ONLINE RESOURCES

Novel genic microsatellite markers from *Cajanus scarabaeoides* and their comparative efficiency in revealing genetic diversity in pigeonpea

SUBHOJIT DATTA[1,2*], PALLAVI SINGH[1], SAHIL MAHFOOZ[1,3], PRAKASH G. PATIL[1], ARVIND K. CHAUDHARY[1], IKECHUKWU O. AGBAGWA[1,4] and N. NADARAJAN[1]

[1]*Indian Institute of Pulses Research, Kanpur 208 024, India*
[2]*Department of Plant Pathology and CAES Genomics Facility, University of California, Davis, CA 95616, USA*
[3]*National Bureau of Agriculturally Important Microorganisms, Mau Nath Bhanjan 275 101, India*
[4]*University of Port Harcourt, East/West Road PMB 5323 Choba, Rivers State, Nigeria*

[Datta S., Singh P., Mahfooz S., Patil G. P. , Chaudhary A. K., Agbagwa I. O. and Nadarajan N. 2013 Novel genic microsatellite markers from *Cajanus scarabaeoides* and their comparative efficiency in revealing genetic diversity in pigeonpea. *J. Genet*. **92**, e24–e30. Online only: http://www.ias.ac.in/jgenet/OnlineResources/92/e24.pdf]

Introduction

Paucity of molecular markers is hindering molecular breeding programmes for genetic improvement in pigeonpea, which is considered to be among the richest source of dietary protein in Asia and Africa. At the time of the start of this study, only 156 microsatellite markers were available in pigeonpea (Burns *et al.* 2001; Odeny *et al.* 2007, 2009). Recently with the publication of draft genome sequence and deep transcriptome studies, the stage has been set to enrich genomic resources to aid molecular breeding in pigeonpea (Dutta *et al.* 2011; Singh *et al.* 2012; Varshney *et al.* 2012). Genic microsatellites or EST-SSRs (simple sequence repeats) derived from expressed sequence tags (ESTs) are useful because these are inexpensive to develop, represent transcribed genes, and often a putative function can be assigned to them.

Compared with genomic sequences, genic SSRs have several advantages as genetic markers. First, if an EST marker is found to be genetically associated with a trait of interest, it may represent the gene affecting the trait directly (Chen *et al.* 2001; Thiel *et al.* 2003). Therefore, EST-derived markers can provide opportunities for gene discovery and enhance the role of genetic markers by assaying variation in transcribed and known-function genes. Second, EST-derived markers are likely to be more conserved and therefore may be more transferable between species than markers derived from genomic sequences (Cordeiro *et al.* 2001; Decroocq *et al.* 2003; Varshney *et al.* 2005). Third, ESTs that show homology to candidate genes can be specifically targeted for genetic mapping and can be useful for aligning genome linkage maps across distantly related species for comparative analysis (Holton *et al.* 2002). Because of these advantages, increasing numbers of genic SSR markers are now being identified and used for a variety of applications in a number of plant species, such as grape (Scott *et al.* 2000), sugarcane (Cordeiro *et al.* 2001), and cereals such as wheat, barley, rye and rice (Varshney *et al.* 2005). Wild relatives of any crop are important sources of resistance to biotic and abiotic constraints. *Cajanus scarabaeoides*, an Indian-origin wild relative of pigeonpea (*Cajanus cajan*), has many desirable characters and is cross-compatible with cultivated pigeonpea, and interspecific gene transfer is possible through conventional hybridization. With an objective to develop and utilize genomic tools for its improvement, a total of 350 EST clones from the root tissue of this wild relative were developed and sequenced. Sixty-three EST sequences were submitted to NCBI. Twenty-six primer pairs were also designed based on sequences flanking the microsatellite motifs using PRIMER v. 3.0 software. The primer pairs were amplified on 45 selected pigeonpea genotypes and polymorphic SSR markers were identified. These new markers will add to the pool of available markers and accelerate molecular breeding for pigeonpea improvement.

