Identification of genetic diversity in pigeonpea \( [Cajanus cajan \text{(L) Mill sp.}] \) varieties through molecular markers

S Rajamani, Rachit Saxena, MV Ramana, Rani Chapara, M Sreekanth, and Rajeev K Varshney

Abstract
The legume crop Pigeonpea \((Cajanus cajan \text{(L) Millsp.)}\) which has drought tolerance belongs to \( Fabaceae \) family under the genus \( Cajanus \). It is mainly cultivated in the semi-arid tropics of Asia, Africa and America. In India, it is widely cultivated as a source of food for proteins, income and for soil improvement in intercropping systems. Natural out crossing leads to significant quality reduction and in turn lead to varietal contamination in redgram. To assess the diversity among nine pigeonpea varieties that are cultivated in Andhra Pradesh, the present investigation was carried out with a genetic fingerprint using 50 polymorphic SSR markers. With an average of 3.68 alleles per marker, a total of 184 alleles were observed. The maximum alleles eight were produced by CcM0481A range of 0.00 to 0.84 with an average of 0.52 was observed with Polymorphic Information Content (PIC). Three clusters were produced in neighbour-joining tree method. The variety LRG-223 was found in separate cluster (cluster I) among the nine varieties, while all other genotypes grouped into another major cluster (cluster II) indicating a distinct back ground of this variety.

Keywords: \( Cajanus cajan \text{(L) Millspaugh, SSR markers, Diversity, DNA fingerprinting} \)

Introduction
The diploid \((2n = 22)\) legume Pigeonpea \([Cajanus cajan \text{(L) Millspagh}]\) is an often cross-pollinated crop with a genome size of 858 Mbp and mainly grown in the tropical and sub-tropical regions of Asia and Africa. In food and nutritional security, it plays a major role as it a major source for proteins, minerals and vitamins. Pigeonpea seeds are mainly used in preparation of ‘dal’, and green pea vegetable as whole grain preparations. In animal feed the leaves, seed husks and pods are used and for firewood the stem and branches are used. Pigeonpea occupies 4.67 Mha in the world acreage with an annual production of 3.30 Mt. India is the largest consumer and producer of pigeonpea with the vernacular names ‘arhar’ and ‘toor’ with an annual production of 2.31 Mt, followed by Myanmar (0.60 Mt). To reveal genetic diversity there is need to analyze genetic relationships in species. In order to implement successful breeding programme there is an immense need to know the genetic diversity among the cultivars as it provides great and precise information on specific and targeted trait availability among the existing germplasm.

One of the useful and great tools for genetic diversity assessment of various crops is molecular markers. Simple sequence repeats (SSRs) are the most popular and more informative among the molecular markers since these can reveal more variation in most of the crops like, rice, wheat, maize, pea and also in pigeonpea. The predominantly cultivated traditional varieties yield less with lower seed quality due to contamination from pathogen propagules. Pigeon pea is an often-cross pollinated crop with an extent of 45% natural out crossing. This natural out crossing is the major source for varietal contamination that causes significant yield losses in farmer’s fields. As it is known that the genetic purity is an essential cause for higher yields there is an immense need to develop high genetic pure lines.

Domestication is the main reason to narrow down genetic variation in crop plants as it implies high continuous selection pressure for specific traits like yields.
Over time this lead crop plants more susceptibility to diseases and insect epidemics. Thus, it is very much essential to study the genetic composition of the existing germplasm that are used to develop modern day cultivars in comparison with their ancestors and related species which provides information on their phylogenetic relationship. This phylogenetic relationship is useful tool for finding of new and useful genes as the accessions with the most distinct DNA profiles are likely to contain greater number of novel genes.

Because of the abundance in genome wide distribution, reliable, reproducible and less cumbersome nature among different kinds of molecular markers, SSR markers are proven as the markers of breeder’s choice in practical breeding (Gupta and Varshney, 2000; Varshney et al., 2005) [5, 8].

Keeping above scenario in view this study was undertaken with detailed molecular and phenotypic characterization of pigeonpea varieties which are under cultivation.

