species were compared, differences were detected. Analysis of the inserts showed the presence of one *Rsa*I restriction site (GA↓TC) in all of the taxa but *L. c*. ssp *culinaris* and *L. tomentosus*. A FASTA search of these inserts against the EMBL and Swiss-Prot databases revealed a high similarity (92–93%) with the sequence of pea seed lectin PSL1. In lentil, an amino acid segment of six residues (Ser-Leu-Glu-Glu-Glu-Asn) was always found between the two chains in the deduced amino acid sequences that was absent in the mature lectin protein. The absence of this six-residue segment in the mature lectin protein suggests that it is removed during post-translational processing (Kaminski et al. 1987). Additionally, our α-chains have four amino acid residues (Ala-Ala-Asp-Ala) at the C-terminus that have not been reported in the sequence of the mature polypeptide. Most likely, these residues are removed post-translationally by a carboxypeptidase, as described by Young et al. (1996). Such a C-terminal cleavage is typical of proteins that accumulate in storage vacuoles (Matsuoka and Neuhaus 1999).

The multiple alignment of the lectin nucleotide sequences showed that cvs. Eston and Laird contained an identical lectin gene. Among the wild species, *L. odemensis* and *L. lamottei* contained highly similar lectin sequences, while *L. nigricans* showed the highest number of nucleotide substitutions with respect to the wild and cultivated species (Table 2). The amino acid sequence deduced from cvs. Eston and Laird clones differed for one and three amino acid residues from the mature β- and α-sequences of Loris et al. (1993), respectively. Conversely, our α-chain sequence is in complete agreement with the one re-determined by protein sequencing and mass spectrometry by Young et al. (1996).

The genomic organisation of the lentil lectin was investigated by Southern blot analysis. The hybridisation patterns, obtained with the labelled 828-bp insert of Est1-

Table 2 Percentages of similarity at the nucleotide level among cloned lectins

	Lai2-74	Ori4-31	Tom133	Ode9-1	Lam8-51	Erv6-22	Nig7-42	PSL1
Est1-17	100	99.8	99.3	98.8	98.7	97.8	96.9	92.9
Lai2-74		99.8	99.3	98.8	98.7	97.8	96.9	92.9
Ori4-31			99.3	99.0	98.9	98.1	97.2	92.9
Tom133				98.8	98.8	97.8	96.9	92.9
Ode9-1					99.9	99.0	97.7	93.5
Lam8-51						98.9	97.6	93.4
Erv6-22							97.7	93.5
Nig7-42								93.4

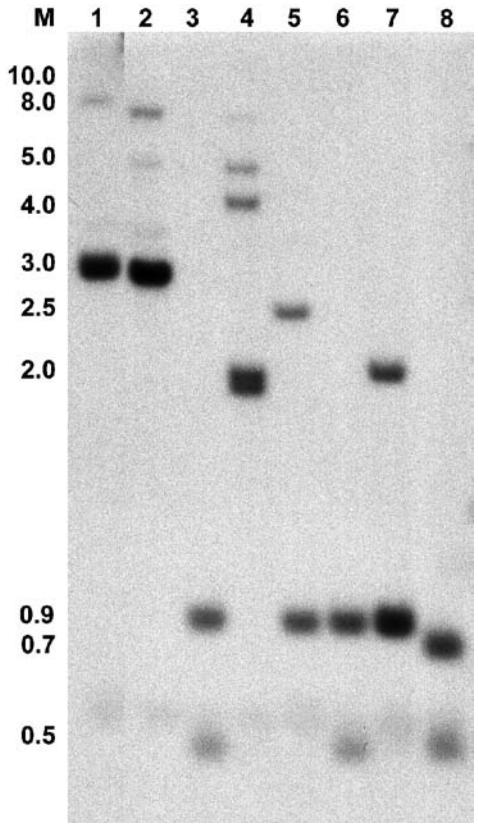

Fig. 2 Southern blot analysis of *Lens* genomic DNAs digested with *Rsa*I restriction enzyme and probed with the 828-bp insert of the Est1-17 clone. M DNA size marker (in kilobases). Numbers at the top correspond to those listed in Table 1

