*Mitochondrial SSRs and their utility in distinguishing wild species, CMS lines and maintainer lines in pigeonpea (*Cajanus cajan *L.)*

Pawan Khera, Rachit Saxena, C. V. Sameerkumar, Kulbhushan Saxena & Rajeev K. Varshney

Euphytica
International Journal of Plant Breeding

ISSN 0014-2336

Euphytica
DOI 10.1007/s10681-015-1504-2

Springer

Euphytica
DOI 10.1007/s10681-015-1504-2

Mitochondrial SSRs and their utility in distinguishing wild species, CMS lines and maintainer lines in pigeonpea (*Cajanus cajan* L.)

Pawan Khera · Rachit Saxena · C. V. Sameerkumar · Kulbhushan Saxena · Rajeev K. Varshney

Received: 23 March 2015 / Accepted: 23 June 2015

Abstract Analysis of the pigeonpea mitochondrial genome sequence identified 25 SSRs. Mononucleotide SSR motifs were the most abundant repeats followed by dinucleotide and trinucleotide repeats. Primer pairs could be designed for 24 SSRs, 23 of which were polymorphic amongst the 22 genotypes consisting of cytoplasmic male sterile (CMS or A) line, maintainer or B line and wild *Cajanus* species representing six different CMS systems viz., A_1, A_2, A_4, A_5, A_6 and A_8. These markers amplified a total of 107 alleles ranging from 2 to 10 with an average of 4.65 alleles per locus. The polymorphic information content for these markers ranged from 0.09 to 0.84 with an average of 0.52 per marker. Hence, the present study adds a novel set of 24 mitochondrial SSR markers to the markers repository in pigeonpea, which would be useful to distinguish the genotypes based on mitochondrial genome types in evolutionary and phylogenetic studies.

Keywords CMS · Genetic diversity · Hybrid · Mitochondria · Molecular marker

Electronic supplementary material The online version of this article (doi:10.1007/s10681-015-1504-2) contains supplementary material, which is available to authorized users.

P. Khera · R. Saxena · C. V. Sameerkumar · K. Saxena · R. K. Varshney (✉)
International Crops Research Institute for the Semi-arid Tropics (ICRISAT), Patancheru 502324, Telangana, India
e-mail: r.k.varshney@cgiar.org

Introduction

Pigeonpea [*Cajanus cajan* (L.) Millspaugh] is one of the most suitable crops under subsistence farming system prevalent in the sub-tropical and semi-arid tropic regions of the world. Globally, pigeonpea is cultivated on 5.32 Mha with total production of 4.32 Mt in Asia, Latin America and southern and eastern Africa (FAO 2013). It is an often cross-pollinated (20–70 %) diploid species ($2n = 2X = 22$) with a genome size of 833.07 Mbp (Varshney et al. 2012). During the last six decades, continuous efforts have been made to increase pigeonpea yield, however only limited success was possible. In this scenario, hybrid technology has been found to be a promising option for increasing the yield. In general hybrids were developed by crossing two inbred lines, where manual emasculation is required. Since manual emasculation is time-consuming and cost-ineffective for developing hybrids, cytoplasmic male sterile (CMS) systems were developed. The CMS system involves a CMS line or A line and a maintainer line or B line. Furthermore, to realize hybrid vigour a restorer line or R line is used to cross with A line, so that only fertile F_1 plants are obtained. This complete hybrid system comprising of A line, B line and R line has been termed as cytoplasmic genic male sterility (CGMS) system.

So far, a total of eight CMS systems (A_1–A_8) have been developed in pigeonpea (Saxena 2013; Saxena et al. 2010a, b). These CMS systems are derived from wild *Cajanus* species, namely, A_1 CMS system from

Published online: 07 July 2015

C. sericeus (Benth. ex Bak.) van der Maesen comb. nov. (Saxena et al. 1997), A_2 CMS system from *C. scarabaeoides* (L.) Thou. (Saxena and Kumar 2003), A_3 CMS system from *C. volubilis* (Blanco) Blanco (Wanjari et al. 1999), A_4 CMS system from *C. cajanifolius* (Haines) van der Maesen comb. nov. (Saxena et al. 2005), A_5 CMS system from *C. acutifolius* (F.V. Muell.) van der Maesen comb. nov. (Mallikarjuna and Saxena 2005), A_6 CMS system from *C. lineatus* (Wight & Arn.) van der Maesen (Saxena et al. 2010a, b), A_7 CMS system from *C. platycarpus* (Benth.) van der Maesen (Mallikarjuna et al. 2006) and most recently A_8 CMS system from *C. reticulatus* (Dryander) F. V. Muell. (Saxena 2013). The CMS systems derived from A_2 and A_4 cytoplasms have been successfully used in developing commercial hybrids. It is noteworthy that the first A_4 cytoplasm based pigeonpea hybrid ICPH 2671, showing an yield advantage of up to 45 % over control (Saxena et al. 2010a, b), was released for commercial cultivation in Central India (Saxena et al. 2013) and many promising hybrids such as ICPH 3762, ICPH 2740, etc. are in pipeline for release (personal communication). However, a number of constraints such as unstable male sterility, poor fertility restoration, lack of proper and complete restorers, etc. have been observed while utilizing other CMS systems in hybrid development.

