Theor Appl Genet (2006) 112: 1416–1428
DOI 10.1007/s00122-006-0243-0

ORIGINAL PAPER

Niroj Kumar Sethy · Bhumika Shokeen
Keith J. Edwards · Sabhyata Bhatia

Development of microsatellite markers and analysis of intraspecific genetic variability in chickpea (*Cicer arietinum* L.)

Received: 9 September 2005 / Accepted: 13 February 2006 / Published online: 14 March 2006

Abstract Paucity of polymorphic molecular markers in chickpea (*Cicer arietinum* L.) has been a major limitation in the improvement of this important legume. Hence, in an attempt to develop sequence-tagged microsatellite sites (STMS) markers from chickpea, a microsatellite enriched library from the *C. arietinum* cv. Pusa362 nuclear genome was constructed for the identification of $(CA/GT)_n$ and $(CT/GA)_n$ microsatellite motifs. A total of 92 new microsatellites were identified, of which 74 functional STMS primer pairs were developed. These markers were validated using 9 chickpea and one *C. reticulatum* accession. Of the STMS markers developed, 25 polymorphic markers were used to analyze the intraspecific genetic diversity within 36 geographically diverse chickpea accessions. The 25 primer pairs amplified single loci producing a minimum of 2 and maximum of 11 alleles. A total of 159 alleles were detected with an average of 6.4 alleles per locus. The observed and expected heterozygosity values averaged 0.32 (0.08–0.91) and 0.74 (0.23–0.89) respectively. The UPGMA based dendrogram was able to distinguish all the accessions except two accessions from Afghanistan establishing that microsatellites could successfully detect intraspecific genetic diversity in chickpea. Further, cloning and sequencing of size variant alleles at two microsatellite loci revealed that the variable numbers of AG repeats in different alleles were the major source of polymorphism. Point mutations were found to occur both within and immediately upstream of the long tracts of perfect repeats, thereby bringing about a conversion of perfect motifs into imperfect or compound motifs. Such events possibly occurred in order to limit the expansion of microsatellites and also lead to the birth of new microsatellites. The microsatellite markers developed in this study will be useful for genetic diversity analysis, linkage map construction as well as for depicting intraspecific microsatellite evolution.

Communicated by A. Charcosset

N. K. Sethy · B. Shokeen · S. Bhatia (✉)
National Centre for Plant Genome Research, Jawaharlal Nehru University Campus, Post Box No. 10531, New Delhi, 110067, India
E-mail: sabhyatab@yahoo.com
Tel.: +91-11-26178612
Fax: +91-11-26716658

K. J. Edwards
Functional Genomics Group, OB120,
School of Biological Sciences, University of Bristol,
Woodland Road, BS8 1UG, Bristol, UK

Introduction

Chickpea (*Cicer arietinum* L.) with a genome size of 750 Mbp is a self-pollinated, diploid ($2n = 2x = 16$), and the third most important grain legume crop of the world. It is cultivated in 44 countries around the world under eight geographically diverse agro-climatic conditions but the crop is majorly cultivated in the Indian subcontinent, West Asia, the Mediterranean, North Africa and the Americas (Croser et al. 2003). Recently the crop has established a new niche in Australia and Canada for its high dietary protein content (25.3–28.9% after dehulling) and ability to fix atmospheric nitrogen in symbiosis with *Rhizobiaceae*. The vulnerability of this crop to biotic stresses (Fusarium wilt, Aschochyta blight, nematodes and pests) and abiotic stresses (drought and cold) severely reduces the yield (Croser et al. 2003). The world average yield (0.8 t ha^{-1}) is far below the yield potential (Singh 1987). Hence in order to improve the productivity of chickpea, the use of DNA-based molecular markers has been proposed for marker-assisted selection (MAS), mapping of quantitative trait loci and positional cloning of genes in chickpea (Winter and Kahl 1995).

Molecular markers, both biochemical (Ahmad et al. 1992; Labdi et al. 1996) and DNA-based like RFLP (Udupa et al. 1993), RAPD (Sant et al. 1999; Iruela et al. 2002; Sudupak et al. 2002; Chowdhary et al. 2002) were unable to address the genetic variation within chickpea. However, microsatellite based marker systems like oligonucleotide fingerprinting (Weising et al. 1992; Sharma

et al. 1995), $(TAA)_n$ repeat motif variation (Udupa et al. 1999; Udupa and Baum 2001), microsatellite-fingerprinting (Sant et al. 1999), and ISSR (Chowdhary et al. 2002) proved that microsatellites were abundant in the chickpea genome and could efficiently be used for detecting the genetic variation within chickpea cultivars (Huttel et al. 1999; Udupa et al. 1999) and for map construction (Winter et al. 1999, 2000; Flandez-Galvez et al. 2003).