*For correspondence. E-mail: subhojit@email.com.

Keywords. genic SSR; diversity; EST sequences; *Cajanus scarabaeoides*.

Materials and methods

Cajanus scarabaeoides (ICRISAT Genebank accession no. ICP15697) cDNA libraries were constructed and ESTs were generated. Total RNA was isolated from 10-day-old root tissue with TRIZOL reagent (Invitrogen, Carlsbad, USA). Double-stranded cDNA was produced from 5 mg of total RNA using the SMART PCR cDNA construction kit (Clontech, Palo Alto, USA) following the manufacturer's instructions. Seven μL of the cDNA product (~110 ng DNA) was used to clone the DNA in pJET1.2 vector (Fermentas, Vilnius, Lithuania) following overnight ligation. The ligation mix was used to transform chemically competent XL1-Blue *Escherichia coli* cells (Stratagene, La Jolla, USA). The transformants were spread on LB agar plates containing 100 mg/mL ampicillin and 30 mg/mL tetracycline for direct picking. Ovenight-grown white colonies (recombinant clones) were picked up into a 96-well plate containing LB broth with ampicillin and tetracycline and grown overnight at 37°C on a rotary shaker at 220 rpm. Glycerol stocks in 96-well format were prepared by combining 38 μL of 60% (v/v) glycerol with 150 μL of culture and frozen at −80°C.

Plasmid DNA was extracted from around 350 randomly selected clones using Qiagen Turbo 96 kits (Qiagen, Hilden, Germany). Purified plasmid DNA was digested with *Bgl*II (Fermentas, Vilnius, Lithuania) and plasmids with average insert sizes of about 1 kb were used for sequencing. DNA sequencing was performed by Bangalore Genei (Bangalore, India) using the standard pJET 1.2 forward primer.

The FASTA files containing the raw sequences were edited using DNASTAR software to remove the vector sequences. Each sequence was trimmed of portions of vector sequence and poly (A/T) tail. EST sequences were submitted to NCBI under accession numbers GR979561–GR979623 and also analysed for homology in GenBank nucleotide and nonredundant protein database. A match was considered significant when the score was higher than 100 (optimized similarity score) with E values $<= 10^{-10}$. Novel ESTs were identified by comparison with sequences in nonredundant nucleotide and EST databases of GenBank using BLASTN algorithm.

Sixty-three EST sequences were examined for SSR mining and marker development. The online software WEBSAT (http://wsmartins.net/websat/) was used for the purpose. The criteria for SSR identification were seven, five, four, three repeat units for dinucleotides, trinucleotides, tetranucleotides, pentanucleotides and higher-order nucleotides, respectively. The designed primers were of 19–25 bp length, annealing temperature was 50–60°C, and product sizes were 100 to 400 bp.

For marker validation and diversity study, a total of 45 genotypes including 34 from cultivated species (*C. cajan*) and 11 genotypes from wild species (table 1) were grown in greenhouse. DNA was isolated from young leaves using modified CTAB method (Abdelnoor *et al.* 1995) and purified with RNAase treatment (10 μg/mL) followed by treatment with phenol: chloroform : isoamyl alcohol (25:24:1). Finally DNA was diluted to a concentration of 5 ng/μL TE

Table 1. Genotypes used in the study.