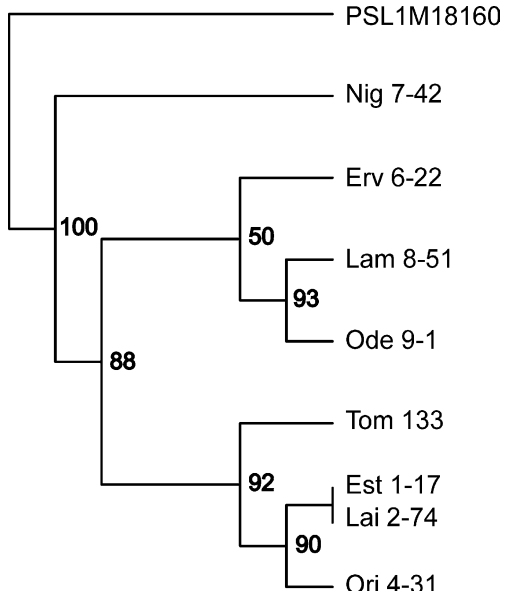

Fig. 3 Phylogenetic tree was calculated on the basis of synonymous codon positions of the lectin sequences isolated in *Lens* taxa, using the General Time Reversible and UPGMA methods. Code names are as indicated in Table 1. On each node the bootstrap value based on 100 replicates is indicated. In the UPGMA tree, the sequence of *Pisum sativum* was used as the outgroup

17 lectin clone, were very simple (Fig. 2). One major hybridising fragment of about 3 kb together with weak hybridising bands were detected in *L. c*. ssp *culinaris*, suggesting the presence of a single gene (Fig. 2, lanes 1, 2). Two major hybridising fragments of different sizes were detected in *L. c*. ssp. *orientalis* (0.9 kb and 0.5 kb), *L. odemensis* (2.5 kb and 0.9 kb), *L. lamottei* (0.9 kb and 0.5 kb), *L. ervoides* (2.0 kb and 0.9 kb) and *L. nigricans* (0.7 kb and 0.5 kb) (Fig. 2, lanes 3, 5–8). In these taxa an internal *Rsa*I restriction site is present in the isolated lectin gene, suggesting that the two fragments result from the cleavage of a single gene. We detected a more complex hybridisation pattern in *L. tomentosus*: four hybridisation fragments with similar intensities were observed (Fig. 2, lane 4). Since no *Rsa* I site was found in its lectin sequence, probably more than one lectin gene type might be present in its genome.

In pea at least four lectin genes are present; one (PLS1) codes for a functional polypeptide, while the others (PLS2, PLS3, PLS4) are pseudogenes and contain several stop codons (Kaminski et al. 1987). In lentil, although we used genomic DNA as the template for the PCR cloning of the lectin gene, the designed primers were apparently able to amplify only nucleotide sequences corresponding to the true seed lectins. No lectin-related genes or pseudogenes were amplified, but their presence on the genome cannot be ruled out. Indeed, in the genome of the cultivated taxon, we detected two or more weak hybridisation fragments together with a strong one. Based on the finding that the lectin gene isolated in cvs. Eston and Laird does not contain any *Rsa* I sites and since identical results were obtained in independent experiments performed with the same enzyme, these less intense fragments most likely correspond to a cross-hybridisation with lectin-related sequences rather than being products of an incomplete digestion of the DNA. While Southern blot analysis of *L. tomentosus* suggested the presence of more then one lectin gene, we identified only one type of lectin sequence among the clones analysed.

Phylogenetic analysis

The extent of variability among the lectin genes isolated here made a phylogenetic analysis possible. The evolutionary distances among the eight lectin nucleotide and amino acid sequences were used to construct a rooted (Fig. 3) and an unrooted phylogenetic tree (not shown). The phylogenetic analyses indicated the existence of two main clusters: one grouping the two cultivars Laird and Eston, the wild progenitor ssp. *orientalis* and *L. tomentosus*, and the other one grouping *L. ervoides*, *L. odemensis* and *L. lamottei*. *Lens nigricans* does not cluster with any of the other samples and shows the greatest distance from all of the taxa analysed (Fig. 3). The tree obtained with the maximum likelihood method on the amino acids alignment was in agreement with these

results (not shown). This great divergence of *L. nigricans* is consistent with the majority of the relevant literature based on crossability, RFLP, cpDNA, ITS sequence analysis and FISH karyotype (Ladizinsky et al. 1984; Havey and Muehlbauer 1989; van Oss et al. 1997; Galasso 2003; Sonnante et al. 2003). A second interesting observation is the clustering of *L. tomentosus* with the cultigen and *L. c.* ssp. *orientalis*. Although not supported by all reports (van Oss et al. 1997; Galasso 2003; Sonnante et al. 2003), it is in agreement with the work of Zimniak-Przybylska et al. (2001), which is based on total seed protein pattern.

