Isolation and Characterization of Protease Inhibitor Gene by Construction of Genomic Library of Black Gram (*Vigna mungo* L.)

HM Jagadeesh*, MT Ila, S Sreenivasa, KP Raghavendra, S Rashmi and KR Koundal

National Research Centre on Plant Biotechnology, Indian Agricultural Research Institute, Pusa Campus, New Delhi-110012

(Received: 24 November 2010; Revised: 7 July 2011; Accepted: 15 July 2011)

Plant protease inhibitors (PIs) are small proteins abundant in storage organs such as seeds and tubers. They have been implicated in binding to proteases of insects and interfere with the digestive process, there by proving their insecticidal activity. Black gram is known to be a source of plant protease inhibitors. A genomic DNA library of Black gram variety Uttara was constructed in Fosmid vector. Primary and secondary screening of library using mung bean protease inhibitor gene as a probe resulted in isolation of recombinant clones containing protease inhibitor gene and promoter. The positive clones were further confirmed by PCR and subsequently sequenced. The nucleotide sequence of genomic clone was analyzed by Blastn algorithm of NCBI database and 5' upstream sequence was analysed *in-silico* by Place and Plant Care databases.

Key Words: Black gram, Genomic library, Plant protease Inhibitors, Promoter

Introduction

Black gram is an important pulse crop of many South Asian countries including India. In India black gram cultivation occupies 3.29 million ha and contributes 1.60 million tones to pulse production. Like other legumes black gram is also known to be a source of plant protease inhibitors. Plant protease inhibitors (PIs) are small proteins which represent a class of well studied compounds of plant defence genes and are abundant in storage organs such as seeds and tubers (Joanitti *et al*., 2006). Majority of PIs in plant kingdom occur in three main families *viz*., Fabaceae, Solanaceae and Poaceae (Ryan, 1990; Lawrence and Koundal, 2002). The PIs form complexes with proteases, thereby inhibiting their proteolytic activity. There have been many examples of PIs active against certain insect species both in *in-vitro* assays against insect gut proteases (Koiwa *et al*., 1998) and *in-vivo* artificial diet bioassays (Urwin *et al*., 1997; Vain *et al*., 1998).

A direct proof of the protective role of PIs against insect herbivores was provided by Hilder *et al.* (1987). The first transgenic tobacco plant expressing PI was reported in 1987 (Hilder *et al*., 1987). Since then a large number of transgenic plants have been developed using PIs. The genes of plant origin have added advantages, as they are correctly transcribed, translated and processed in the host plant after genetic transformation without any codon bias. Further more, it has been demonstrated that mechanical wounding, insect chewing and microbial infection enhanced the level of PIs content significantly in local as well as remote tissues (Telang *et al*., 2003, Srinivasan *et al*., 2009). Insect-resistant crops have been one of the major successes of applying plant genetic engineering technology to agriculture; Cotton (*Gossypium hirsutum*) resistant to lepidopteran larvae (Caterpillars) and maize (*Zea mays*) resistant to both lepidopteran and coleopteran larvae (rootworms) have become widely used in global agriculture and have led to reductions in pesticide usage and lower production costs increasing the crop yield substantially (Toenniessen *et al*., 2003). Single gene based genetic engineering is having major concerns of development of insect resistance. Many strategies to insect resistance management particularly to *Bt* toxins were proposed and reviewed (Gould, 1998) and holds good for all other transgenic too. Several strategies including tissue or time-specific expression of toxins, gene stacking/ pyramiding, mixtures, rotation or mosaics of transgenic plants, combination of toxins with different modes of action, refuge strategy and "trap plants" strategy were proposed for mitigating the advent of resistance in insects. Gene stacking/pyramiding demands genes and promoters for resistance management. Hence isolation and deployment of genes differing for mode of action is need of the hour. Plant protease inhibitors (PIs) which are known for insecticidal activity should be explored to have sustainable transgenic technology. Since the PIs are ubiquitous, there is tremendous scope for finding specific PI which is toxic to a particular insect pest (Marchetti *et al*., 2000, Dunse *et al*., 2010). As genetic engineering of plants increases in complexity, an increased need for

**Author for Correspondence: E-mail: jagadeeshbt@gmail.com*

versatile promoters for direct expression of desirable transgenes in plants has arisen and indigenous to plant system promoters adds an advantage. So, In this study we have reported successful development of genomic library of Black gram variety Uttara, and isolation of protease inhibitor gene and 5' upstream sequences of the protease inhibitor gene.