The 1–6 bp tandemly repeated simple sequences in the genomes of living organisms are called microsatellites. They represent an abundant class of repeats which find ubiquitous genome wide distribution. These repeat classes are attributed with high allelic variation, co-dominant Mendelian inheritance and easy conversion to PCR based assays, which makes them an ideal marker system for all breeding purposes. In recent years microsatellites have been preferred markers in species exhibiting low levels of genetic variability, inbred populations, recently derived populations and geographically close populations. They have also been preferentially used in genome mapping projects, cultivar identification, pedigree analysis and MAS (Gupta and Varshney 2000). However, the time, cost and labor-intensive generation of functional microsatellites for a given species limits the availability of these markers (Zane et al. 2002; Squirrell et al. 2003). These limitations have been overcome by the use of microsatellite enrichment techniques, which have led to enrichment of upto 90% as compared to conventional microsatellite isolation techniques (Billotte et al. 1999; Zane et al. 2002).

For chickpea, 174 microsatellite markers have been developed and characterized by Hüttel et al. (1999), Winter et al. (1999) and another 10 by Sethy et al. (2003) so far. More recently, a set of 233 and 13 new microsatellite markers were generated for chickpea by Lichtenzveig et al. (2005) and Choudhary et al. (2006) respectively. Of all the reported STMS markers a maximum of 118 have been positioned in the interspecific map with *C. reticulatum* (Winter et al. 2000) where as a maximum of 51 STMS markers have been positioned on the intraspecific chickpea genetic linkage map (Flandez-Galvez et al. 2003). The number of available STMS markers for chickpea genome mapping is still limited and only low marker density has been achieved (Pfaff and Kahl 2003). Moreover, the reported microsatellite sequences from chickpea have been isolated either by conventional genomic library screening procedure (Hüttel et al. 1999; Winter et al. 1999; Sethy et al. 2003) or generated from BAC libraries (Lichtenzveig et al. 2005). No attempts so far have been made to isolate nuclear microsatellite markers from the chickpea genome using the enrichment procedure, though this provides an attractive choice for targeted microsatellite development. The microsatellite marker development in chickpea has been majorly targeted at AT/TA, ATT/TAA, TTA/AAT and related motifs (Winter et al. 1999), whereas $(CT/GA)_n$ and $(CA/GT)_n$ repeat motifs which are reported to be abundant and polymorphic in the chickpea genome (Weising et al. 1992; Hüttel et al. 1999; Winter et al. 1999), occurring once in every 60 kb of the genome (Hüttel et al. 1999), have been largely unexploited.

In the present study, a microsatellite enriched library was constructed from chickpea in order to isolate (CT/GA) and (CA/GT) microsatellite motifs. A set of 74 functional STMS markers were developed and characterized. The levels and patterns of intraspecific microsatellite polymorphism in a collection of chickpea accessions were studied. The molecular basis of microsatellite length variations in chickpea was also investigated.

Materials and methods

Plant material and DNA extraction

Thirty six chickpea accessions from 11 countries (Table 1) and one accession of *C. reticulatum* (ICC17121) were obtained from the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India. The *C. arietinum* cv. Pusa362 was collected from Indian Agriculture Research Institute (IARI), India. All the plants were grown at the NCPGR field site. Nuclear DNA from young leaves of Pusa362 was isolated according to Malmberg et al. (1985). Genomic DNA from the 36 chickpea accessions were isolated from 21 days grown seedlings using CTAB method (Doyle and Doyle 1987). The genomic DNA from *C. reticulatum* was isolated using the Plant DNA MiniElute kit (SigmaAldrich, USA). The quality and quantity of all DNA samples were checked on agarose gels by comparison with known amounts of uncut λ DNA.

Construction of microsatellite enriched library and screening

The microsatellite enriched library was constructed using nuclear DNA of chickpea cv. Pusa362 according to the method of Edwards et al. (1996) for the identification of $(CA/GT)_n$ and $(CT/GA)_n$ repeats. Following blue-white selection on luria-agar plates containing ampicillin, IPTG and X-gal the recombinants were grided in duplicate onto 150 mm petridishes and were transferred to Hybond N membrane (Amersham Biosciences, USA). The membranes were pre-hybridized for 3 h at 52°C in 6× SSC, 0.1% (w/v) SDS and 10 µg/ml calf thymus DNA and probed with $\gamma[^{32}P]$-ATP-labeled $(CA)_{10}$ and $(CT)_{10}$ oligonucleotides. Each oligonucleotide (1 µM) was labeled using 5 U of polynucleotide kinase (New England Biolabs, USA), denatured at 95°C for 5 min and pooled together. The filters were hybridized at 52°C for 16 h in a hybridization oven and then washed twice with 1× SSC, 1% (w/v) SDS for 15 min followed by 0.1× SSC and 1% (w/v) SDS for 15 min, at