	Pigeonpea genotypes	Species		Genotype	Species
Cultivar					
1	UPAS 120	*C. cajan*	24	IPA - 16F	*C. cajan*
2	Type 7	*C. cajan*	25	IPA - 70	*C. cajan*
3	ICP 8863	*C. cajan*	26	ICP - 7626	*C. cajan*
4	ICPL 87119	*C. cajan*	27	ICP - 8840	*C. cajan*
5	BDN 2	*C. cajan*	28	IPA - 69	*C. cajan*
6	BSMR 853	*C. cajan*	29	ICP - 10958	*C. cajan*
7	MAL 13	*C. cajan*	30	ICPL - 20102	*C. cajan*
8	PUSA 9	*C. cajan*	31	ICPL - 20107	*C. cajan*
9	DA 11	*C. cajan*	32	ICPL - 20116	*C. cajan*
10	NDA 1	*C. cajan*	33	ICPL - 20125	*C. cajan*
11	MA 6	*C. cajan*	34	ICPL - 20135	*C. cajan*
12	MA 3	*C. cajan*		Wild relatives	
13	GT 1	*C. cajan*	35	ICP-1629-1	*C. cajanifolius*
14	PUSA 992	*C. cajan*	36	ICP-1629-2	*C. cajanifolius*
15	Paras	*C. cajan*	37	ICP- 15685	*C. scarabaeoides*
16	PUSA 33	*C. cajan*	38	ICP-15724	*C. scarabaeoides*
17	ICPL 87	*C. cajan*	39	ICP-15697	*C. scarabaeoides*
18	GS 1	*C. cajan*	40	ICP 15661	*C. platycarpus*
19	Manak	*C. cajan*	41	ICP 15760	*C. sericeus*
20	GT 100	*C. cajan*	42	ICP 15761	*C. sericeus*
21	Bahar	*C. cajan*	43	ICP 15666	*C. platycarpus*
22	IPA – 8F	*C. cajan*	44	ICP15624	*C. albicans*
23	IPA – 15F	*C. cajan*	45	ICP 15622	*C. albicans*

and used for marker evaluation and diversity study. PCR was performed in a 5.0-μL reaction volume containing 1× PCR buffer (10 mM Tris HCl pH 9.0, 1.5 μM $MgCl_2$, 50 mM KCl, 0.01% gelatin), 0.2 mM each of dNTP (Bangalore Genei, Bangalore, India), 0.1 U of *Taq* DNA polymerase (Bangalore Genei) and 5 pmol each of forward and reverse primers and 5 ng of template DNA. PCR cycle comprised an initial denaturation for 3 min at 95°C followed by 40 cycles of 94°C 20 s, annealing for 20 s (temperature depending on marker) and 72°C for 20 s, and a final extension was carried out at 72°C for 5 min.

All PCR amplicons were resolved by electrophoresis on 3% agarose gel and visualized with ethidium bromide under UV illumination to identify the informative SSR loci across the selected genotypes. GeneRuler 100-bp DNA ladder (MBI Fermentas) was used to estimate allele size (figure 1). The gels were run for 4 h at 45 V and SSR fingerprint profiles were recorded using BioRad Gel Doc XR v. 2.0. The SSR bands were scored as present (1) or absent (0) for each primer–genotype combination and a binary raw matrix was generated for all banding patterns. The amplification data generated by microsatellite markers were analysed using SIMQUAL subprogram of NTSYS-PC software v. 2.1 (Rohlf 1998) to generate Jaccard's similarity coefficient (Jaccard 1908). These similarity coefficients were used to construct a dendrogram depicting genetic relationships among the cultivars by employing the unweighted pair group method of arithmetic averages (UPGMA) algorithm and SAHN (sequential, agglomerative, hierarchical and non-overlapping) clustering. Polymorphism information content (PIC) was calculated for each marker using the following equation:

$$\text{Polymorphism information content } (\text{PIC}_i) = 1 - \sum_{j=1}^{n} P_{ij}^2,$$

where P_{ij} is the frequency of the jth allele for the ith marker, and summation is over n alleles.