It is now widely known that the CMS in plants arises due to aberrations in the mitochondrial (mt) genome (Tuteja et al. 2013). Therefore, understanding the genetic relationships of nuclear and mitochondrial interactions at genome level in CGMS system is critical to hybrid pigeonpea breeding. Recently the mitochondrial genome of 545.7 kb has been assembled for a pigeonpea A line ICPA 2039 derived from A_4 cytoplasm (Tuteja et al. 2013). However, little is known about the nature and organization of simple sequence repeats (SSRs) in the pigeonpea mitochondrial genome. Microsatellite or SSR markers are known to be abundant, hypervariable and ubiquitous across prokaryotic and eukaryotic genomes. The high level of polymorphism combined with their high reproducibility and co-dominance nature have established them as the markers of choice for genetic studies and breeding applications in crops (Gupta and Varshney 2000). In the past, SSRs were developed by using a range of methods such as SSR-enriched libraries, BAC-end sequencing and mining of sequence data (Varshney et al. 2005a, b). However in silico mining of sequence data from nuclear or mitochondrial genome is the most cost-effective and fast method (Rajendrakumar et al. 2007). Therefore SSRs have been developed from the mitochondrial genomes in a number of crops such as sorghum (Nishikawa et al. 2002), rice (Nishikawa et al. 2005; Rajendrakumar et al. 2007) and cotton (Zhang et al. 2012).

In view of the importance of CMS for hybrid breeding, and the possible use of mt-genome-derived SSR markers, the present study reports identification, categorization and development of mt-genome derived SSR markers in pigeonpea. The mt-SSR markers so developed were used to understand the genetic relationships among A lines, B lines and wild *Cajanus* species accessions from six different CMS systems.

Materials and methods

Plant material and DNA isolation

A total of 22 genotypes including 8 A lines, 8 B lines and 6 wild species accessions were used in the present study. These 22 genotypes represented the following six different CMS systems in pigeonpea, viz., A_1 (*C. sericeus*), A_2 (*C. scarabaeoides*), A_4 (*C. cajanifolius*), A_5 (*C. acutifolius*), A_6 (*C. lineatus*) and A_8 (*C. reticulatus*) (Table 1).

Genomic DNA was isolated from freshly harvested young leaves from 2-week-old seedlings using the standard DNA isolation protocol as mentioned in Cuc et al. (2008). The quantity and quality of DNA samples was assessed on 0.8 % agarose gel and the DNA was then diluted to 5 ng μL^{-1} for genotyping.

SSR identification and primer designing

A high-quality DNA sequence of pigeonpea mitochondrial genome ICPA 2039 (SRA053693) from GenBank (http://www.ncbi.nlm.nih.gov; Tuteja et al. 2013) was used for mining SSRs using MISA software (http://pgrc.ipk-gatersleben.de/misa/; Thiel et al. 2003). The search criteria used were mono (N) $\geq$ 10, di (NN) $\geq$ 6, tri (NNN) $\geq$ 5, tetra (NNNN) $\geq$ 5, penta (NNNNN) $\geq$ 5 and hexanucleotides (NNNNNN) $\geq$ 5 repeats as well as compound SSRs (which were interrupted by few bases).

Table 1 A list of 22 genotypes including 8 A lines, 8 B lines and 6 wild species accessions representing 6 CMS systems used in the present study

CMS system	Genotype	Characteristic
A_1	ICPW 162	Progenitor wild species *Cajanus sericeus*
	ICPA 2067	CMS line or A line
	ICPA 2068	CMS line or A line
	ICPA 2032	CMS line or A line
	ICPB 2067	Maintainer line or B line
	ICPB 2068	Maintainer line or B line
	ICPB 2032	Maintainer line or B line
A_2	ICPW 89	Progenitor wild species *C. scarabaeoides*
	ICPA 2052	CMS line or A line
	ICPB 2052	Maintainer line or B line
A_4	ICPW 29	Progenitor wild species *C. cajanifolius*
	ICPA 2039	CMS line or A line
	ICPB 2039	Maintainer line or B line
A_5	ICPW 2	Progenitor wild species *C. acutifolius*
	A Line*	CMS line or A line
	B Line*	Maintainer line or B line
A_6	ICPW 42	Progenitor wild species *C. lineatus*
	ICPA 2209	CMS line or A line
	ICPB 2209	Maintainer line or B line
A_8	ICPW 74	Progenitor wild species *C. reticulatus*
	ICPA 2212	CMS line or A line
	ICPB 2212	Maintainer line or B line

CMS cytoplasmic male sterility

* Accessions don't have designated code

Identified sequences containing SSRs were used for designing primer pairs using Primer3 software (Untergrasser et al. 2012). Following criteria were used for designing primers pairs:, amplicon length in the range of 100–300 bp, optimal melting temperature set to 60 °C and optimal primer size of 20 bp. Rest of the options were default values of Primer3 software.