Table 1 List of chickpea accessions used in this study

Accession	Country
1. ICC2065	India
2. ICC4872	India
3. ICC5337	India
4. ICC7255	India
5. ICC13077	India
6. ICC14778	India
7. ICC16487	Pakistan
8. ICC4495	Turkey
9. ICC8261	Turkey
10. ICC11879	Turkey
11. ICC10885	Ethiopia
12. ICC12537	Ethiopia
13. ICC12866	Ethiopia
14. ICC13892	Ethiopia
15. ICC7712	Spain
16. ICC13780	Spain
17. ICC16834	Spain
18. ICC6802	Iran
19. ICC9137	Iran
20. ICC13357	Iran
21. ICC5504	Mexico
22. ICC12037	Mexico
23. ICC14199	Mexico
24. ICC6263	Russia and CIS
25. ICC6306	Russia and CIS
26. ICC7668	Russia and CIS
27. ICC8740	Afghanistan
28. ICC9848	Afghanistan
29. ICC9895	Afghanistan
30. ICC3429	Egypt
31. ICC5092	Egypt
32. ICC7229	Egypt
33. ICC9590	Egypt
34. ICC7716	Greece
35. ICC8511	Greece
36. ICC11859	Greece

The ICRISAT accession number and source country are mentioned

52°C. The filters were then exposed to X-ray films (Kodak India Limited). Those recombinants producing intense signals were selected for sequencing.

Microsatellite identification and primer design

Plasmid DNA from the putative positive clones was extracted from 5 ml of overnight culture using alkaline lysis method (Sambrook et al. 1989). The plasmids were sequenced using the Big Dye Terminator reaction kit on the ABI 3700 prism automated DNA sequencer (Applied Biosystems, USA). The microsatellite motif in each sequence was identified and each sequence was compared against the local and NCBI sequence databases in order to identify the redundant clones. The unique sequences were retained and submitted to GenBank to obtain the accession numbers, AY446327–AY446401 (Table 2). Primers were designed based on the sequences flanking the microsatellites motifs using the PRIMER 3.0 software (Rozen and Skaletsky 1998) using the following criteria: (1) primer length 18–26 nucleotide with an optimal length of 20 nucleotides (2) primer $T_m = 50–65$°C with an optimal of 55°C (3) amplified PCR product size 100–300 bp, and (4) an optimal 40% G + C content. The presence of structures such as hairpin or short repeat motifs were also considered while designing the primers. All primers were synthesized from BioBasic Inc., Canada.

PCR amplification of microsatellite loci and electrophoresis

Functionality of each primer pair was tested by amplification of Pusa362 genomic DNA. PCR amplification was done in 20 µl reaction volume containing 200 mM Tris–HCl (pH 8.4), 500 mM KCl, 2 mM $MgCl_2$, 0.2 mM of each dNTPs, 0.5 µM of each primer, 40 ng of genomic DNA and 0.5 units of Taq DNA polymerase (Invitrogen, USA). Reactions were carried out in an Icycler (BIORAD Laboratories, USA) thermal cycler using the touchdown amplification profile: an initial denaturation at 94°C for 3 min followed by 18 cycles of 94°C for 30 s, 64°C for 50 s with 0.5°C decrease in each subsequent cycle, 72°C for 50 s. This was followed by 20 cycles of 94°C for 30 s, 55°C for 50 s, 72°C for 50 s and final extension of 10 min at 72°C. Amplification products were separated on either 3% Metaphor agarose (Cambrex, USA) gels or 6% polyacrylamide gels and stained with ethidium bromide. All the primer pairs were used to amplify DNA from nine chickpea accessions and one *C. reticulatum* accession (ICC17121) in order to establish their usefulness for intra- as well as interspecific polymorphism detection. Of these, 25 polymorphic primer pairs were further used to amplify DNA of 36 geographically diverse chickpea accessions (Table 1) in order to assess the informativeness of the STMS primers for detecting microsatellite polymorphism within chickpea.

Data analysis

For each microsatellite locus, sizes of the alleles were estimated by comparison with standard size DNA markers and scored across all the 36 chickpea accessions both manually and with the help of gel documentation system (Alpha Innotech Corp., USA) in a binary matrix where 1 represented the presence of a band and 0 the absence of common bands. Genetic parameters like alleles per locus, effective alleles per locus (N_e), observed heterozygosity (H_O), expected heterozygosity ($He = 1-\Sigma p_i^2$, where p_i is the frequency of the ith allele), Fixation index (F_{IS}) and Shannon's information index (I) were calculated using POPGENE Version 1.32 (Yeh and Boyle 1997). Pairwise genetic similarity was calculated among the 36 accessions using Jaccard's similarity coefficient. The similarity matrix ($D = 1-S$) was used for constructing the dendrogram using the UPGMA (Unweighted Pair Group Method with Arithmetic Averages) algorithm on NTSYS-pc (version 2.1, Rohlf