Results and discussion

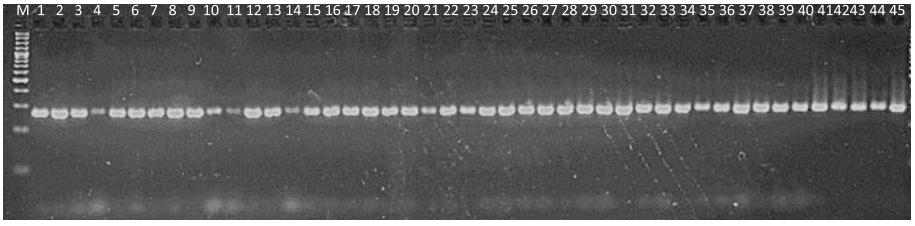
The SSRs identified in the present study could be characterized by many different types of repeat motifs. A total of 153 different repeat motifs were found during the search which included 72 mononucleotide repeats (MNRs), four dinucleotide repeats (DNRs), two trinucleotide repeats (TNRs), four tetranucleotide repeats (TTNRs), 44 pentanucleotide repeats (PNRs) and 27 hexanucleotide repeats (HNRs).

Out of 63 nonredundant SSR-ESTs, only 26 were used for primer design. The remaining ESTs were inappropriate for primer design mainly due to short unique domains flanking the microsatellite core. Details of primer sequences and expected product sizes with SSR motifs are given in table 2. This suggests that ~41% of the total ESTs investigated in the study represent potential candidates for SSR-marker development.

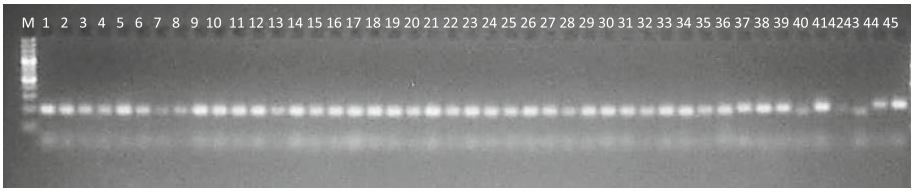
Moreover, based on BLASTX analysis, a putative function could be assigned to 18 (69.2%) potential markers assuming a threshold of $<1.00\ E^{-05}$ (table 2). Most of the similarity search results showed homology with *Ricinus communis*, *Arachis hypogea* and *Lotus japonicus*.

Of the 26 primer pairs tested on 45 pigeonpea genotypes, only 20 gave good amplification for all the genotypes, whereas six markers did not produce amplification at all. Of the 20 working primer pairs, 10 gave PCR products of the expected sizes, and nine primer pairs resulted in PCR products larger than expected, and PCR products from the remaining one primer pairs were smaller than expected.

Eight (40%) of the 26 SSRs markers were polymorphic among the 45 genotypes. PIC ranged from 0.05 to 0.7 with



a. CS 21



b. PP 1

Figure 1. SSR profile of a genic SSR (a) and a genomic SSR (b). Lane M, 100-bp DNA ladder; lanes 1 to 45, pigeonpea genotypes as list in table 1.

Table 2. Genic SSR markers and BLASTX similarity results.