Polymerase chain reaction

Polymerase chain reaction (PCR) mix contains 10 µl reaction volume constituting 1.0 µl of 10× PCR buffer, 1.0 µl of 2 mM dNTPs, 1.0 µl of 2 pM primer (Eurofins, Bangalore, India), 0.4 µl of 25 mM $MgCl_2$, 0.06 U of Taq polymerase (Kappa Biosystems, Woburn, MA, USA), 1.5 µl (5 ng) of template DNA, 1.0 µl of fluorescent dyes (FAM, NED, VIC or PET) and 4.04 H_2O in 96-well microtitre plate (Axygen Inc., Union City, CA, USA). The DNA fragments were then amplified using a touch-down PCR programme consisting of an initial denaturation step at 95 °C for 3 min, followed by 5 cycles, each cycle involving denaturation for 25 s at 94 °C, annealing for 20 s at 60 °C (the annealing temperature for each cycle being reduced by 1 °C per cycle) and extension for 30 s at 72 °C. The touch-down PCR was followed by 40 cycles, each cycle with denaturation for 20 s at 94 °C, annealing for 20 s at 55 °C and extension for 30 s at 72 °C, and subsequently the final extension step at 72 °C for 20 min using Gene-Amp PCR System 9700 thermal cycler (Applied Biosystems, Foster City, California, USA).

SSR fragment analysis

SSR fragment analysis was performed through capillary electrophoresis using 1 µl amplified PCR product, 7.5 µl of Hi-Di formamide, 0.05 µl of internal lane standard GeneScan 500 (Applied Biosystems) labelled with LIZ after denaturation at 95 °C for 5 min. LIZ 500 internal lane standard and the GeneScan Filter Set D were used for size fractionation of amplicons labelled with different fluorescent dyes. Capillary electrophoresis was done using ABI 3700 Genetic Analyzer

(Applied Biosystems, Foster City, CA, USA). Allele sizing and scoring of the electrophoresis data was carried out using the GeneMapper 4.0 software (Applied Biosystems, Foster City, CA, USA).

Analysis of genotyping data

Allelic data recorded for each marker was subjected to Allelobin software (Prasanth et al. 2006) in order to get allele calls based on the repeat motif of each SSR. Several features of the mtSSR markers such as polymorphism information content (PIC) value, major allele frequency, and allele numbers were calculated using PowerMarker version 3.25 software (Liu and Muse 2005). DARwin version 5.0.158 (Perrier and Jacquemoud-Collet 2006) was used for principal coordinate analysis (PCoA) and construction of neighbour-joining tree.

Results

Distribution of SSRs

The pigeonpea mitochondrial genome ICPA 2039 (Tuteja et al. 2013) was used for mining SSRs using *MI*cro*SA*tellite (MISA) software (Thiel et al. 2003). A total of 25 SSRs were identified with a frequency of 0.046 SSR per kb of the mitochondrial genome. Identified SSRs were designated as ***C****ajanus* ***c****ajan* **Mit**ochondrial (CcMt) SSRs. Among the CcMt SSR classes, mononucleotides constituted the major proportion at 60 % of the total SSRs identified (15 out of 25 SSRs identified). Only two mononucleotide SSR motifs namely poly "A" and poly "T" were present, however the SSR motif poly "A" was more abundant with its presence in 9 mononucleotide SSRs, the poly "T" being present in only 6 SSRs. Dinucleotide SSR motifs were the second most abundant repeats constituting 36 % of the total SSRs identified (9 out of 25 SSRs identified), with TA/AT repeat motif in six SSRs followed by TC/CT in two SSRs and AG in a solitary SSR. Only one trinucleotide SSR with "TAA" motif was identified.

Development of SSR markers

SSR containing sequences were used for primer pair designing. As a result, the primer pairs were designed for 24 SSRs. Amplification conditions for all 24 primer pairs were optimized initially on two pigeonpea genotypes viz., ICPA 2039 (A line) and ICPB 2039 (B line). Both the genotypes showed 100 % amplification for all the primer pairs. Subsequently, primer pairs for all 24 SSRs were used for assessing the polymorphism in pigeonpea germplasm (Table 2).

Polymorphism assessment in different CMS sources

A total of 22 genotypes consisting of A lines, B lines and wild *Cajanus* species representing the following six different CMS systems (Table 1) were used for assessment of polymorphism: A_1 (*Cajanus sericeus*), A_2 (*C. scarabaeoides*), A_4 (*C. cajanifolius*), A_5 (*C. acutifolius*), A_6 (*C. lineatus*) and A_8 (*C. reticulatus*). Out of 24 CcMt SSR markers screened, a solitary marker (CcMt13) was found to be monomorphic. Remaining 23 SSRs showed polymorphism and generated a total of 107 alleles with an average of 4.65 alleles per SSR marker. Allele numbers identified by the polymorphic SSR markers ranged from 2 for 5 markers (CcMt07, CcMt12, CcMt18, CcMt22 and CcMt24) to 10 for one marker (CcMt19). The PIC values for the polymorphic markers varied from 0.09 (CcMt12) to 0.84 (CcMt03), with an average of 0.52 per marker. Furthermore, the major allele frequency for these SSRs ranged from 0.21 (CcMt03) to 0.95 (CcMt12) with a mean of 0.54 (Table 3).