Materials and Methods

Plant Material

Seeds of black gram (*Vigna mungo* L.) variety Uttara were collected from Indian Institute of Pulse Research, Kanpur. Seeds were sown in plastic trays and seedlings were raised in dark condition, and the 10 days old etiolated seedlings were used for genomic DNA isolation.

Construction of Genomic DNA Library

Isolation and Shearing of Genomic DNA

The good quality and high molecular weight DNA (~30 - 50 Kb) was isolated by the CTAB method (Doyle and Doyle, 1990). Around 25 µg genomic DNA of black gram was randomly sheared by passing it through a 200 µl small-bore pipette tip. The DNA aspirated and expelled from the pipette tip for about 100 times. Around 1-2 µl of sheared DNA was examined on a 20 cm agarose gel using the Fosmid Control DNA as a size marker.

End-Repair of the Insert DNA and Size Selection of the End-Repaired DNA

The end repair of the insert DNA was done using End-Repair Enzyme Mix according to the EPICENTER® Biotechnologies and size fractionation (more than 25 Kb) of end-repaired DNA using GELASE method as per the manufacturer's instructions. The concentration of extracted DNA was determined by running an aliquot of the DNA on an agarose gel using dilutions of known amounts of the Fosmid Control DNA as standard.

Ligation of the Insert DNA and Packaging of Recombinant Fosmid DNA

It is suitable to clone large fragments of genomic DNA in Copy Control™ Fosmid vector. The Copy Control Cloning System combines the clone stability afforded by single-copy cloning with the advantages of high yields of DNA obtained by "on-demand" induction of the clones to high-copy number. A 10:1 molar ratio of CopyControl PCC2FOS Vector to insert DNA was optimal. To ligate more than 25 Kb inserts to the Fosmid vector, 0.25 µg of insert for every 0.5 µg of vector was used. The ligation was incubated at room temperature for 4 hours.

Tittering the Packaged Recombinant Fosmid Clones

The titter of the phage particles (Packaged Copy–Control Fosmid Clones) was determined to calculate the number of plates and dilutions required to obtain a library. Different serial dilutions of the 1 ml of packaged phage particles were made with Phage Dilution Buffer (PDB) in sterile microfuge tubes. The infected EPI300-T1^R cells spread on an LB plate + 12.5µg/ml chloramphenicol and incubated at 37°C overnight to select for the CopyControl Fosmid clones. The colonies were counted and the titter of the packaged phage particles was calculated using following formula:

$$\frac{\text{(Number of colonies) x (Dilution factor) X (1,000 µl/ml)}}{\text{(Volume of phage plated [µl])}}$$

Screening of Genomic Library

The primary screening of genomic library for isolation of black gram PI gene and promoter was done using mung bean (*Vigna radiata*) PI as probe according to the methods of Davis *et al*. (1986) and Sambrook *et al*. (1989). The recombinant fosmid clones were inoculated in 96 well ELISA plates containing LB freezing media with chloramphenicol antibiotic (Master plate and replica plate) and incubated overnight at 37°C with shaking. Master plates were stored for alignment of positive clones and replica plates were centrifuged at 10,000 rpm for 5 min and pelleted colonies were dissolved in 20 µl of autoclaved milliQ water. About 5 µl of dissolved bacterial colonies were spotted onto Hybond-N+nylon membranes (Amersham) using 96 well grids plate and dried completely. Three solutions were used for washing of membranes after air drying, denaturation solution (1.5 M NaCl, 0.5 M NaOH), neutralization solution (1.5 M NaCl, 0.5 M Tris-HCl: pH 8.0) and buffer solution (2 x SSC). Each membrane was incubated and processed in a stepwise fashion; i.e., denaturation 2-5 min, neutralization 2-5 min and 2 x buffer solution 2-5 min with occasional gentle shaking. After washing, the membranes were dried again and crosslinked the DNA on to the membranes using a UV crosslinker. For primary and secondary screening the probes were labeled with P^{32} radioisotopes by random-primer labeling kit as per manufacturer's instructions (Stratagene, La Jolla, CA). Only those colonies which showed convincing hybridization signals on the autoradiogram were marked and the reference number of each colony was noted down.