	GenBank accession no.	Primer sequences (5-3')	SSR motif	T_m (°C)	Expected size (bp)	Putative function
1	GR979561	F: GTTGACCCGAGGCTTTTATATGT	$(TGTGT)_2$	60.6	185	Peptidyl-prolyl cis-trans isomerase
		R: AGACTCCATTCACTCTATGAACCAC		59.9		
2	GR979561	F: GTGTCGTTTTACCTTTTCTCTTGTC	$(CCCTT)_2$	59.6	131	Peptidyl-prolyl cis-trans isomerase
		R: ACAGACTCCATTCACTCTATGAACC		59.9		
3	GR979563	F: GGGCCTGACTATTATTATTGACTTG	$(CAAAA)_2$	59.4	335	No homology
		R: GGACCAACTCCTTCAACTTCAG		59.3		
4	GR979561	F: GTACAGCACATAGCTTTCTCAACAA	$(CGGGT)_2$	60	385	
		R: ACCAACACCTTCATCACTTTCTTC		60.7		
5	GR979566	F: GGGGAAACAGGTAAAGGGG	$(CCCGG)_2$	60.5	257	SH3 family domain protein
		R: ATTGAAGGTAATGACGCCAAG		59.8		
6	GR979567	F: GTCATTCAGATCACAGTCCCCT	$(CTGCA)_2$	60.3	124	Pleckstrine homology domain
		R: TTTCTCCTTGCTCAGGCTTTAC		60.0		
7	GR979561	F: GCTTGTTTGTGGGATTTCCTTA	$(GGGGA)_2$	60.3	398	
		R CCTTCCTTTGCTTTCTTGGTTA		59.7		
8	GR979573	F: TTTGTCTTGCTACTGCTTCTGG	$(GAACG)_2$	59.6	378	Defensin precursor (Vignaunguiculata), protease inhibitor (*Glycine max*)
		R: CACACACTCCCTCGAATCATAA		59.9		
9	GR979575	F: CTTCTCATCAAGTCCCCAAAAC	$(TTTGC)_2$	59.9	389	Unknown (*Glycine max*), transporter, putative (*Ricinus communis*)
		R: AAGTCAAGAACAAATGTCGGGT		59.9		
10	GR979577	F: GTAGTTGACTTGCTTACTCGTTAGA	$(AGAGC)_2$	59.4	209	Conserved hypothetical protein (*Ricinus communis*)
		R: GAGAGGGAGAGGGAGAGAAAAG		59.5		
11	GR979577	F: CTTTCCTTTTCTCTCCCTCTCC	$(CTCC)_2$	59.8	236	Conserved hypothetical protein (*Ricinus communis*)
		R: GTGGTTGTGGTGTTTTGTTTGA		60.7		
12	GR979578	F: TACAAGATAGCGAAAGCCTAAGTG	$(TC)_5$	59.5	351	Unknown (*Glycine max*)
		R: AGAGTTAAGCAGTGACCAAAAGAC		59.9		
13	GR979578	F: ATAGCGAAAGCCTAAGTGAGAAAT	$(AAAAC)_2$	60.6	347	Unknown (*Glycine max*)
		R: GGGAGAGTTAAGCAGTGACCAA		60.6		
14	GR979585	F: TGGGTACACGAAAGTGACGA	$(AGAAC)_2$	60.2	270	
		R AGACTCGCGACTATGGCAG		60.0		
15	GR979588	F: GAGCAAAATCACACTATCATCTACG	$(TTCAT)_2$	59.3	334	NADPH:quinone oxidoreductase, putative (*Ricinus communis*)
		R: GATTCCATTCCAGTTGTCATACAC		59.6		

Table 2 (*contd.*)