Materials and Methods

Plant material: The nine redgram genotypes used under this study with different maturity groups Viz., LRG 223, LRG 275, ICPL 85063, ICPL 87119, LRG 187, LRG 105, LRG 41, LRG160 and LRG 52.

DNA isolation

From each genotype five individual seedlings of two weeks aged old were selected and DNA was extracted from leaves of those plants. To extract high quality DNA, leaf material after sampling was transported on ice and stored at -80 °C till subjected to DNA extraction. Approximately 200 mg genomic DNA was extracted according to the protocol of cetyltrimethyl ammonium bromide (CTAB) method (Murray and Thompson, 1986).

Reagents used in CTAB method were extraction TE Buffer: 10 mMTris-HCl (pH 7.6), 5 M NaCl, 1.5 M NaCl, 2% (w/v) CTAB, Buffer 100 mM Tris-HCl (pH 7.5), β-mercaptoethanol- added immediately before use, 25 mM EDTA 0.3% (v/v), RNase A stock solution (10 mg/mL), 95% ethanol (v/v), 70% ethanol (v/v), 0.1 mM EDTA.

Preparatory steps

Before grinding, 95% ethanol solution and the mortar and pestle (to minimize frozen tissue thawing) at -20 °C were pre chilled and before beginning the extraction the water baths were pre-heated at 65 °C and 37 °C. For 1 g of leaf tissue, prepared 10 ml extraction buffer in a 50 ml Falcon tube by adding 0.3% (v/v) β-mercaptoethanol and pre-heated in the 65 °C water bath.

Grinding and tissue disruption

Using liquid nitrogen, 100mg of frozen leaf tissue was grounded into a fine powder. In the new 1.5ml effendorf tube the grounded sample powder was placed and was mixed in the pre-heated extraction buffer. After that the samples were kept for incubation into the water bath at 65 °C and mixed by inversion for every 10 min up to 30 minutes to 1 hour. Centrifuged the sample tube for 5 min at 5000×g after incubation to pellet and remove un-lysed leaf tissue flowed by decanted the supernatant into a new 1.5 ml effendorf tube.

Protein Extraction and RNase treatment

One volume of chloroform: isoamyl alcohol (24:1) was added to the solution and mixed by inversion for 5 min. Later centrifuged the sample for 10 min at 5000xg and the upper aqueous phase was pipetted into a new effendorf tube, by taking care to avoid the aqueous/organic layer interface. To this solution added 10μL of RNaseA (10 mg/ml) and incubated at 37 °C for 30min with periodic, gentle mixing. After incubation, added 1 volume of chloroform: isoamyl alcohol (24:1) to the solution and mixed by inversion for 5 min followed by centrifugation for 10 min at 5000xg and pipetted the aqueous phase into a new Falcon tube, again taking care to avoid the organic layer.

Precipitation

Half volume of 5 M NaCl added to the sample and mixed gently by inversion followed by three volumes of cold 95% ethanol and mixed gently by inversion. After that placed the tubes into a -20 °C freezer and incubated for 30min to 1 hour. As both the CTAB and NaCl could be precipitated from solution and prevent DNA isolation hence, caution was taken not to leave the sample at -20 °C for more than 1 hour.

To pellet the DNA the effendorf tubes were centrifuged for 10 min at 5000xg after incubation. Decanted away the supernatant carefully and washed the DNA pellet with 1ml of 70% ethanol. Swirled the solution gently and centrifuged again for 10 min at 5000xg. After that the supernatant was decanted carefully and at room temperature kept DNA pellet for air dry for 15 min. After proper drying, DNA was suspended in 100-200 μl of TE buffer.