Genetic relationships among A line, B line and wild species accessions

The allelic data generated on 23 polymorphic CcMt SSR markers on 22 genotypes could separate all individual genotypes and were used to calculate genetic dissimilarity matrix and to construct dendrogram (Fig. 1) and factorial analysis (Supplementary Fig. 1) using the DARwin software. Five clusters (Cl) were identified based on the dendrogram (Fig. 1). The Cl I contained 8 genotypes, Cl II contained 4 genotypes, Cl III contained 2 genotypes, Cl IV contained 4 genotypes while Cl V contained 3 genotypes. However, one genotype ICPB 2052 was not part of any cluster and an outgroup in the dendrogram. The Cl I represents by and large genotypes related to A_1 CMS system (3 A lines, 3 B lines and wild representative ICPW 162 which is donor of

Table 2 Details of mitochondrial SSR markers in pigeonpea

Sl. No.	Marker name	SSR motif	Start position[a]	End position[a]	Forward primer (5′–3′)	Tm (°C)	Reverse primer (5′–3′)	Tm (°C)	Product size (bp)
1	CcMt01	$(A)_{10}$	1915	1924	AGGAGACAGCTGATGGTGCT	60.0	CCTTGGGGGATACACCTTCT	60.2	132
2	CcMt02	$(TC)_6$	4702	4713	CCCAGCTAAAGCGAATCTCA	60.5	ACTGGTGCTTACACCGAAGC	60.3	106
3	CcMt03	$(T)_{10}$	50,593	50,602	TATCTTGCTTGAGAGGGGGA	59.8	GAACTCAAAGAGCGAGCCAC	60.1	215
4	CcMt04	$(AT)_8$	53,553	53,568	TAGCGAAGGAGAGGGTGAGA	60.1	GCAATGATTTCAGACGCAGA	60.0	257
5	CcMt05	$(AT)_6$	85,067	85,078	GGAGGTTGAGGTCTGTGTCC	59.5	TGAGGCTAGGCCTTATTGGA	59.9	181
6	CcMt06	$(A)_{11}$	102,804	102,814	AGGCCCAACAACCTAAAAGG	60.3	AAGGTACGGAGTCGCTCAAC	59.5	228
7	CcMt07	$(A)_{10}$	156,298	156,307	CTCGAGTTTCTTGGTCGGTC	59.8	TTTCTCAGGGGATGAGGTTG	60.0	153
8	CcMt08	$(A)_{11}$	167,208	167,218	TAAAGCAGTGTGCGTCATCC	59.9	CGCCCACCTCATTCATATCT	59.9	115
9	CcMt09	$(A)_{10}$	191,482	191,491	GAGCATAAGGATTAGCAGAATCG	59.4	CAAGTGCTACGAGTGCTTCG	59.8	104
10	CcMt10	$(A)_{10}$	194,635	194,644	TGGGTCCAAAGGACAGGTAG	60.0	AGAAAATCTGCCTCCAGCAA	60.0	139
11	CcMt11	$(CT)_6$	202,827	202,838	CCAACCTCTCGAAATCCAAA	60.0	TCAGATGATTTGTGGACGGA	60.1	107
12	CcMt12	$(TA)_6$	227,980	227,991	ACAGACCAAGCAAGGGCTTA	59.9	AGGGGCTTAAAGGAGTTTCG	59.7	147
13	CcMt13	$(AT)_6$	242,115	242,126	CTGAGAGAAGGCCTGTGACC	60.0	ATTATTTCCACCCCCTCGTC	60.0	273
14	CcMt14	$(T)_{10}$	248,345	248,354	ACCCACAATCAGCCAAGTTC	60.0	CGAGGTCTCAACGAAAGGAG	60.0	185
15	CcMt15	$(T)_{10}$	264,267	264,276	ACTCCGTGCAGGAAAGACTC	59.5	GAGCGGCCTGTAATAAGTGC	60.0	238
16	CcMt16	$(T)_{10}$	278,202	278,211	CGGGGTTTGATAGTTGCAGT	60.0	CAATCCCTCTTTCTGAAGCG	60.0	111
17	CcMt17	$(A)_{11}$	333,351	333,361	AAAAGCGTGTTCCCTTCAGA	59.9	TAAAGGAAAGGCTCGACGAA	60.0	142
18	CcMt18	$(AG)_6$	348,442	348,453	TTCCACTTGCTTTCTCGCTT	60.1	AGTTGGTGAAACCTGAGCGT	60.0	187
19	CcMt19	$(TA)_6$	368,690	368,701	CCTCTTCCCAGGAATGAACA	60.0	AAGCAAATAGATAGCCCCCG	60.4	168
20	CcMt20	$(T)_{11}$	440,248	440,258	TGCTCTCTTGCTACGGGATT	60.0	ACGCTGGTCCATCACTTACC	60.0	203
21	CcMt21	$(A)_{11}$	505,653	505,663	CCTCGTGGGGAAAGATAAGA	59.1	TATCTCCCCCTTGCCTTTTT	60.0	234
22	CcMt22	$(TA)_7$	529,454	529,467	CAGAACGAATCTCTTATCGCC	58.9	ACTCCAGTCCTACCCCGTCT	60.0	175
23	CcMt23	$(A)_{10}$	530,826	530,835	GTTTCGGTACGAAAGCCTGA	60.2	TCCGTCTTGCCTTTGCTACT	60.0	202
24	CcMt24	$(TAA)_5$	544,996	545,010	AGTCCCGGGATCGTAAAGTT	59.8	TGTCACAAGGAAGGTGGTCA	60.1	146

SSR simple sequence repeats, *Tm* melting temperature

[a] The start position and end position denotes the starting number of bp for the SSR in the 545,742 bp assembly of ICPA 2039 pigeonpea mitochondrial genome (Tuteja et al. 2013)