Secondary screening was done to identify truly positive clones. The positive clones were further confirmed by PCR using mung bean protease inhibitor heterologous gene specific primers and subsequently sequenced (Chromous Biotech, Pvt. Ltd, India). The nucleotide sequence of genomic clone was analyzed by BLASTn algorithm of NCBI database. Further 5' upstream sequence of genomic clone was analysed *in-silico* by Place (Higo *et al*., 1999) and Plant Care (Lescot *et al*., 1999) databases.

Results and Discussion

Construction of Genomic DNA Library Isolation and End-Repair of the Insert DNA

For genomic library construction high molecular weight DNA isolation was very crucial. After passing genomic DNA through a 200 µl small-bore pipette tip for 100 times the correct size sheared DNA of more than 25 Kb was obtained. The End repair of the insert DNA generated blunt-ended 5'-phosphorylated DNA. The size-fractionation of the end-repaired DNA was done by running on a 20-cm long, 1% LMP (Low Melting Point) agarose gel at 30-35 V overnight (Fig. 1). This electrophoresis was performed in absence of ethidium bromide. The desired size DNA (≥25 Kb) was excised from the gel and extracted by GELase method.

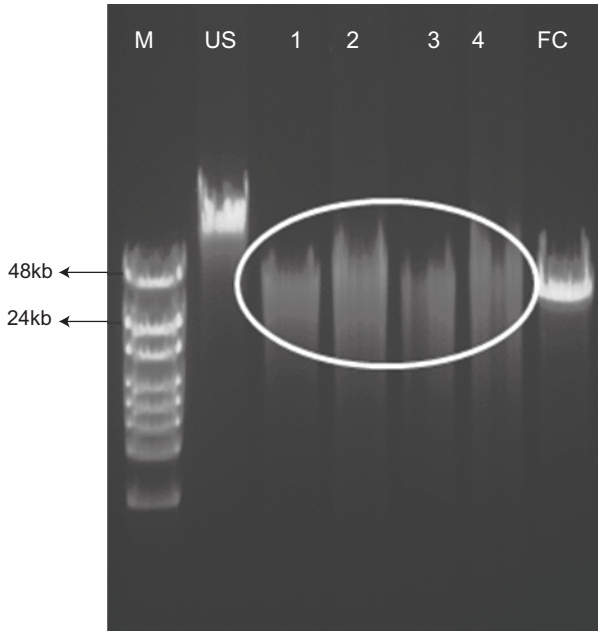

Fig. 1. Size fractionation of sheared genomic DNA

Ligation of the Insert DNA and Packaging of Recombinant Fosmid DNA

The gel extracted insert DNA was ligated with Copy Control™PCC2FOS Fosmid vector. Genomic library construction depends on efficient ligation. Since successful ligation can only be determined after infecting the phages with the host bacteria, higher the number of cfu/ml of the library, higher the efficiency of ligation. Using EPICENTRE® Biotechnologies Fast-Link DNA Ligase enzyme and buffer resulted in a titter of 3 x 10^5 to 1 x 10^6 cfu/ml, which was high. The recombined DNA was packaged using MaxPlax Lambda Packaging Extracts with high efficiency. Ligations were carried out at DNA concentrations of 0.2 µg/µl which favors concatemers and efficient packaging. The care was taken ensuring that the colony forming units were the same for every dilution. The titter of colonies of Uttara library was around 3 x 10^5 to 1 x 10^6 cfu/ml in different tubes with the same ligation product.