DNA quality and quantity assessment

After DNA extraction, added 1 μl RNase solution (10 mg/ml) followed by set aside the tubes in water bath at 37 °C 30 min for degrading the RNA. By using a NanoDrop UV/Vis spectrophotometer and 0.8% (w/v) agarose gel both the absorbance ratio of 1.8-2.0, and observed no evidence of substantial band shearing or contamination (either RNA or polysaccharide). The fluorescent phosphoramidite dye 6-carboxy-fluorescein (FAM) was used to label the 5′ end of all forward primers. PCR reaction volume was prepared to 20 μL, consisting of genomic DNA (2 μl, 25 ng), each primer (0.3 μL, 20 μM), Ex-Taq Buffer (2.0 μL) and Ex-Taq DNA polymerase (0.2 μL, 1.0 U). PCR thermal cycling was programmed for an initial denaturation step at 94 °C for 5 min, followed by 32 cycles of (30 s at 94 °C; 35 s for annealing and 40 s at 72 °C), and a final extension of 3 min at 72 °C.

SSR Markers

In pigeonpea diversity studies a large numbers of SSR markers were developed and applied (Burns et al., 2001; Odeny et al., 2009; Saxena et al., 2010) [1, 12, 3, 4, 14]. Unlike AFLPs, the abundance and co-dominance nature of SSR markers across the genome were reported by many researchers. It was also reported that SSRs were multi-allelic, open to high-through put applications and can detect more polymorphism (Gupta and Varshney 2000; Salgado et al., 2006) [5, 6]. Due to their robustness, SSRs were highly useful in assessing genetic purity and also used in characterising different pigeonpea hybrids (Saxena et al., 2010; Datta et al., 2010; Upadhyaya et al., 2011) [4, 7, 9]. SSRs are tandem repeats and are usually two to six bases long and occurring abundantly in a genome. Usually the diversity at the SSR loci is due to the variable number of repeat units. During DNA synthesis this variation was caused by slip-strand mispairing, and resulted in a gain or loss of one or more repeat units (Semagn et al., 2006) [10].

"2755"
Primers were designed to flank the SSR loci to allow PCR amplification. The amplification products were separated by capillary electrophoresis, which incorporates fluorescence detection systems. Usually the forward primers were synthesized with a fluorochrome attached to the 5’ end for separation with fluorescent detection systems. In the PCR step, alternatively two different forward primers were used. The first primer was designed to contain an M13 sequence at the 5’ end, in addition to the unique primer sequence and was used in the first few PCR cycles (Shuelke, 2000) [11].

Many more studies have been led to development and utilization of many SSR markers in pigeonpea diversity analyses (Burns et al., 2001; Odeny et al., 2007; Odeny et al., 2009) [1, 12, 13, 14].

### Results and Discussion

For any successful breeding programme, it is essential to have information regarding genetic variation in the available material for selection of suitable genotypes for including them in breeding programme. In the present study fifty SSR primers were selected to know the diversity (Table 1). Among these primer pairs, twenty-four contain di-nucleotide motif repeats, nine contained tri-nucleotide motif repeats and seventeen contain composite nucleotide repeats as listed in Table 1.

The PIC values indicated the effective number of alleles that could be detected per marker in a set of individuals. In the present study, the PIC values were ranged from 0.73-0.88 with an average of 0.77. A greater PIC value of 0.95 was found with the primer pair CcM0257, while CcM0402 primer pair recorded the minimum PIC value of 0.73.