Table 2 Characterization of chickpea STMS markers in *C. arietinum* cv. Pusa362

Sl no	Locus	Primer sequence (5′–3′)	Repeat motif	Length (bp)	N_A	GenBank accn. no.
1	NCPGR21	TCTACCTCGTTTTTCGTGCC/TTGCTCCTTCAACAAAACCC	$(CT)_{15}$	137	2	AY446327
2	NCPGR23	CCAAAGGTGCGATGAAAATC/GCGCGTGGACTACTCTTCAT	$(GT)_5gc(GT)_4$	184	1	AY446328
3	NCPGR27	ACCCCATTTTTGGGTTTTCT/TGCATCCAACTTGTTGTCTTG	$(CA)_{12}$	278	1	AY446329
4	NCPGR28	TGATGGAAGGTGATGTGGAA/GAGGGGGAAACGTTTTCTTT	$(AT)_6(GT)_5gc(GT)_7$	224	3	AY446330
5	NCPGR29	ATTCCTCACCACCATCAAGG/TCGTGAGCCGAAATACAAGA	$(CA)_{10}$	245	1	AY446331
6	NCPGR30	GGCAATGAGCAACTTTTCCT/GCTTTGAAAAACGGGGTGT	$(CA)_{14}(CGCA)_3(CA)_2N_{58}(CA)_{10}$	173	1	AY446332
7	NCPGR32	CGCAGGTAAAGCTCCTCTCA/CCCCTTTTTCACCCTGTAAG	$(CA)_{12}$	250	1	AY446334
8	NCPGR33	ACATCTTGAAGTGCCCCAAC/TGCAAGCAGACGGTTACAAG	$(GA)_{20}$	248	1	AY446335
9	NCPGR34	TGGAAGGTGTTTTAGTGGGTG/GACTAACTGGCCCCCAAAA	$(CT)_{17}$	240	1	AY446336
10	NCPGR36	GTGGAGCCAAAAATCGACAT/AACTTTATTTTCATTTGTCCATCAA	$(GT)_{12}$	204	3	AY446337
11	NCPGR37	AGTCGCATCTCTGCCAAAGT/CATTCCCTGACCTGCTGC	$(CT)_2gc(CT)_{19}$	191	1	AY446338
12	NCPGR38	CGGTTAGTTGGGGACTTTCA/CCCGACGTAAAATGGAAGAA	$(CA)_{11}(CT)_8$	197	1	AY446339
13	NCPGR39	GCGCGTGGACTACTCATGT/GGTGGCTGCCTTTTCTTTC	$(GA)_{17}taaa(GA)_2ta(GA)_8$	152	2	AY446340
14	NCPGR40	TGAACGAATCATGGCAAGAG/GCCCTCCTTCTTGCTTACAA	$(GA)_{12}gt(GA)ca(GA)_3gtaagt(GA)_4gtgg$ $(GA)_{10}gt(GA)gt(GA)_6$	193	2	AY446341
15	NCPGR41	GGGAGGAGGATCAAAATTAC/CAACTATAAAGAGGCATGTTCC	$(CT)_8(CA)_{17}$	262	1	AY446342
16	NCPGR42	CCCCTAGTAGCAAATATTTTGACC/TTTGAATGCATTTCTTCATAGCA	$(CT)_{27}$	170	1	AY446343
17	NCPGR43	GAAGTCGAGATGCTGAAAAG/AATTCTAGAAGGGAAGGGTG	$(CT)_{12}at(CT)_3$	255	1	AY446344
18	NCPGR44	AAATGTTTTGTATGATCGGG/CTAAACAAGTGCATACAGCG	$(CT)_{25}$	212	1	AY446345
19	NCPGR45	TGTTTTCAAATCAAACAGGC/GATACACACCAAGGCACAGT	$(CT)_2gtcat(CT)_5cc(CT)_2cc(CT)_{17}$	223	1	AY446346
20	NCPGR46	CCCAAAACTGAAATGGAAAC/GGCAGTTACTACCAAGGCAT	$(CT)_6at(CT)_6at(CT)(CA)_{14}$	217	1	AY446347
21	NCPGR47	AGCGGAACCTTGTTTACATA/ATTCTTTAGGGATACCAGCC	$(CT)_{19}t(CT)_3$	147	1	AY446348
