Isolation and analysis of genomic clones coding for legumin storage proteins and their copy number determination in chickpea, *Cicer arietinum* L.

Lekha D Kumar, A D Mandaokar & K R Koundal
National Research Centre, Plant Biotechnology, Indian Agricultural Research Institute, New Delhi 110 012, India

Received 8 March 1999; revised 28 June 1999

The genes encoding legumin storage proteins of chickpea were identified initially by Southern blot hybridization of restricted total genomic DNA of chickpea with pea legumin cDNA probes (pDUB6 and pDUB8). Since these probes showed strong homology to chickpea DNA, they were used for screening chickpea genomic library to isolate and ascertain the positive clones. Confirmation and detailed analysis of these clones were then done by restriction analysis and Southern hybridization. The gene copy number of legumin was estimated and found to be 10 to 12 in chickpea genome.

Chickpea is the major pulse crop in India and provides a substantial portion of the dietary protein. But unfortunately, like most legumes, chickpea seed storage proteins being deficient in sulphur containing amino acids, methionine and cysteine suffer from an essential amino acid imbalance. Thus, improvement of nutritional quality of these proteins is very essential. Based on their sedimentation coefficient storage proteins of legumes are classified into two major groups i.e. 11S and 7S, otherwise known as legumin and vicilin respectively. These proteins are salt soluble, and deficient in sulphur containing amino acids and synthesized by multi gene family. The advent of recombinant DNA technology has made it possible to isolate storage protein genes and transfer it across species barriers. The isolation of genes are now possible by several ways. The most commonly used approach is by constructing and screening of genomic libraries using suitable probes.

Materials and Methods

Plant material—Seeds of chickpea (*Cicer arietinum* L.) var. 256 were surface sterilized using 0.01% mercuric chloride and germinated on paper towels in dark. A week old etiolated seedlings were harvested, frozen in liquid N2 and used for the study.

Isolation and southern blot hybridization of genomic DNA—Chickpea genomic DNA was isolated from 7 days old etiolated seedlings according to the modified protocol. Genomic DNA was further purified by CsCl2 gradient (0.90g/ml of DNA) ultracentrifugation in Vti 65 rotor at 60,000 rpm for 16 hr at 20°C. The DNA band was collected and dialysed against TE buffer (10mM Tris, 1 mM EDTA, pH 8.0) at 4°C overnight and precipitated using 0.1 volume of 3M sodium acetate and 2 volumes of ethanol. The yield of pure DNA was estimated by taking absorbance at 260 nm. Isolated chickpea genomic DNA (10 µg) was restricted with the independent sets of enzymes, each containing BamH1, EcoRI, EcoRV, HindIII, PspI, Pst1 and Sau3A. The restricted samples were electrophoresed on 0.7% agarose gel blotted on to gene screen plus filters. The membranes were hybridized using pea legumin cDNA probes, pDUB6 (leg6) and pDUB8 (leg8) separately (obtained from Prof. Donald Boulter, Department of Biological Sciences, University of Durham, UK). The hybridization was carried out at 65°C overnight. The membranes were washed with 2xSSC, 1xSSC and 0.5xSSC all with 0.1% SDS at 65°C. The washed filters were exposed to X-ray film in dark for 2 days and film was developed using Kodak developer and fixer.

Isolation of genomic clones from chickpea genomic library — The chickpea genomic library constructed earlier was amplified to get a titre value of 1.1x10^6 pfu/µl. A total number of 5x10^5 plaques were screened at a time for each of the probes. The plaques were lifted on to nitro-cellulose membrane filters and screened using radiolabelled pea legumin cDNA probes (pDUB6 and pDUB8) separately. The hybridization and washings were carried out as
described for the Southern blot hybridization of genomic DNA. Sixteen plaques lighted up by screening filters with leg8 and 12 plaques were identified using leg6. To separate out the positive clones further from the mixture of primary clones, secondary and tertiary screenings were done. These well separated clones from the secondary screening were ultimately spotted in duplicates on to the gene screen plus membrane discs and probed separately with pDUB6 and pDUB8 probes.