### Table 1: Details of 50 genic-SSR loci showing polymorphism among pigeonpea cultivars

| S No | Marker name | Forward primer (5’-3’) | Reverse primer (5’-3’) | Number of alleles | PIC values |
|------|-------------|------------------------|-----------------------|-------------------|------------|
| 1    | CcM0021     | TGAATTTTTTCCAGGATTTTACA | GCGCAATAATAAAGGCCAG | 280               | (TAA)10    |
| 2    | CcM0057     | CAATGTTGCAATAGAACCA   | GTTCATAAAATCCTGGGCA  | 269               | (AAT)6     |
| 3    | CcM0126     | TGGTCAATTCCTCCTCA    | CCAATGAAATAGAACCACTCA| 218               | (TAT)12    |
| 4    | CcM0133     | TGGTCAATTCCTCCTCA    | CCAATGAAATAGAACCACTCA| 176               | (TAT)12    |
| 5    | CcM0193     | TAATACACCCCTTTGAGGA | TGCAAAACACATCTTTGGAA| 190               | (TA)20     |
| 6    | CcM0195     | CCAATAAAGCATAAAACCCACA | TACGCTAGATTTGCTATGAGA | 223               | (AT)11     |
| 7    | CcM0207     | TTTTGGCGGTACTTATAACC | TTATGCGGGAACACACTGA | 235               | (TA)15     |
| 8    | CcM0246     | ATGAGCTCAACGATCAAGA | ATGAAAGAAACATACGCGG | 226               | (AT)15     |
| 9    | CcM0252     | CATAAGAAGACCTCTCACA | CGTTCAGCAAAGAAAGAGA | 234               | (AT)23     |
| 10   | CcM0295     | GCCATGAGGAGGAGATGAA | CTCTCTCAAGGAGACCTGGA | 241               | (AG5TG)15  |
| 11   | CcM0303     | CACGGCTTGGATGCGACA | TACAGCAGAAATTCACAGC | 239               | (TG)10     |
| 12   | CcM0306     | TACGCTGTACCTCTCTACATT | TACGCTGCAATACGGCATA | 280               | (TA)5      |
| 13   | CcM0402     | CCAATGTTGCAAGGAGACCG | GCAGATCCTCAATCTCCTCC | 179               | (TGA6AG)5  |
| 14   | CcM0431     | CCAATACCTCAATCCTAAACA | TCAATGAAACGGCCATGAAA | 126               | (CATA5A5)9 |
| 15   | CcM0471     | AATTTTTACCTACATCCACCTAAAA | TTTTAGCGCATATTTGACATTT | 275               | (AT)12     |
| 16   | CcM0473     | GGCGGTCTTGGAGGATGCT | AAAGAAGAAGATACGGGAAAT | 261               | (AAT10)10 |
| 17   | CcM0481     | CTCCCAAAATGCTCCACCA | TTCTTGGCGGTGAGGACGT | 245               | (AT)12     |
| 18   | CcM0493     | AGCGAAATGACCTCGGCACT | GCCTGAGTACCTCGTGCC | 191               | (TC)14     |
| 19   | CcM0516     | ATGTCTGTGTTGTTTCGACA | TCTCGTACATCTCAGTTC | 190               | (TA)17     |
| 20   | CcM0522     | TTGCTGCTTTGGTCTATGAG | AGTCACTACCTACACATCA | 117               | (TA21)11   |
| 21   | CcM0594     | GCTTTCAGTCTCTCTGAGTG | GAACTTGACTTCATTTCC | 185               | (GA9GT)9   |
| 22   | CcM0673     | TGACCAAGCCTACTACCAA | CATGCGACGACAGGAATC | 272               | (AT)6AG9   |
| 23   | CcM0724     | AGTTCCTATACATCTACCAACAAC | CAATATGAAGATCGACATTT | 275               | (AT)10     |
| 24   | CcM1010     | TGTGCCTCGTATCACAAGAC | GCACGCTTTCCTGTAGTGG | 245               | (TA)12     |
| 25   | CcM1011     | ATTGGTGGTGGCTGACACG | GGACTCTTTCTCAGTCCTG | 170               | (TGC5GT)14 |
| 26   | CcM1118     | ATGTGCGTGGCTGACAGCC | GGAATGCGTCACCTAAA | 277               | (AG)9GT19  |
| 27   | CcM1175     | TTGTCCTTCTATCCGCAAC | TGCGAGAGCATTATACCTTTCA | 204               | (TA)20     |
| 28   | CcM1182     | GCCTGCCATGGCTTACTAGA | GCCCTACATGTTCCATGTC | 270               | (GA)6      |
| 29   | CcM1232     | GCTGAGGCTCCTGCTAGTCA | GCCCTACATGTTCCATGTC | 270               | (GA)6      |
| 30   | CcM1251     | CAATGAGCAGACGACGACG | CGGAGATTGCTGGTCTTTTCT | 228               | (CCA)9     |
| 31   | CcM1282     | CCAATGCTGATACATGTATGAAACA | CCGATCTGTCCAAATACCGAIG | 257               | (AT)18     |
| 32   | CcM1439     | TGAGGCGTACCTCAGAA | CAAGATGCAAGATAAAAGACAA | 270               | (C11T2C5)10 |
| 33   | CcM1506     | CGCTTTCTGGAGGAGAAGG | ATCCCAATACCTCAGAATG | 263               | (G13)10TG5 |
| 34   | CcM1508     | CGCTTTCGAGGACTGAGGAGAAG | TTGCAATTACCTCAGAATG | 127               | (TG)5     |
| 35   | CcM1825     | TGAATGTTGCGGAAACACAT | TGCGAGAGGAAACAATACCTG | 263               | (G13)10TG5 |
| 36   | CcM1866     | GAACCTCCTTAAGTTTCTGC | TTGCTGATGCTTGGGGCAA | 127               | (TG)6     |
| 37   | CcM1895     | GGGAGAGAGGCAGGCAGAGGT | GTTGACGAGGATGTGGGGGA | 272               | (AG)12    |
| 38   | CcM2379     | CCGGAAAATGCTCATTGGA | TGCTAGCAGAATATTTCGTTG | 151               | (TC)10     |
| 39   | CcM2453     | CCAAAATTAGGGAGGACCT | TGCTCGCAGATGCTGCTTG | 229               | (TA)12     |
| 40   | CcM2672     | TGGTCGTGGAGGAAGCT | TTGTCGAGCAGAGGGCA | 226               | (AG)10     |
| 41   | CcM2704     | AAAATTGCTGAATGGATATGTTGGA | TGGGATCATATGCCTCAACTCA | 127               | (AT)10     |
| 42   | CcM2751     | CTGGAGCTATGGTCACGACA | TAAAGGCAATCAAGGTGGGCA | 205               | (TA)13     |
| 43   | CcM2781     | TGTTAGTGCTCAGGATCACCTT | AAGTGTGATCGTCAATCTTAAATG | 221               | (CTT1TAT)6 |
| 44   | CcM2948     | CCAGAGACCTTCAGAAGGCG | TGAGGAAATTGCGGCGTCT | 220               | (TAA8GAT)5 |
| 45   | CcM2982     | CCAGAGACCTTCAGAAGGCG | TGAGGAAATTGCGGCGTCT | 220               | (TAA8GAT)5 |
Allele scoring and analysis