22	NCPGR48	TGGGCTATGAATTAAGATGG/TAATTGATGAGGGAGAGAGC	$(CT)_{22}N_{13}(CT)_3cg(CT)_3$	206	1	AY446349
23	NCPGR49	TGTCTCAGATGTAGTGGCCT/TTCATGGCAAGAAATTGAAC	$(CT)_{17}c(CT)_3ttcttt(CT)_6tt(CT)_7$	234	2	AY446350
24	NCPGR50	ATGATGGATTTTCGGAATGT/AAAAATGCTGGAAGGAACTG	$(GA)_{26}$	209	2	AY446351
25	NCPGR51	CATAATGCAAGGGCAATTAG/CTCTTATCTTCATGTTGCCG	$(GA)_{20}$	203	1	AY446352
26	NCPGR52	CAAGCTCTTTCAGAATTTGC/TACTGGTGGAAAAATGGATG	$(GA)_2aa(GA)_{25}$	245	1	AY446353
27	NCPGR53	CCCTCCTTCTTGCTTACAAA/TAATGGTGAACGAATCATGG	$(CT)_5ca(CT)ca\ (CT)_{10}ca(CT)_4\ ca(CT)ta$ $(CT)_4\ gtca(CT)_{12}$	194	2	AY446354
28	NCPGR54	GAAGTCGAGATGCTGAAAAG/AATTCTAGAAGGGAAGGGTG	$(CT)_{16}$	255	1	AY446355
29	NCPGR55	TCCATTGGATACATCACAGG/GGGCAAATTCAGTATTTTGG	$(GA)_{16}$	204	1	AY446356
30	NCPGR56	CATGACAATAATGGTGAACG/GATCTTGACTTCTGTTTGTGC	$(GA_{)12}gt(GA)ca\ (GA)_3N_6(GA)_4gt\ (GA)_9gt$ $(GA)gt\ (GA)_6$	162	1	AY446357
31	NCPGR57	CGATGATATTCTCAGCGAAC/TGTATGAAAACACTTTGACTCATT	$(GT)_{14}(GA)_{24}\ (GT)_2$	219	1	AY446358
32	NCPGR58	TGAAGATCTCCAACGGTAAC/TTTCTTTTGATGTGTTCTTGG	$(GT)_{13}(GA)_{27}$	215	4	AY446359
33	NCPGR59	CTTGACCAGAGGCATTTATC/AACATAATGGTGTCCAAAGC	$(GT)_{12}(GA)_{13}$	267	1	AY446360
34	NCPGR60	AGAAATCACAAACCTCTTCG/GCTTGGATCTTCAAAACTTG	$(GT)_2c(GT)cgat\ (CT)_5ca(CT)_{26}$	259	3	AY446361
35	NCPGR61	AAATGGTTTGAGAGGTGATG/AAGGGAGAAGGAAGAAAAGA	$(AT)_5gtat\ (GTCT)_2(GT)_{20}$	226	3	AY446362
36	NCPGR62	TCTTAGACTCGGACCTGGTA/TTCGTTTTTCTCTTACGCTC	$(GT)_{14}(GA)_4\ gg(GA)_2$	295	2	AY446363
37	NCPGR63	CTCTCTTCTCCTCCCAAAAT/GGGGATTTGTTTTAAGTCCT	$(GA)_{11}gc(GA)_{12}$	270	1	AY446364
38	NCPGR64	GGCGCGTGGACTAACTAAT/CACATGATTTCTACATGGTGTT	$(GT)_{23}$	239	1	AY446365
39	NCPGR65	CGTGGACTAACGTTCACTG/GGTTCTCCCTCTGTTCTTCT	$(GA)_7gg(GA)_{29}$	296	1	AY446366
40	NCPGR66	GAAGCCATTGTTTTGTTGGT/ATTATAGCACCCCGCAAC	$(CT)_{24}(CA)_2$	281	1	AY446367
41	NCPGR67	GCGTGGACTAACTAGAGGTC/ATGGAATCCAGGACGTTAAT	$(AC)_{40}(AT)_6$	300	1	AY446368
42	NCPGR68	GCGTGGACTACTCACTAGCA/GAAAAACGGGGTGTTACATT	$(CA)_{11}(CGAG)_4(CA)_2N_{34}(CA)_3N_{16}(CA)_{10}$	195	4	AY446369