Acknowledgements The authors are grateful to all the donors of seed samples. This research is partially supported by grants from the MIUR Italy project Cluster C03-Ingegneria Molecolare, legge 488/92

References

Bayaa B, Erskine W, Hamdi A (1994) Response of wild lentil to *Ascochyta fabae* f.ssp. *Lentis* from Syria. Genet Res Crop Evol 41:61–65

Bollini R, Chrispeels MJ (1978) Characterisation and subcellular localisation of vicilin and phytohemagglutinin, the two major reserve proteins of *Phaseolus vulgaris* L. Planta 142:291–298

Burnette WN (1981) Western blotting: electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. Anal Biochem 112:195–203

Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: Version II Plant Mol Biol Rep 1:19–21

Fories A, Lebrun E, Van Rapenbusch R, de Neve R, Strosberg AD (1981) The structure of lentil (*Lens culinaris*) lectin. J Biol Chem 256:5550–5560

Galasso I (2003) Distribution of highly repeated DNA sequences in species of the genus *Lens* Miller. Genome (in press)

Havey MJ, Muehlbauer FJ (1989) Variability for restriction fragment lengths and phylogenesis in lentil. Theor Appl Genet 77:839–843

Kaminski PA, Buffard D, Strosberg AD (1987) The pea lectin gene family contains only one functional gene. Plant Mol Biol 9:497–507

Ladizinsky G, Braun D, Goshen D, Muehlbauer FJ (1984) The biological species of the genus *Lens* L. Bot Gaz 145:253–261

Lanave C, Preparata G, Saccone C, Serio G (1984) A new method for calculating evolutionary substitution rates. J Mol Evol 20:86–93

Liener IE (1997) Plant lectins: properties, nutritional significance, and function. In: Am Chem Soc (ed) Antinutrients and phytochemicals in food. Am Chem Soc, Washington, D.C., pp 31–43

Loris R, Steyaert J, Maes D, Lisgarten J, Pickersgill R, Wyns L (1993) Crystal structure determination and refinement at 2.3 Å resolution of the lentil lectin. Biochemistry 32:8772–8781

Loris R, Hamelryck T, Bouckaert J, Wyns L (1998) Legume lectin structure. Biochim Biophys Acta 1383:9-36

Matsuoka K, Neuhaus JM (1999) *Cis*-elements of protein transport to the plant vacuoles. J Exp Bot 50:165–174

Saccone C, Lanave C, Pesole G, Preparata G (1990) Influence of base composition on quantitative estimates of gene evolution. Method Enzymol 183:570–583

Sonnante G, Galasso I, Pignone D (2003) ITS sequence analysis and phylogenetic inference in the genus *Lens* Mill. Ann Bot 91:49–54

Sparvoli F, Daminati MG, Lioi L, Bollini R (1996) In vivo endoproteolytically cleaved phaseolin is stable and accumulates in developing *Phaseolus lunatus* L. seeds. Biochim Bioph Acta 1292:15–22

Sparvoli F, Gallo A, Marinelli D, Santucci A, Bollini R (1998) Novel lectin-related proteins are major components in Lima bean (*Phaseolus lunatus* L.) seeds. Biochim Bioph Acta 1382:311–323

Swofford DL (1998) PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Assoc, Sunderland, Mass.

Van Damme EJM, Peumans WJ, Barre A, Rougè P (1998) Plant lectins: a composite of several distinct families of structurally and evolutionary related proteins with diverse biological roles. Crit Rev Plant Sci 17:575–692

Van Oss H, Aron Y, Ladizinsky G (1997) Chloroplast DNA variation and evolution in the genus *Lens* Mill. Theor Appl Genet 94:452–457

Vitale A, Bollini R (1995) Legume storage proteins. In: Kigel J, Galili G (eds) Seed development and germination. Marcel Dekker, New York, pp 73–102

Young NM, Watson DC, Thibault P (1996) Post-translational proteolytic processing and the isolectins of lentil and other Vicieae seed lectins. Glycocon J 13:575–583

Zimniak-Przybylska Z, Przybylska J, Krajewski P (2001) Electrophoretic seed globulin patterns and species relationships in the genus *Lens* Miller. J Appl Genet 42:435–447