After PCR, the amplified DNA fragments were successfully separated by capillary electrophoresis on an ABI3730 automatic sequencer. Gene Mapper® software was used to score the allele (fragment) sizes as presented in Figure 1.

Dendrogram Analysis

Dendrogram Analysis was done using Darwin software. By using similarity coefficient values a dendrogram was constructed (Fig. 2) and all the nine varieties were grouped into two major clusters. Among the nine varieties, LRG-223 was fallen in to a separate cluster (cluster I), while rest of varieties were grouped into two major clusters indicating a distinct background of this variety. Again, these major clusters were divided into two sub clusters. The sub cluster A1 exhibited two sub clusters having 3 varieties LRG 275, ICPL 85063 and ICPL 8711 (independent from LRG 275 and ICPL 85063). These varieties were genetically closely related as these were grouped under sub cluster A1. Two sub-sub clusters namely A2-1 and A2-2 were formed from sub cluster A2. The varieties LRG 105 and LRG 187 were fallen under the cluster A2-2 while LRG 41, LRG 160 and LRG 52 were under A2-2 cluster. LRG 223 and LRG 52 were the varieties located at the two extremes of the dendrogram.

Based on the present set of markers used in the present investigation the varieties ICPL 85063, ICPL 8711 from Hyderabad LRG 275 from Lam, Guntur of sub cluster A1 showed maximum similarity. The variety LRG 275 of sub cluster A1 belongs to Lam, Guntur with different ecological region exhibited close association with the varieties ICPL85063, ICPL 8711 from Hyderabad. The geographical isolation is not only the factor for genetic diversity. This was demonstrated from the inclination of varieties in clusters irrespective of geographic boundaries. Despite of their common origin i.e. Lam, Guntur, the varieties LRG 187 and LRG 105 from cluster A2-1; LRG-41, LRG 160 and LRG 52 (sub cluster A2-2) exhibited diversity at genetic level. The variety LRG 223 exhibited less similarity with other varieties and was genetically more distinct and diverse. The present polymorphism study using random SSR markers was analysed and understood effectively about the extent and distribution of the genetic variation available within the redgram varieties.