Table 2 (Contd.)

Sl no	Locus	Primer sequence (5′–3′)	Repeat motif	Length (bp)	N_A	GenBank accn. no.
43	NCPGR69	GACCGAATGTCCATAAATCA/GGAGCTGGAAAAACTACAGC	$(GA)_{36}$	252	1	AY446370
44	NCPGR70	TATCCAAAGCACATCTCACA/CTTAGTTTTGGTAGGGGGTC	$(GA)_{33}ta(GA)ta(GA)_{10}$	270	1	AY446371
45	NCPGR71	CAACGACACTTAAGCAATCA/AGCAATTGGCCTTACATTAG	$(GA)_{38}$	249	1	AY446372
46	NCPGR72	TTAACCCATTAGCGTGACTT/GATCAGCTTCTTGCTTTCAT	$(GA)_{21}$	250	1	AY446373
47	NCPGR73	GGATGAACGAGAGTTGGTAT/TATCTGTCAATTGAGTTGCG	$(CA)_{10}ga(CA)ta(CA)_{10}cg(CA)_{30}aa(GA)_2$	282	1	AY446374
48	NCPGR74	TCCGTCCACACATTTCTACT/CTTTTAGTTGGTCGAAGCC	$(GA)_{39}aa(GA)_2$	231	2	AY446375
49	NCPGR75	AACTGAAATGGAAACACAGG/GAAAGCTGACTCCTCTACCA	$(CT)_{15}(CA)_{14}$	192	1	AY446376
50	NCPGR76	GAAAGCTGACTCCTCTACCA/GAAAATGCTCTCAGTCAAGG	$(GT)_{13}(GA)_2ta(GA)_7ta(GA)_6$	245	3	AY446377
51	NCPGR77	TGGACTAACAAATACGACGA/AGGCCACCCTAAATTTTATT	$(GA)_{36}ta(GA)ta(GA)_{10}$	225	1	AY446378
52	NCPGR78	CTCTGTGAGGAGGAAGATGA/AGAAGTTAAAGCAATGCACC	$(CT)_{21}$	235	1	AY446379
53	NCPGR79	ATTGGTTTGAGAAGTGATGG/AGAAAAGATGGAGTTCGTGA	$(GT)_{14}gc(GT)gc(GT)gc(GT)_6$	203	2	AY446380
54	NCPGR80	TGGACTAACCCTTCTTTCTTC/TTATATTATGCAGGACCGCT	$(AT)_4(AC)_2at(AC)_{24}$	256	1	AY446381
55	NCPGR81	CCGAATGTCCATAAATCAAT/TGTTTGACTGGGATAACTCC	$(AG)_{31}$	211	1	AY446382
56	NCPGR82	ATTGGTTTGAGAAGTGATGG/GGGAAGTTCAGGACTCTTTT	$(GA)_{29}$	280	1	AY446383
57	NCPGR83	GCTTGACCTATTTATGGTCTG/AGGTGATGTGGAAATGATGA	$(CA)_{45}cg(CA)cg(CA)_{15}$	314	1	AY446384
58	NCPGR84	GCATGAGAGATGGGTCATTA/GAGGCGCGTGGACTAACA	$(AT)_8(GT)_{19}$	216	1	AY446385
59	NCPGR85	CGCGTGGACTAACATAGAGT/GAAGTGGGTGTTGTGTTTTT	$(CA)_{10}aag\ (CT)_4tt(CT)_4$	180	1	AY446386
60	NCPGR86	CTACTGCAGAAAAATCAGGG/ATAGTTCTTGACCAGAGGCA	$(CT)_{13}(CA)_{11}$	208	2	AY446387
61	NCPGR87	CCTTGGATTTCTTTTCTTGA/ATTGGTTTGAGAAGTGATGG	$(CA)_9cg(CA)cg(CA)_{16}$	243	3	AY446388
62	NCPGR88	AAAATCAAAGTTGGGGGTAT/CGGGGACTAACTTTTCATCT	$(GA)_{19}$	177	2	AY446389
63	NCPGR89	AAAGGGCCTTCAAGTTGTAT/ACTTTTGGAGTGAGAGGCT	$(GA)_{28}$	263	1	AY446390
64	NCPGR90	TAGCATACCATTGTCAACCA/AAGAGCACATACGGTTTTGT	$(GA)_{36}$	204	2	AY446391
65	NCPGR91	ATTGAATCCTTTCTGAACCG/CTGTTCTCTTTTCTCCTCCG	$(AG)_{12}at\ (AG)_{17}$	266	1	AY446392
66	NCPGR92	GCTTGGATCTTCAAAACTTG/GATTAGAAATCCCCAACCTT	$(GA)_{33}$	265	1	AY446393
67	NCPGR93	CAAAGTTTGTTGCTAGGATTC/GAAGATCTCCGACGATGATA	$(CA)_2\ (CT)_{24}\ (CA)_{13}$	299	2	AY446394
68	NCPGR94	GGTTTGATGTGTTCTTGGCT/CCCTCAATTCCCTCGATTTA	$(CT)_{25}$	176	1	AY446395
69	NCPGR95	AGCCCTTGGATTTCTTTTCT/GTCAATTGAGTTGCAACGAG	$(CA)_{10}(CGCA)_3(CA)_{13}$	282	3	AY446396
70	NCPGR96	CGTGGACTAACCAAACAAAT/CAAGGTCAATTCGTAGAAGG	$(CA)_{11}ga(CA)_5ta(CA)_5cg(CA)_{33}(TA)_6$	268	3	AY446397
71	NCPGR97	CGCGTGGACTAGCAAACA/AAGGAATACGAATCATCCCT	$(GA)_{24}gg(GA)_{14}$	262	1	AY446398
72	NCPGR98	CATCTTATTTTTCATTTTAGAGGAGG/AGGAAGTGTTATGGAGATGCC	$(GA)_{20}gg(GA)_{14}$	141	1	AY446399
73	NCPGR99	ATCATGAAGCAAATCCTCAC/TGAACCCAACATAGCATACA	$(GA)_{18}$	227	1	AY446400
74	NCPGR100	CCATTTTCTACAATCTCATGTCT/GTAGAAAGAGCCAAGAGGCA	$(CT)_{15}N_{42}(CT)_2cc(CT)_5tt(CT)_6at(CT)_7$	263	2	AY446401

The microsatellite repeat motifs, primer sequences, expected size of the amplification product (bp) and numbers of amplified alleles (N_A) along with GenBank numbers are mentioned

1998) software. The reliability of the dendrogram was evaluated with 1,000 bootstraps using the WinBoot software (Yap and Nelson 1996).