Screening of Genomic Library

The genomic library was screened by using mung bean PI heterologous gene specific probe. After primary screening positive signals were found and the reference number of each positive colony was traced out. The each positive colony lifted from the primary screening was passed through secondary screening with the same probe. After secondary screening each positive colonies were found to be very distinct and it was very easy to locate the individual positive colonies (Fig. 2). The positive clones were further confirmed by PCR using mung bean protease inhibitor heterologous gene specific primers and four PCR positive clones were obtained (Fig. 3). The PCR positive clones sequenced at Chromous Biotech, Pvt. Ltd, India. Based on a haploid genome size of 5.74 x 10^8 bp, the coverage of the library was estimated to be 4.5 haploid genome equivalents,

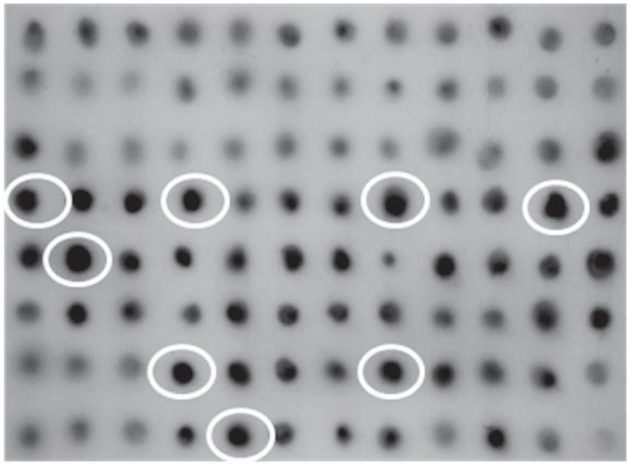

Fig. 2. Secondary screening of black gram genomic library

Fig. 3. Colony PCR confirmation of genomic clones

resulting in a 99.5 per cent probability of recovering any specific sequence (Clarke and Carbon, 1976).

Nucleotide Sequencing and Sequence Analysis

The NCBI Blastn algorithm analysis showed that 1514 bp nucleotide sequence of genomic clone contains partial coding sequence of black gram protease inhibitor gene (Fig. 4) and it showed 97 per cent homology with *Vigna radiata* trypsin protease inhibitor. Similarly it showed 93% homology with Cicer arietinum trypsin inhibitor. The 5' upstream sequence of genomic clone was analysed *in-silico* by Place and Plant Care databases. The 5' upstream sequence contains transcription control regions like TATA signals between 624-631 bases, GAAT and CAAT signals between 498-554 bases. Similarly MYB1AT dehydration responsive elements, a pollen specific expression element, endosperm specific expression elements and wound responsive elements were also found. The longest ORF analyzed in partial coding sequence was between 746 and 1414 bases in the second reading frame and the ATG for this ORF starting at 746 bp. This ORF on translation gave a protein with 222 amino acids (Fig. 4). The molecular weight of this protein was calculated to be 23.9 KD of which 21 amino acids were negatively charged, 16 positively charged with an isoelectric point of 5.0. The protein has an N-terminal signal sequence of 19 amino acids which is important in its translocation as it has been

```
 746 atgaaatccattgtattcctcctatttgcaattctaaccaatcct
     M  K  S  I  V  F  L  L  F  A  I  L  T  N  P
 791 ttatttgcattctccaataacaatgctattgagcaagtgttggat
     L  F  A  F  S  N  N  N  A  I  E  Q  V  L  D
 836 acaaatggtaacccccttatccctggtgatgaatactacatttttt
     T  N  G  N  P  L  I  P  G  D  E  Y  Y  I  F
 881 ccagcaagtgataaccctaaaactggaggactaaccctaaacaaa
     P  A  S  D  N  P  K  T  G  G  L  T  L  N  K
 926 attagtgatgcagagtgtcctgtaactgtcctacaaaataatgcg
     I  S  D  A  E  C  P  V  T  V  L  Q  N  N  A
 971 acaagaggtttaccagttaaattcaccttatcaggaagtaataac
     T  R  G  L  P  V  K  F  T  L  S  G  S  N  N
1016 actggtaataatatcttgactaatactgatcttgaaattgagttc
     T  G  N  N  I  L  T  N  T  D  L  E  I  E  F
1061 actaagaagccaaattgcgttgaatcatcaaaatggattgtgttt
     T  K  K  P  N  C  V  E  S  S  K  W  I  V  F
1106 gttgatgattttactcctcaaggttgtgttggtattggtggtcct
     V  D  D  F  T  P  Q  G  C  V  G  I  G  G  P
1151 gaaaatcatcttggtttagaaatactcaatggcaaatttttaatt
     E  N  H  L  G  L  E  I  L  N  G  K  F  L  I
1196 gtgagacatgcttctggatatgtttataggtttggatttgttttg
     V  R  H  A  S  G  Y  V  Y  R  F  G  F  C  L
1241 gatgtgagtggtgattgtggtttgcttggattgaatacttttgat
     D  V  S  G  D  C  G  L  L  G  L  N  T  F  D
1286 tcgagagaaggtggatcacgtttaatttaactatatttaattct
     S  R  E  G  G  S  R  L  I  L  T  I  F  N  S
1331 tataatgttgtatttgttgatgttgcttctgttaaatctggacgt
     Y  N  V  V  F  V  D  V  A  S  V  K  S  G  R
1376 attatgcctcttaaaggtggatttgatgcgtcagtttag 1414
     I  M  P  L  K  G  G  F  D  A  S  V  *
```