Table 3 Polymorphism features of pigeonpea mitochondrial SSR markers

Sl. No.	Marker name	Allele No.	Major allele frequency	PIC value
1	CcMt01	3	0.57	0.52
2	CcMt02	4	0.55	0.47
3	CcMt03	8	0.21	0.84
4	CcMt04	9	0.29	0.80
5	CcMt05	5	0.40	0.63
6	CcMt06	4	0.45	0.55
7	CcMt07	2	0.67	0.35
8	CcMt08	3	0.68	0.43
9	CcMt09	4	0.71	0.37
10	CcMt10	6	0.67	0.49
11	CcMt11	6	0.36	0.69
12	CcMt12	2	0.95	0.09
13	CcMt14	3	0.60	0.47
14	CcMt15	5	0.40	0.68
15	CcMt16	5	0.57	0.56
16	CcMt17	4	0.66	0.49
17	CcMt18	2	0.52	0.37
18	CcMt19	10	0.29	0.79
19	CcMt20	7	0.35	0.78
20	CcMt21	8	0.34	0.78
21	CcMt22	2	0.50	0.38
22	CcMt23	3	0.90	0.17
23	CcMt24	2	0.87	0.20
	Mean	4.65	0.54	0.52

PIC value polymorphism information content value

A_1 cytoplasm), except for ICPA 2052 (A_2 cytoplasm). The Cl II, had mixture of genotypes i.e. ICPA 2039 (A_4 cytoplasm) and ICPW 29 (wild species donor of A_4 cytoplasm), ICPW 89 (wild species donor of A_2 cytoplasm) and ICPB 2209 (B line of A_6 cytoplasm). The Cl III had two genotypes, one A line and the other wild species accession of pigeonpea having A_8 cytoplasm. The Cl IV had 3 genotypes from different cytoplasms such as A_5 (A line) and A_6 (ICPW 42 and ICPA 2209) cytoplasm as well as a B line of A_8 cytoplasm. Cl V had 3 genotypes, a wild genotype ICPW 2 (wild species of A_5 cytoplasm), a B line of A_5 cytoplasm and ICPB 2039 (B line for A_4 cytoplasm).

Discussion

Simple sequence repeats are hypervariable and available in abundance in a range of crop genomes (Gupta and Varshney 2000). The polymorphism in SSRs is believed to be the outcome of replication slippage (Moxon and Wills 1999), which commonly occurs at a higher rate than mutation in the non-repetitive DNA (Wierdl et al. 1997). Owing to their high polymorphism, co-dominance and reproducible nature, they have consequently become the marker of choice for genetic analyses such as molecular mapping, diversity studies and breeding applications in crops (Gupta and Varshney 2000). In recent years, the availability of genome sequence information has eliminated technical limitations and enabled researchers to accelerate the process of SSR development (Varshney et al. 2005b).

In case of pigeonpea, until 2010 efforts towards identification of SSRs were limited to identification of SSR markers from SSR-enriched libraries (Burns et al. 2001; Odeny et al. 2007; Saxena et al. 2010a, b). However, with the advent of sequencing technologies,