Sequencing of microsatellite length variants

Fragments representing allelic size variants at NCPGR81 and NCPGR91 loci were eluted from polyacrylamide gel, cloned into pGEM-T vector (Promega, USA) and transformed into DH5α cells. Following blue-white selection, plasmids were isolated and sequenced as mentioned earlier. For each individual fragment, four to six random clones were sequenced. Multiple sequence alignments were performed using the sequence data of variant alleles along with the originally cloned allele from Pusa362 using CLUSTAL W (1.82) at EBI. All the sequences were deposited in the GenBank to obtain the accession numbers (see figure legends).

Results

Characterization of microsatellites from the chickpea enriched library

The microsatellite enriched library was constructed by digestion of Pusa362 nuclear DNA with *Rsa*I. For enrichment, the digested genomic DNA was hybridized with membrane bound synthetic (CA) and (CT) oligonucleotides, followed by elution and ligation with the modified cloning vector pJV1 (Edwards et al. 1996). This yielded a large number of recombinants of which 1,000 clones were subjected to screening using radio-labeled $(CA)_{10}$ and $(CT)_{10}$ oligonucleotide probes. Screening resulted in the identification of 135 putative recombinants, which were sequenced.

Sequence analysis of the 135 putative clones revealed that all (100%) clones contained microsatellite motifs. However, 34 sequences were found to be redundant and 10 sequences had less than 5 repeat units and hence these 44 sequences were not considered for primer designing. As expected, the remaining 91 microsatellite sequences comprised of $(GA/CT)_n$ and $(GT/CA)_n$ rich motifs and were classified as 28 perfect, 25 imperfect, 24 compound and 14 interrupted microsatellites based on the structural organization of the repeat motifs. The homology searches with BlastN tool at NCBI database showed no significant homology with any of the earlier reported microsatellite sequences. Primers for these unique loci were designed based on the regions flanking the microsatellite motifs. However, of the 91 SSR motifs, primers could be designed only from 83 sequences as the remaining 8 were too close to the cloning site. The 83 primers were validated by amplification of *C. arietinum* cv. Pusa362 DNA, out of which 9 primers either did not amplify or produced anomalous fragments. Finally 74 functional STMS primer pairs were identified and their sequences are listed in Table 2.

Table 3 Number of alleles (N_a), effective number of alleles (N_e), size range of amplified products, observed heterozygosity (Ho), expected heterozygosity (He), Shannon's informative index (I) and Fixation index (F_{IS}) values obtained after screening of 36 chickpea accessions at 25 microsatellite loci

Locus	N_a	N_e	Size range (bp)	Ho	He	I	F_{IS}
NCPGR21	9	7.48	131–159	0.40	0.87	2.18	0.53
NCPGR33	5	4.36	245–258	0.51	0.78	1.51	0.33
NCPGR37	7	3.76	165–219	0.24	0.74	1.53	0.66
NCPGR41	8	5.14	241–268	0.10	0.81	1.83	0.86
NCPGR44	3	2.05	212–221	0.18	0.51	0.84	0.63
NCPGR45	6	4.07	163–226	0.32	0.76	1.51	0.57
NCPGR46	4	3.07	210–224	0.24	0.68	1.22	0.63
NCPGR48	7	5.52	206–228	0.35	0.83	1.83	0.57
NCPGR50	6	3.79	196–268	0.91	0.74	1.41	-0.24
NCPGR51	10	5.24	188–285	0.70	0.82	1.86	0.13
NCPGR52	8	5.35	223–269	0.10	0.82	1.91	0.86
NCPGR53	4	3.10	184–206	0.67	0.68	1.22	0.00
NCPGR55	6	4.89	204–226	0.08	0.80	1.63	0.89
NCPGR56	5	2.30	159–198	0.10	0.57	1.03	0.80
NCPGR57	7	5.80	205–277	0.27	0.83	1.84	0.67
NCPGR60	4	2.75	228–278	0.43	0.64	1.05	0.32
NCPGR67	8	7.00	277–314	0.13	0.86	2.01	0.84
NCPGR80	6	4.96	243–272	0.40	0.80	1.67	0.49
NCPGR81	9	6.05	175–219	0.35	0.84	1.86	0.57
NCPGR86	2	1.30	138–218	0.27	0.23	0.39	-0.15
NCPGR89	3	2.75	227–292	0.08	0.64	1.05	0.87
NCPGR90	11	8.88	155–228	0.24	0.89	2.28	0.72
NCPGR91	7	5.08	228–262	0.35	0.81	1.84	0.56
NCPGR94	8	6.36	178–210	0.32	0.85	1.96	0.61
NCPGR98	6	4.36	138–168	0.40	0.78	1.58	0.47
Average	6.36	4.62	–	0.32	0.74	1.56	–
St. Dev.	2.48	1.79	–	0.20	0.14	0.44	–

This library was enriched only for two dinucleotide repeats (CA and CT) and of the 74 characterized loci, CT repeats were found to be more abundant (43 clones, 58.6%) as compared to CA repeats (21 clones; 28.4%). The remaining 10 clones (13.5%) contained CA and CT compound motifs. The numbers of microsatellite motifs were found to be highly variable at these loci with the maximum number of uninterrupted GA and GT units being 39 (NCPGR 74) and 45 (NCPGR 83), respectively. However, the majority of the repeat motifs comprised of 12–30 repeat units (Table 2).