Analysis of the legumin genomic clones by Southern hybridization — In order to confirm the presence of legumin gene, the recombinant phage DNA was isolated from all the clones which showed hybridization with both the probes and were subjected to further analysis by restriction digestion and Southern analysis. All the clones showed a complex restriction pattern when digested with SalI. The restricted samples separated on 0.7% agarose gel in duplicates gave five to six bands in different clones suggesting internal sites for SalI. These blots were separately hybridized to pDUB6 and pDUB8 cDNA probes.

Estimation of gene copy number of legumin in chickpea—Total genomic DNA (10 μg) of chickpea was digested separately with EcoRV, HindIII and EcoRI enzymes. The gene copy equivalent to 10 μg of chickpea DNA (C=1.75 pg) was estimated and electrophoresed on 0.7% agarose gel. DNA was then blotted on nylon membrane and hybridized with 1.5 kb EcoRV/XhoI DNA fragment of chickpea legumin gene, leg3. The intensity of hybridization was compared with known amount of legumin DNA using gel scanner.

Results and Discussion

Identification of legumin sequences in chickpea using pea legumin cDNA probes — Genomic DNA of chickpea restricted with different set of enzymes and hybridized separately to pea legumin cDNA probes, pDUB6 and pDUB8 (Fig.1) showed differential hybridization with leg 8 (Fig. 2) and leg 6 (Fig.3). Though the probes were heterologous, they showed very strong hybridization with 1-2 kb and 4-5 kb fragments of genomic DNA suggesting strong homology with chickpea genomic DNA. Such kind of intergenic homologies have been shown in other legume genera like Vicia faba and Glycine max. These probes were therefore, used for screening the genomic library of chickpea.

Isolation of legumin genomic clones from chickpea genomic library—Chickpea genomic library constructed in EMBL3 vector was used for screening for legumin gene. Considering the insert size of 15-20 kb, approximately 5x10^5 independent recombinants from the genomic library have been screened in order to have a reasonable chance of including the desired sequences taking into account that legumin storage protein genes belong to multigene family. Hybridization with leg6 and leg8 separately showed a total of 10 positive clones. These 10 clones were then spotted in duplicate and hybridized separately to both the probes. All the 10
clones hybridized with leg8 (Fig. 4A) and designated as LD1 to LD10. However, when hybridized with leg6 cDNA probe only 6 out of 10 clones showed positive signal (Fig. 4B) and these were designated as LD1 to LD6. This indicated that these clones carry sequences homologous to the sequences encoding the β-subunit of the legumin polypeptide of pea and since leg8 (which codes for α subunit of legumin) also hybridized to these clones suggesting that LD1 to LD6 contain at least one full length gene sequence. The clones LD7 to LD10 showed strong hybridization with leg8 probe but did not show any hybridization with leg6 probe. This indicated that these clones carry sequences homologous to only 5' end of the legumin gene (encoding α subunit) and not to 3' end of the gene (encoding β subunits). The presence of such

intergeneric homology has been demonstrated between other legume genera like *Pisum* and *Glycine* and also between *Vicia* and *Glycine*. It would be interesting to know the extent of homology between legumin genes of chickpea and the rest of other legume species for which the sequences are known. This can give an insight into their phylogenetic relationship and the position of *Cicer* in the evolutionary tree with reference to other legumes.

**Analysis of legumin genomic clones by Southern hybridization** — In order to confirm the presence of legumin gene, the recombinant phage DNA was isolated from all the six clones (LD1 to LD6) which showed hybridization with both the probes and were subjected to further analysis. The recombinant phage DNA of all the six clones showed a complex restriction pattern when digested with SaII, suggesting that there are some internal sites of SaII in the genomic insert and Southern hybridization with pea legumin probe indicates that some of these clones contain more than one legumin gene.

The clone LD1, LD2 and LD3 upon restriction with SaII gave different restriction pattern (Fig. 5A), indicating that these three clones are different from each other. Southern hybridization with legumin cDNA (Leg6) probe also showed the different hybridization pattern of these clones (Fig. 5B). The LD1 clone showed the presence of single hybridizing band (Fig. 5B lane 1) indicating the presence of legumin gene in large fragment. Clone LD3 also showed hybridization at single fragment of about 9 kb (Fig. 5B, lane 3). Interestingly, clone LD2 showed hybridization at 3 bands of 9, 5 and 1 kb (Fig. 6B lane 2). As these fragments can not be the part of single gene and therefore at least two legumin genes must be presented in LD2 clone. The less intensity of 5 and 1 kb fragments could be because of the presence of internal site within a gene. The intensity differences between 9 kb and these two fragments also suggest the presence of legumin genes from different family in this clone.