Fig 2: Neighbour joining Tree showing genetic relationship

| LRG 41          | LRG 105          |
|-----------------|------------------|

Computer screen shot of GeneMapper® peaks

Marker: CcM0471
LRG 223
Fig 1: Gene Mapper profile for an amplified SSR marker showing polymorphism

Author’s contribution
Conceptualization of research (S. Rajamani); Designing of the experiments (S. Rajamani); Contribution of experimental materials (S. Rajamani); Execution of field/lab experiments and data collection (S. Rajamani, M.V. Ramana, M. Sreekanth, Rachit Saxena and Rajeev Varshney); Analysis of
Conflict of interest
The authors declare that they have no conflict of interest.

Acknowledgments
The authors acknowledge technical assistance from Dr. Rachith Saxena, ICRISAT, Hyderabad.

References
1. Burns MJ, Edwards KJ, Newbury HJ, Ford-Lloyd BV, Baggott CD. Development of simple sequence repeats (SSR) markers for the assessment of gene flow and genetic diversity in pigeonpea (Cajanus cajan). Mol Ecol Notes. 2001; 1:283-285.
2. Odeny A, Jayashree B, Ferguson M, Hoisington D, Crouch J, Gebhardt C et al. Development, characterization and utilization of microsatellite markers in pigeonpea. Plant Breed. 2007; 126:130-136.
3. Odeny DA, Jayashree B, Gebhardt C, Crouch J. New microsatellite markers for pigeonpea (Cajanus cajan (L.) millsp.). BMC Res Notes. 2009; 2:35.
4. Saxena RK, Prathima C, Saxena KB, Hoisington DA, Singh NK, Varshney RK et al. Novel SSR markers for polymorphism detection in pigeonpea (Cajanus spp.). Plant Breed. 2010; 129:142-8.
5. Gupta PK, Varshney RK. The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica, 2000; 113:163-185.
6. Salgado KCPC, Vieira MGGC, Pinho EVR, Guimaraes CT, Pinho RGV, Souza LV. Genetic purity certificate in seeds of hybrid maize using molecular markers. Revista Brasileira de Sementes. 2006; 28:169-175.
7. Data J, Lal N, Kaashyap M, Gupta PP. Efficiency of Three PCR based Marker Systems for Detecting DNA Polymorphism in Cicer arietinum L and (Cajanus cajan (L) Millsp.). GEBJ, 2010, 5.
8. Varshney RK, Sigmund R, Borner A, Korzun V, Stein N, Sorrells ME et al. Interspecific transferability and comparative mapping of barley EST–SSR markers in wheat rye and rice. Pl. Sci. 2005; 168:195-202.
9. Upadhyaya HD, Reddy KN, Sharma S, Varshney RK, Bhattacharjee R, Singh S et al. Pigeonpea composite collection and identification of germplasm for use in crop improvement programmes. Plant Genetic Resources. 2011; 9:97-108.
10. Semagn K, Bjonstad A, Ndjjondjop MN. An overview of molecular marker methods for plants. Afri. J of Biotech. 2006; 5(25):2540-2568.
11. Shuelke M. An economic method for the fluorescent labeling of PCR fragments. Nature Biotechnology. 2000; 18:233-234.
12. Burns MJ, Edwards KJ, Newbury HJ, Ford LBR, Baggot CD. Development of simple sequence repeat (SSR) markers for the assessment of gene flow and genetic diversity in pigeonpea (Cajanus cajan). Molecular Ecology Notes. 2001; 1:283-285.
13. Odeny DA. The potential of pigeonpea (Cajanus cajan (L.) Millsp.) in Africa. Natural Resources Forum. 2007; 31:297-305.
14. Odeny DA, Gebhardt C. Towards molecular breeding of pigeonpea (Cajanus cajan (L.) Millsp.): a case for increased production in Africa. Acta Horticulture (ISHS). 2009; 806:151-154.