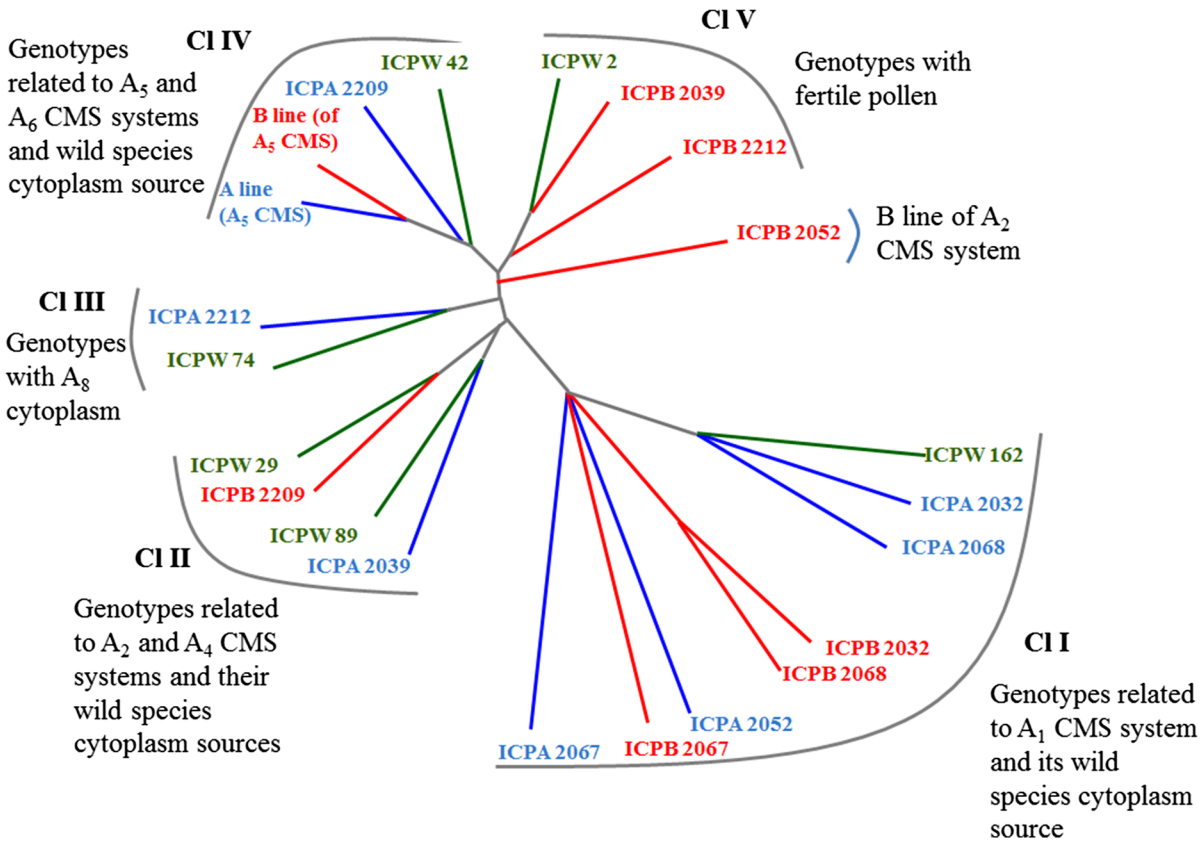

Fig. 1 Clustering pattern of pigeonpea genotypes (8 A lines, 8 B lines and 6 wild species accessions) obtained from survey of 23 mitochondrial SSR markers. Genotypes with *blue*, *red* and *green* colours indicate A lines, B lines and wild species accessions, respectively

the efforts towards identification of SSRs in pigeonpea increased leading to initial identification of 3583 genic SSRs through transcriptome sequencing (Raju et al. 2010), 3072 SSRs through BAC end sequencing (Bohra et al. 2011) and later to a collection of 309,052 SSRs through draft genome sequencing (Varshney et al. 2012); however, the present report is the first report of mt-SSRs in pigeonpea.

The present study reports a set of 25 novel mitochondrial SSRs identified through in silico analysis of mitochondrial genome. Mononucleotide SSR motif emerged as a major class of SSRs, followed by dinucleotide and trinucleotide. While comparing organellar genomes of major cereals including rice, wheat, maize and sorghum, mononucleotides were the most frequent repeats in an earlier study by Rajendrakumar et al. (2008), the commonest mononucleotide repeat being poly "A" and poly "T". These mononucleotide repeats were also relatively more abundant in mitochondrial genomes of algae and angiosperms (Kuntal and Sharma 2011). The frequency of SSRs in pigeonpea mitochondrial genome observed during the present study was relatively low, when compared with that in mitochondrial genomes of cereals (Rajendrakumar et al. 2007, 2008). In the case of legumes other than pigeonpea, the mitochondrial genomes of mung bean (*Vigna radiata*) (Alverson et al. 2011), faba bean (Negruk 2013) and soybean (Chang et al. 2013) have been sequenced. However, SSRs in the mitochondrial genomes of these above other legumes have yet to be discovered. The low frequency of SSRs in legumes may be attributed to a low proportion of repetitive DNA in legumes relative to that in cereals and eudicots (Alverson et al. 2011; Tuteja et al. 2013).

All mtSSRs were considered for designing the primer pairs. The percentage of primers designed was high at 96 % (24 out of 25 were successfully designed). While conducting polymorphism survey using these novel 24 markers on 22 genotypes (which

represents six different cytoplasm in *Cajanus*), 96 % (23 out of 24 markers) of the markers showed polymorphism. This polymorphism rate is comparable to that with the SSRs from the nuclear genome, wherein 95 % polymorphism was observed (Odeny et al. 2007). In other studies, involving nuclear SSRs, polymorphism was low, with reports of 81.3 %, (Saxena et al. 2010a, b) and 50 % (Burns et al. 2001). The high level of polymorphism rate observed may be attributed to the use of wild species from secondary gene pool of pigeonpea. The average PIC value and number of alleles were 0.52 and 4.65, respectively, which were also comparable to those reported by Odeny et al. (2007), at 0.60 and 4.8 for the respective values. In another study, the average number of alleles and PIC value were respectively 3.4 and 0.32, which are relatively low (Saxena et al. 2010a, b).

In the present study, the average PIC value of mononucleotide and dinucleotide SSRs were almost same. However, higher level of polymorphism for dinucleotides was reported in the past (Ashworth et al. 2004; Saxena et al. 2010a, b). Furthermore, due to longer repeat length, better level of polymorphism has been observed by dinucleotide repeats, and they are in general considered better SSR markers than mononucleotide repeats (Temnykh et al. 2001). Since it is difficult to resolve polymorphism in mononucleotide SSRs on agarose gel, capillary electrophoresis was used to resolve the polymorphism for the SSR markers identified during the present study. It was interesting to note that all the 14 mononucleotide SSR markers showed polymorphism. In fact, CcMt03 (mononucleotide SSR) showed highest PIC of 0.84. Similar results were obtained by (Saxena et al. 2010a, b), where two mononucleotide SSR markers showed relatively high level of polymorphism. These observations suggest that in pigeonpea, it may be desirable to develop and deploy mononucleotide SSRs also.

Genotyping data were used for elucidating genetic dissimilarity among the 22 genotypes. The hypothesis used for deciphering the genetic relationship was that, the CMS occurs due to mitochondrial aberration which is maternally inherited, and since all the CMS systems in pigeonpea are from the wild species, the mitochondrial variations between the CMS line and its wild progenitor would be minimal. Chimeric open reading frames (ORFs) caused due to mitochondrial genome rearrangement are the leading cause of CMS trait in plants (Hanson and Bentolila 2004). In the case of pigeonpea, a total of 13 potential chimeric ORFs were identified in the male-sterile line ICPA 2039. Out of these 13 candidate ORFs, eight were within other mitochondrial genes, while five were having parts of different mitochondrial genes (Tuteja et al. 2013).