Domoney et al. described interclass homologies within a species at about 50-60% at the nucleotide level. In *Phaseolus vulgaris* the genomic members of *phaseolin* gene family are very closely related with 71% to 95% homology. Genetic evidence which showed that *phaseolin* genes are closely linked led them to suggest that duplications and other events must have led to the formation of *phaseolin* multigene

---

Fig 3—Agarose gel electrophoresis of chickpea genomic DNA digested with various restriction enzymes (A) and Southern blot hybridisation (B) with pDUB6 probe. Lanes 1-8 indicate DNA restriction with BamH1, EcoRI, EcoRV, HindIII, HinfI, MspI, PstI, Sau3AI respectively. Lane 9 - undigested DNA, M - Molecular weight marker λDNA digested with HindIII.

Fig 4— Autoradiogram showing tertiary screening of chickpea genomic library probed with pDUB6 (A) and pDUB8 (B).
family. In *Pisum sativum* also, the genes coding for legA and legJ which belong to different classes of legumin share sequence homology\(^\text{11}\). In *Vicia faba* legumin A and B subfamilies were found to be recognizably 50\% homologous\(^\text{12}\), though they showed considerable sequence divergence at the amino acid level.

In case of LD1 and LD3 the number of genes cannot be assessed as there is only single hybridizing band. Since intensity in those bands is high it could be possible that more than one gene is present in these two clones as legumin genes are present in clusters\(^\text{15}\). This possibility can be tested by digesting these clones with different restriction enzymes and then hybridizing with legumin probe.

The other three clones, LD4, LD5 and LD6 were restricted with Sall and electrophoresed on agarose gel (Fig. 6, 1A and 1B). The clone LD5 was found to be different from the LD4 and LD6 whereas these two clones showed identical restriction pattern. To have further insight these restricted clones were blotted on hybrid membrane and hybridized separately with leg8 and leg6. As shown in Fig 6B, these three clones have differential hybridization pattern. Though the restriction pattern of LD4 and LD6 was similar (Fig. 6A) hybridization with either of the probes gave different pattern indicating that all the three clones are different. Hybridization of these three clones with both the probes gave similar pattern. Since leg8 and leg6 codes for 5' and 3' end of legumin polypeptide respectively and hybridization of both the probes to same restriction fragment suggests the presence of complete legumin gene(s) in all the three clones. The detailed characterization of these fragments with sequencing will reveal the exact nature of the legumin genes.

**Estimation of gene copies of legumin in chickpea** —Total genomic DNA of chickpea digested with restriction endonucleases and gene copy equivalent to 10 \(\mu\)g of DNA is shown in Fig. 7a. This DNA was then blotted and hybridized with 1.5 kb fragment of leg3 gene (Fig. 7b). The probe hybridized to number of restriction fragments of chickpea genomic DNA and intensity of bands indicates the presence of family of legumin genes. The intensity of
hybridization was compared with known amount of legumin DNA using gel scanner. The number of genes thus corresponds to 10-12 copies in chickpea genome.

The copy number obtained in this study differs with those available for *Glycine max* where five copies of glycinin (11S) protein gene have been estimated. However, the results presented here is in good agreement with those published for *pea* in which eight legumin genes per haploid genome have been reported. Legumin genes are classified into class I and class II family and high degree of sequence homology within the class and considerable degree of homology between the classes have been reported. This suggests the presence of conserved sequences among legumin genes and thus 1.5 kb fragment of leg3 which belongs to class I gene family could also have hybridized to class II gene family of chickpea. Therefore, 10-12 copies of chickpea legumin genes obtained in this study might be the members of both class I and class II families of chickpea. Further isolation and characterization of genes from class II family will give complete picture of legumin genes of chickpea.

The organization of seed storage protein genes is similar to that of other plant genes. These genes occur singly and in clusters on the same as well as on different chromosomes as has been reported in *Vicia faba* (15). Hence the genomic clones analysed in this study may contain complete gene or similar linked legumin genes in a single clone. Additional study may provide further insight to establish their structural and functional relationship.

Acknowledgement

The authors are grateful to Dr. R.P. Sharma, Project Director, NRCPB for providing the encouragement and facilities to carry out this research.

References