Fig. 4. The longest ORF of coding sequence of Black gram Protease Inhibitor gene (669 bp)

reported in many protease inhibitors such as soybean (Hammond *et al*., 1984), cowpea (Hilder *et al*., 1989), pea (Domoney *et al*., 1995) and maize (Rohrmeier and Lehle, 1993). Comparison for conserved domain in the protein by NCBI CD-search showed that the regions from 27-202 amino acids have the conserved domain similar to trypsin inhibitor Kunitz family of protease inhibitors. The active centre has four amino acid residues (from 85 to 93 amino acids) that are located within the conserved region (Fig 5).

In conclusion, a good quality of black gram Fosmid library representing 4.5 genome equivalents has been successfully constructed and characterized for isolation of protease inhibitor gene as well as promoter. Further this library will provide a good genomic resource for isolation of other agronomically important genes and promoters which can be further deployed for genetic transformation of crops for desired traits.

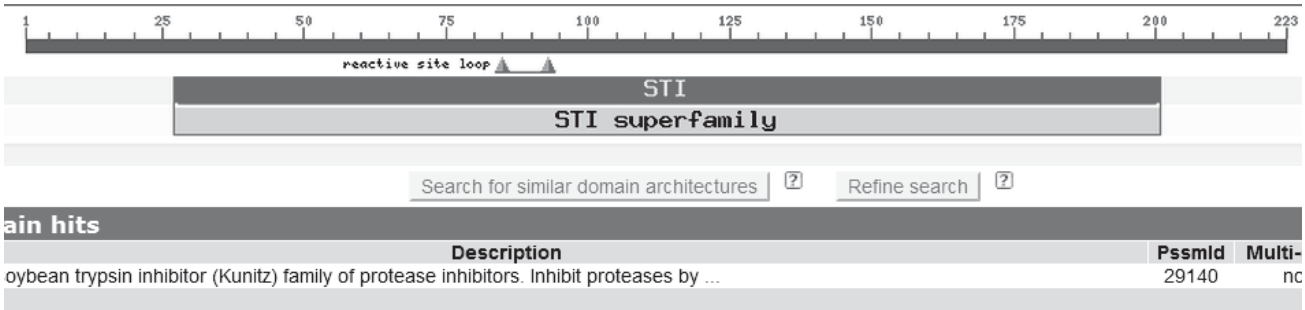

Fig. 5: Comparison for conserved domain of Protease Inhibitor gene using NCBI CD-search

References

Clarke L and Carbon J (1976) A colony bank containing synthetic Colel hybrids representative of the entire *E.coli* genome. *Cell* **9:** 91-99.

Davis LG, Dibner MD and Battey JF (1986) Basic methods in molecular biology. Elsevier Science Publishing Co. Inc. NY.

Domoney C, Welham T, Sidebottom C and Firmin, JL (1995). Multiple isoforms of Pisum trypsin inhibitors result from modification of two primary gene products. *FEBS Letters* **360:** 15-20.

Doyle JJ and Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus* **12:** 13-15.