Phylogenetic analysis grouped 22 genotypes into 5 clusters. Clear clusters were observed for genotypes related to the CMS system of A_1 and A_8 cytoplasm. This suggests that these markers could by and large distinguish differences in the CMS system rather than at the genotype level. Genotypes related to A_2 and A_4 CMS system grouped together into a single cluster and therefore suggest that A_2 and A_4 CMS system have some homology. As earlier mentioned, in pigeonpea only these two CMS system have been successfully utilized to develop pigeonpea hybrids, with ICPH 2671 based on A_4 cytoplasm (Saxena et al. 2013) and IPH 09-5 based on A_2 cytoplasm (Pulses Newsletter IIPR 2013). Genotypes with A_5 and A_6 cytoplasm were grouped together into another cluster, suggesting that these two CMS system might have common ancestry.

In summary, the present study adds a novel set of 24 mitochondrial SSR markers to the marker repository in pigeonpea. These markers would be useful not only in differentiating A lines and their corresponding B line, but also for the study of the origin and phylogeny of pigeonpea.

Acknowledgments Authors thank United States Agency for International Development (USAID) and Biotechnology Industry Partnership Programme (BIPP), Department of Biotechnology, Government of India for financial support. Authors also thank M. Sriswathi and Abdul Gafoor for providing technical support. This work has been undertaken as part of the CGIAR Research Program on Grain Legumes. ICRISAT is a member of CGIAR Consortium.

References

Alverson AJ, Zhuo S, Rice DW, Sloan DB, Palmer JD (2011) The mitochondrial genome of the legume *Vigna radiata*

and the analysis of recombination across short mitochondrial repeats. PLoS One 6:e16404

Ashworth VETM, Kobayashi MC, De La Cruz M, Clegg MT (2004) Microsatellite markers in avocado (*Persea Americana* Mill.): development of dinucleotide and trinucleotide markers. Sci Hortic 101:255–267

Bohra A, Dubey A, Saxena RK, Varma Penmetsa R, Poornima KN, Kumar N, Farmer AD, Srivani G, Upadhyaya HD, Gothalwal R, Ramesh S, Singh D, Saxena KB, Kavi Kishor PB, Singh NK, Town CD, May GD, Cook DR, Varshney RK (2011) Analysis of BAC-end sequences (BESs) and development of BES-SSR markers for genetic mapping and hybrid purity assessment in pigeonpea (*Cajanus* spp). BMC Plant Biol 11:56

Burns MJ, Edwards KJ, Newbury HJ, Ford-Lloyd BV, Baggott CD (2001) Development of simple sequence repeat (SSR) markers for the assessment of gene flow and genetic diversity in pigeonpea (*Cajanus cajan*). Mol Ecol Notes 1:283–285

Chang S, Wang Y, Lu J, Gai J, Li J, Chu P, Guan R, Zhao T (2013) The mitochondrial genome of soybean reveals complex genome structures and gene evolution at intercellular and phylogenetic levels. PLoS One 8:e56502

Cuc LM, Mace ES, Crouch JH, Quang VD, Long TD, Varshney RK (2008) Isolation and characterization of novel microsatellite markers and their application for diversity assessment in cultivated groundnut (*Arachis hypogaea* L.). BMC Plant Biol 8:55

Food and Agricultural Organization of the United Nation, FAO (2013) FAO statistical database. http://faostat3.fao.org/faostat-gateway/go/to/home/E. Accessed 7 Jan 2013

Gupta PK, Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica 113:163–185

Hanson MR, Bentolila S (2004) Interactions of mitochondrial and nuclear genes that affect male gametophyte development. Plant Cell 16:154–169

Kuntal H, Sharma V (2011) In silico analysis of SSRs in mitochondrial genomes of plants. Omics 15:783–789

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129

Mallikarjuna N, Saxena KB (2005) A new cytoplasmic male sterility system derived from cultivated pigeonpea cytoplasm. Euphytica 142:143–148

Mallikarjuna N, Jadhav D, Reddy P (2006) Introgression of *Cajanus platycarpus* genome into cultivated pigeonpea genome. Euphytica 149:161–167

Moxon ER, Wills C (1999) DNA microsatellites: agents of evolution? Sci Am 280:94–99

Negruk V (2013) Mitochondrial genome sequence of the legume *Vicia faba*. Front Plant Sci 4:128

Nishikawa T, Salomon B, Komatsuda T, von Bothmer R, Kadowaki KI (2002) Molecular phylogeny of the genus *Hordeum* using three chloroplast DNA sequences. Genome 45:1157–1166

Nishikawa T, Vaughan DA, Kadowaki KI (2005) Phylogenetic analysis of *Oryza* species, based on simple sequence repeats and their flanking nucleotide sequences from the mitochondrial and chloroplast genomes. Theor Appl Genet 110:696–705

Odeny DA, Jayashree B, Ferguson M, Hoisington D, Crouch J, Gebhardt C (2007) Development characterization and utilization of microsatellite markers in pigeonpea [*Cajanus cajan* (L.) Millsp.]. Plant Breed 126:130–137