	GenBank accession no.	Primer sequences (5-3')	SSR motif	T_m (°C)	Expected size (bp)	Putative function
16	GR979589	F: CCTTCATTTATTCTTCCTCAGCATC	$(TTGGT)_2$	61.5	252	NADPH:quinoneoxidoreductase, putative (*Ricinus communis*)
		R: CATACCTGTCATTTTGGCATACAG		60.6		
17	GR979592	F: CGGTTTCGCTCTTTCAATTACT	$(TCCTT)_2$	59.8	235	Glutathione peroxidase 1 (*Lotus japonicus*)(stress)
		R: AGATTTAAGATAAGCGTAGAGTGG		59.5		
18	GR979595	F: GTTCTAACCATAGGAGAGACAGTGC	$(GCA)_4$	59.7	390	Unknown (*Glycine max*)
		R: GTTCAAAGTGGTGGTGGATCTAT		59.3		
19	GR979596	F: TCAAGAAAATTAGGCCACCTTC	$(TTTTG)_2$	59.6	279	60S ribosomal protein L18A (RPL18aB) (*Arabidopsis thaliana*)
		R: GCAGTGGTATCAACGCAGAGTA		60.3		
20	GR979603	F: ATTGCGAGTGCGAGAGGTT	$(AAGA)_3$	60.9	111	No homology
		R: ACACGAGTTTCTCTTTGTCTTTGTC		60.2		
21	GR979613	F: AGCACTCTTCATTCTTCTGTTGCT	$(TTGCT)_2/(GAAAA)_2$	60.9	364	60S acidic ribosomal protein P1, putative (*Ricinus communis*)
		R: ATATCATCGTCACTCTCCTCTTCAG		60.4		
22	GR979616	F: CAGTCGGACACTCACTGCTG	$(TATTG)_2$	60.7	152	40S ribosomal protein S7, putative (*Ricinus communis*)
		R TGTACTCGGTGTTGTTTCGC		59.8		
23	GR979618	F: TACTAAGCAGTGGTATCAACGCAA	$(GTATC)_2$	61	268	*Glycine max* cDNA, clone: GMFL01-21-C10
		R: GAAATGTCACGACTCAGAAATCAC		60.9		
24	GR979620	F: CTCTTTCTTCTTCTGCTTCTCTGTCT	$(GCAAG)_2$	60.3	328	60S ribosomal protein L35 (RPL35A) (*Arabidopsis thaliana*)
		R: GAGGATCAAGGTTTACGAACTGA		59.4		
25	GR979622	F: GGATTTCTCGTCGTCTTCTTCT	$(TCT)_4/(TTCTC)_2$	59	338	Cu-Zn superoxide dismutase (*Arachis hypogaea*) (stress)
		R: ACATTCCCTAGATCACCAGCAT		60		
26	GR979623	F: ATCACATGAAGTTGAAACGTGC	$(GGTCT)_2$	60	156	Cu-Zn superoxide dismutase (*Arachis hypogaea*) (stress)
		R: ATCGAGTGCCGACATATAATCC		60.1		

the average being 0.128. A total of 947 alleles were detected and the average number of alleles per SSR marker was 3.30, with a range of 1–4.

To compare the discriminatory potential of genic markers with those of genomic markers, 20 already published markers of Burns *et al.* (2001) were used for the PCR and