Dunse KM, JA Stevensa, FT Laya, YM Gasparb, RL Heath and MA Anderson (2010) Coexpression of potato type I and II proteinase inhibitors gives cotton plants protection against insect damage in the field. *PNAS,* **107:** 15011-15015.

Gould F (1998) Sustainability of transgenic insecticidal cultivars: integrating pest genetics and ecology. *Ann. Rev. Entomol.,* **43:** 701-726.

Hammond RW, DE Foard and BA Larkins (1984) Molecular cloning and analysis of a gene coding for the Bowman-Birk protease inhibitor in soybean. *J. Biol. Chemistry* **259:** 9883-9890.

Higo K, Y Ugawa, M Iwamoto and T Korenaga (1999) Plant cis-acting regulatory DNA elements (PLACE) database. *Nucleic Acids Res.* **27:** 297-300.

Hilder VA, RF Barker, RA Samour, AMR Gatehouse, JA Gatehouse and D Boulter (1989) Protein and CDNA sequences of Bowman-Birk protease inhibitors from the cowpea (*Vigna unguiculata* Walp.). *Plant Mol. Biol.* **3:** 701-710.

Hilder VA, AMR Gatehouse, SE Sheerman, RF Barker and D Boulter (1987) A novel mechanism of insect resistance engineered into tobacco. *Nature* **300:** 160-163.

Joanitti GA, SM Freitas and LP Silva (2006) Proteinaceous Protease Inhibitors: Structural Features and Multiple Functional Faces. *Curr. Enzyme Inhibition* **3:** 199-217.

Koiwa K, RE Shade, K Zhu-Salzman, L Subramanian, LL Murdock, SS Nielsen, RA Bressan and PM Hasegawa (1998) Phage display selection can differentiate insecticidal activity of soybean Cystatins. *Plant J.* **14:** 371-379.

Lawrence PK and KR Koundal (2002) Plant protease inhibitors in control of phytophagous insects. *Electronic J. Biotechnol.* **5:** 93-109.

Lescot M, P Déhais, G Thijs , K Marchal, Y Moreau, Y Van de Peer , P Rouzé and S Rombauts (2002) Plant Care, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res*. **30:** 325-7.

Marchetti S, M Delledonne, C Fogher, C Chiaba, F Chiesa, F Savazzani and A Giordano (2000) Soybean Kunitz, C-II and PI-IV Inhibitor Genes Confer Different Levels of Insect Resistance to Tobacco and Potato Transgenic Plants. *Theor. Appl. Genet.* **101:** 519-526.

Rohrmeier T and L Lehle (1993) WIP1, a woundinducible gene from maize with homology to Bowman-Birk proteinase inhibitors. *Plant Mol. Biol.* **22:** 783-92.

Ryan CA (1990) Protease inhibitors in plants: genes for improving defenses against insects and pathogens. *Ann. Rev. Phytopathol.* **28:** 425-449.

Sambrook J, EF Fritsch and T Maniatis (1989) Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory Press, New York.

Srinivasan T, KR Rajesh Kumar and PB Kirti (2009) Constitutive expression of a trypsin protease inhibitor confers multiple stress tolerance in transgenic tobacco. *Plant Cell Physiol.* **50:** 541-553.

Telang M, A Srinivasan, A Patankar, A Harsulkar, V Joshi and A Damle (2003) Bitter gourd protease inhibitors: potential growth inhibitors of *Helicoverpa armigera* and *Spodoptera litura*. *Phytochemistry* **63:** 643-652.

Toenniessen GH, JCO Toole and J De Vries (2003) Advances in plant iotechnology and its adoption in developing countries. *Curr. Opin. Pla. Biol.* **6:** 191-198.

Urwin PE, CJ Lilley, MJ McPherson and HJ Atkinson (1997) Resistance to both cyst and root knot nematodes conferred by transgenic *Arabidopsis* expressing a modified plant cystatin. *Plant J.* **12:** 455-461.

Vain P, B Worland, MC Clarke, G Richard, M Beavis, H Liu, A Kohli, M Leech, J Snape and P Christou (1998) Expression of an engineered cysteine proteinase inhibitor (Oryzacystatin-I delta D86) for nematode resistance in transgenic rice plants. *Theor. Appl. Genet.* **96:** 266-271.