Perrier X, Jacquemoud-Collet JP (2006) DARwin software. Centre de cooperation internationale en recherche agronomique pour le développement (CIRAD), Paris. http://darwin.cirad.fr/. Accessed 3 Jan 2014

Prasanth VP, Chandra S, Jayashree B, Hoisington D (2006) AlleloBin—a program for allele binning of microsatellite markers based on the algorithm of Idury and Cardon (1997). International Crops Research Institute for the Semi-Arid Tropics, Patancheru

Pulses Newsletter IIPR, January–March (2013) Research advisory committee meeting held. http://www.iipr.res.in/pdf/newsletter_2013_5june.pdf. Accessed 12 Feb 2014

Rajendrakumar P, Biswal AK, Balachandran SM, Srinivasarao K, Sundaram RM (2007) Simple sequence repeats in organellar genomes of rice: frequency and distribution in genic and intergenic regions. Bioinformatics 23:1–4

Rajendrakumar P, Biswal AK, Balachandran SM, Sundaram RM (2008) In silico analysis of microsatellites in organellar genomes of major cereals for understanding their phylogenetic relationships. In silico Biol 8:87–104

Raju NL, Gnanesh BN, Lekha P, Jayashree B, Pande S, Hiremath PJ, Byregowda M, Singh NK, Varshney RK (2010) The first set of EST resource for gene discovery and marker development in pigeonpea (*Cajanus cajan* L.). BMC Plant Biol 10:45

Saxena KB (2013) A novel source of CMS in pigeonpea derived from *Cajanus reticulatus*. Indian J Genet Plant Breed 73:259–263

Saxena KB, Kumar RV (2003) Development of a cytoplasmic-nuclear male sterility system in pigeonpea using *C. scarabaeoides* (L.) Thours. Indian J Genet Plant Breed 63:225–229

Saxena KB, Kumar RV, Singh L, Raina R (1997) Development of a cytoplasmic nuclear male-sterility system in pigeonpea. Progress report fourth consultative group meeting on cytoplasmic male-sterility in pigeonpea, 3–4 March 1997. Bhaba Atomic Research Centre, Trombay, p 19

Saxena KB, Kumar RV, Srivastava N, Shiying B (2005) A cytoplasmic-nuclear male sterility system derived from a cross between *Cajanus cajanifolius* and *C. cajan*. Euphytica 145:291–296

Saxena KB, Sultana R, Mallikarjuna N, Saxena RK, Kumar RV, Sawargaonkar SL, Varshney RK (2010a) Male-sterility systems in pigeonpea and their role in enhancing yield. Plant Breed 129:125–134

Saxena RK, Prathima C, Saxena KB, Hoisington DA, Singh NK, Varshney RK (2010b) Novel SSR Markers for polymorphism detection in pigeon pea (*Cajanus* spp.). Plant Breed 129:142–148

Saxena KB, Kumar RV, Tikle AN, Saxena MK, Gautam VS, Rao SK, Khare DK, Chauhan YS, Saxena RK, Reddy BVS, Sharma D, Reddy LJ, Green JM, Faris DG, Nene YL, Mula M, Sultana R, Srivastava RK, Gowda CLL, Sawargaonkar SL, Varshney RK (2013) ICPH 2671—The world's first commercial food legume hybrid. Plant Breed 132:479–485

Temnykh S, DeClerck G, Lukashova A, Lipovich L, Cartinhour S, McCouch S (2001) Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genet Res 11:1441–1452

Thiel T, Michalek W, Varshney R, Graner 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

Tuteja R, Saxena RK, Davila J, Shah T, Chen W, Xiao YL, Fan G, Saxena KB, Alverson AJ, Spillane C, Town C, Varshney RK (2013) Cytoplasmic male sterility-associated chimeric open reading frames identified by mitochondrial genome sequencing of four *Cajanus* genotypes. DNA Res 20:485–495

Untergrasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. Nucleic Acids Res 40:e115

Varshney RK, Graner A, Sorrells ME (2005a) Genomics-assisted breeding for crop improvement. Trends Plant Sci 10:621–630

Varshney RK, Graner A, Sorrells ME (2005b) Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23:48–55

Varshney RK, Chen W, Li Y, Bharti AK, Saxena RK, Schlueter JA, Donoghue MTA, Azam S, Fan G, Whaley AM, Farmer AD, Sheridan J, Iwata A, Tuteja R, Penmetsa RV, Wu W, Upadhyaya HD, Yang SP, Shah T, Saxena KB, Michael T, McCombie WR, Yang B, Zhang G, Yang H, Wang J, Spillane C, Cook DR, May GD, Xu X, Jackson SA (2012) Draft genome sequence of pigeon pea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. Nat Biotechnol 30:83–89

Wanjari KB, Patil AN, Manapure P, Manjaya JG, Manish P (1999) Cytoplasmic male-sterility in pigeonpea with cytoplasm from *Cajanus volubilis*. Ann Plant Physiol 13:170–174

Wierdl M, Dominska M, Petes TD (1997) Microsatellite instability in yeast: dependence on the length of the microsatellite. Genetics 146:769–779

Zhang X, Meng Z, Zhou T, Sun G, Shi J, Yu Y, Zhang R, Guo S (2012) Mitochondrial SCAR and SSR markers for distinguishing cytoplasmic male sterile lines from their isogenic maintainer lines in cotton. Plant Breed 131:563–570