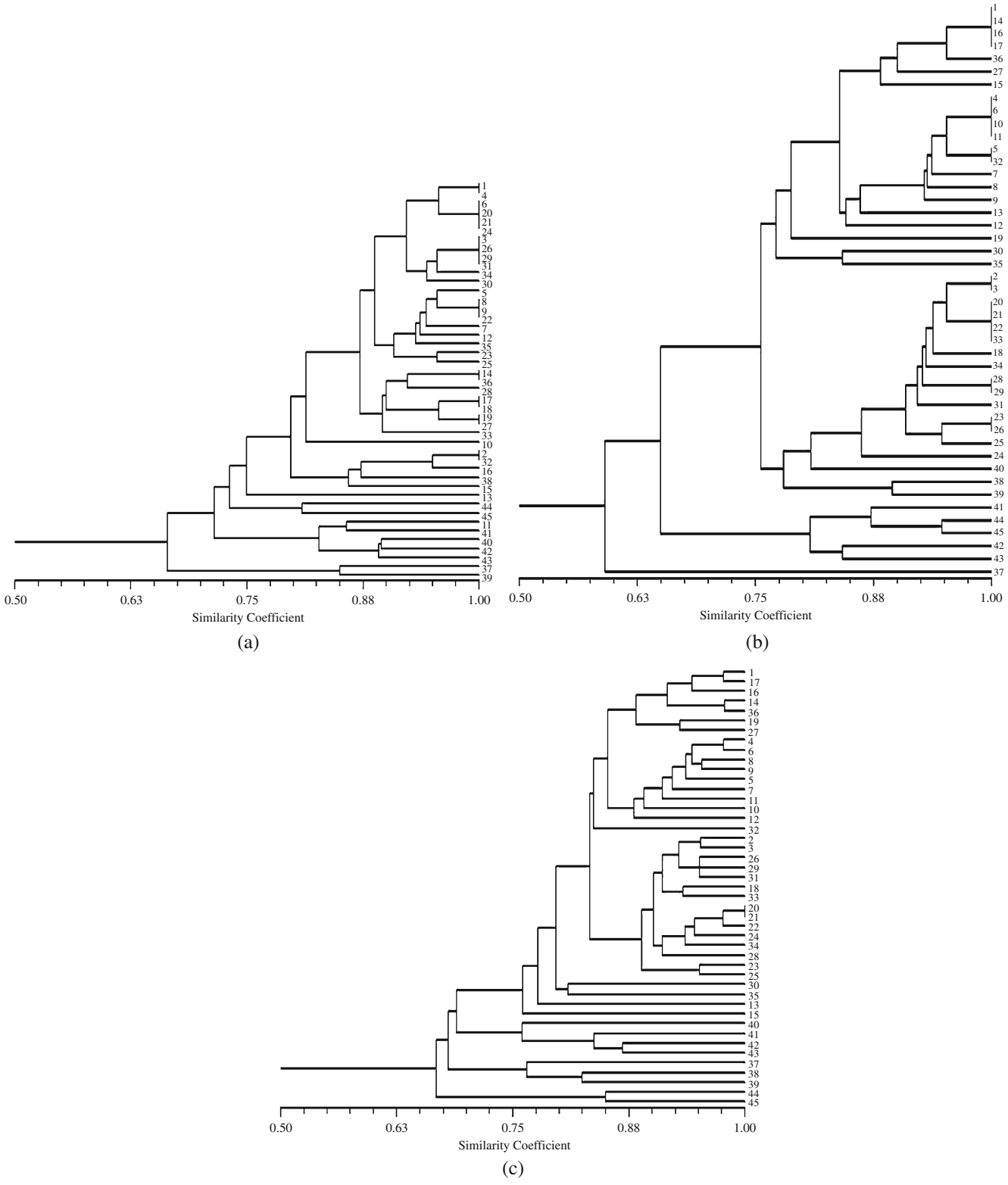


Figure 2. Dendrograms showing clustering pattern of pigeonpea genotypes using (a) 20 genic SSR markers, (b) 20 genomic SSR markers, and (c) all 40 SSR markers used in the study.

diversity analysis among the 45 pigeonpea genotypes used in the study. A total of 40 SSR markers (20 genic and 20 genomic) were used to study the genetic diversity among the 45 genotypes which include different cultivars as well as wild varieties. The genic markers grouped the genotypes into two main clusters, one containing two genotypes of *C. scarabaeoides* (ICP-15685 and ICP-15697) and the other containing the other 43 genotypes (figure 2a). The second cluster subclustered all the wild varieties together with the exception of only two (ICP-1629-2 and ICP-15724) which were clustered with the cultivars.

The dendrogram (figure 2b) obtained using genomic markers on the same set of genotypes showed some variation only in the clustering of the cultivars, whereas the wild varieties were clustered together in one cluster with the exception of only one. A Mantel test (Mantel 1967) performed between the two dendrogram resulted in a *t* value of 5.38 and *r* value of 0.5.

A combined dendrogram (figure 2c) was constructed to study the ability of genic and genomic markers in estimating the diversity between the pigeonpea genotypes. The 40 SSR markers were able to cluster the cultivar and wild varieties separately and all the three *C. scarbaeoides* accessions (ICP-15685, ICP-15724 and ICP-15697) together into one main cluster.

Availability of different databases of ESTs provides an opportunity for development of gene-derived molecular markers which have use in crop improvement programmes. Since analysis of ESTs is a simple strategy to study the transcribed parts of genomes, even large, complex, and highly redundant genomes like those of wheat, barley, maize, tomato, lentil, pea and chickpea are amenable to large-scale analysis. Thus EST-SSR have the potential to expedite evolutionary analyses in a wide variety of taxa and represent the best possible way to analyse species where limited resources are available. These markers will also complement the already available SSR markers. Our results also demonstrate that the designed EST-SSRs show cross-species transferability. Thus the study provides genic SSR markers not only for wild pigeonpea species, but also for genetic studies involving related species that constitute the important gene pool for improvement of pigeonpea.

References

Abdelnoor R. V., Barros E. G. and Moreira M. A. 1995 Determination of genetic diversity within Brazilian soybean germplasm using random amplified DNA amplification techniques and comparative analysis with pedigree data. *Braz. J. Genet.* **18**, 265–273.

Burns M. J., Edwards K. J., Newburg H. J., Ford Loyd B. V. and Baggott C. D. 2001 Development of simple sequence repeat (SSR) markers for the assessment of gene flow and genetic diversity in pigeonpea (*Cajanus cajan*). *Mol. Ecol. Notes* **4**, 283–285.

Chen X., Salamini F. and Gebhardt C. 2001 A potato molecular function map for charbohydrate metabolism and transport. *Theor. Appl. Genet.* **102**, 284–295.

Cordeiro G. M., Casu R., McIntyre C. L., Manners J. M. and Henry R. J. 2001 Microsatellite markers from sugarcane (Saccharum spp.) ESTs cross transferable to erianthus and sorghum. *Plant Sci.* **160**, 1115–1123.

Decroocq V., Fave M. G., Hagen L., Bordenave L. and Decroocq S. 2003 Development and transferability of apricot and grape EST microsatellite markers across taxa. *Theor. Appl. Genet.* **106**, 912–922.

Dutta S., Kumawat G., Singh B. P., Gupta D. K., Singh S., Dogra V. *et al*. 2011 Development of genic-SSR markers by deep transcriptome sequencing in pigeonpea (*Cajanus cajan* (L.) Millspaugh]. *BMC Plant Biol.* **11**, 17.

Holton T. A., Christopher J. T., McClure L., Harker N. and Henry R. J. 2002 Identification and mapping of polymorphic SSR markers from expressed sequences of barley and wheat. *Mol. Breeding* **9**, 63–71.

Jaccard P. 1908 Nouvelle recherches sur La distribution florale. *Bulletin de la Société Vaudoise des Sciences Naturelles* **44**, 223–270.

Mantel N. 1967 Detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**, 209–220.

Odeny D. A., Jayashree B., Ferguson M., Hoisington D., Crouch J. and Gebhardt C. 2007 Development, characterization and utilization of microsatellite markers in pigeonpea. *Plant Breeding* **126**, 130–136.

Odeny D. A., Jayashree B., Gebhardt C. and Crouch J. 2009 New microsatellite markers for pigeonpea (*Cajanus cajan* (L) Millsp.). *BMC Res. Notes*.

Rohlf F. 1998 NTSYSpc: Numerical Taxonomy System, Version 2.2 (2). Exeter Publishing Ltd: Setauket, New York, USA.

Scott K. D., Eggler P., Seaton G., Rossetto M., Ablett E. M., Lee L. S. and Henry R. J. 2000 Analysis of SSRs derived from grape ESTs. *Theor. Appl. Genet.* **100**, 723–726.

Singh N. K., Gupta D. K., Jayaswal P. K., Mahato A. K., Dutta S., Singh S. *et al*. 2012 The first draft of the pigeonpea genome sequence. *J. Pl. Biochem. Biotech.* **21**, 98–112.

Thiel T., Michalek W., Varshney R. K. and Graver A. 2003 Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **106**, 411–422.

Varshney R. K., Graner A. and Sorrells E. 2005 Genic microsatellite markers in plants: features and applications. *Trends Biotechnol.* **23**, 48–55.

Varshney R. K., Chen W., Li Y., Bharti A. K., Saxena R. K., Schlueter J. A. *et al.* 2012 Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nat. Biotech.* **30**, 83–89.

Received 4 October 2012; accepted 22 January 2013
Published on the Web: 12 April 2013