Microsatellite polymorphism within chickpea

The functionality of the 74 STMS primer pairs was established by validation in *C. arietinum* cv. Pusa362 in which all the primers amplified the expected sized fragments. Single alleles were amplified with 50 primer pairs, whereas 24 primers (33%) produced 2–4 alleles (Table 2). The 74 STMS markers were also used to amplify genomic DNA from 9 accessions of chickpea and 1 of *C. reticulatum*. All the primers amplified allelic fragments in *C. arietinum* where as only 69 (93%) amplified in *C. reticulatum*. Comparison of chickpea with *C. reticulatum* amplification profiles revealed that 53 primers were polymorphic between the two species. Intra-specific polymorphic amplification profiles were obtained with 49 primer pairs whereas 25 primer pairs produced monomorphic banding patterns (Data not shown). Of these 49 polymorphic primer pairs in chickpea, a set of 25 were further used to analyze the genetic diversity in the 36 chickpea accessions listed in Table 1. A total of 159 alleles were produced at the 25 loci with an average of 6.4 alleles and 4.6 effective alleles per locus (Table 3). A maximum of 11 alleles were detected for primer pair NCPGR90 while minimum of 2 alleles were detected for NCPGR86 with the alleles varying in size from 131 to 314 bp. The expected heterozygosity values ranged from 0.23 to 0.89 (mean 0.74) and the observed heterozygosity values ranged from 0.08 to 0.91 with an average of 0.32 (Table 3). Based on heterozygosity values, excess of homozygotes were observed for 22 loci where as for two loci (NCPGR50 and NCPGR86) heterozygote excess were observed apparent from the Fis values (Table 3).

Genetic relationships among chickpea accessions

The allelic data obtained with the 25 primer pairs across 36 accessions was scored and computed to obtain the dendrogram which was able to clearly distinguish all the cultivars except the two accessions, ICC9848 and ICC9845 from Afghanistan (Fig. 1). No distinct clustering pattern was obtained, however, accessions from the same country or geographically close locations seemed to cluster together indicating their genetic similarity (Fig. 1). The three Indian accessions (ICC5337, ICC7255 and ICC13077) and one accession each from Turkey (ICC11879) and Pakistan (ICC16487) were most diverse forming two distinct clusters, distant from all other accessions. Most of the nodes, except a few, were supported with low bootstrap values (≤ 50) indicating weak clustering.

Sequence analysis of microsatellite alleles

In order to assess the basis of length variation at microsatellite loci in chickpea, two loci namely NCPGR81 and NCPGR91, comprising of long arrays of $(GA)_n$ repeats and amplifying 9 and 7 size variant alleles respectively in the 36 chickpea accessions, were identified. Fragments corresponding to all the allelic size variants (ranging from 175–219 bp for the NCPGR81 locus and 228–262 bp for the NCPGR91 locus) were cloned and sequenced (Figs. 2, 3) to investigate the patterns of length variation (a) within multiple alleles in the same accession and (b) between size variant alleles across different accessions. Sequence alignments revealed that the size variant alleles across various accessions as well as the multiple alleles within one accession possessed variable number of repeat units with conserved primer binding sites. The number of $(AG)_n$ repeat units at NCPGR81 locus varied from 13 (ICC11879B) to 34 (ICC7712B) units (Fig. 2). In alleles having long stretches $(AG)_n$ of the microsatellite motifs, for example accessions ICC7712A and ICC11879A (32 repeats each), ICC10885 (35 repeats), a point mutation from G $\rightarrow$ A in the ninth repeat motif (position 124; Fig. 2) was observed which converted the perfect repeat into an imperfect one. Moreover, in the ''B'' allele of Spanish accession ICC7712, a point mutation from A $\rightarrow$ T (position 104) upstream of the AG repeats was observed resulting in the formation of a compound repeat structure $(TA)_3(GA)_{33}$ at this locus (Fig. 2). Sequence analysis of the length variant alleles at the NCPGR 91 locus also revealed the presence of variable numbers of $(AG)_n$ repeats. At this locus also, alleles containing more than $(AG)_{17}$ repeats were converted into an imperfect $(AG)_n$at$(AG)_n$ motif (Fig. 3). In addition, isolated point mutations were observed at positions 28, 46 and 123, in accessions ICC8740, ICC12866 and ICC13780, respectively (Fig. 3).

Discussion

Microsatellite enrichment in chickpea

Microsatellite markers are characterized by a high degree of variability making them powerful tools for population genetic analyses. This class of markers was reported to be abundant and polymorphic in the chickpea genome (Weising et al. 1992; Sharma et al. 1995; Huttel et al. 1999; Udupa et al. 1999; Litchtenzveig et al. 2005), yet the numbers of available microsatellite markers in chickpea are not enough for the saturation of

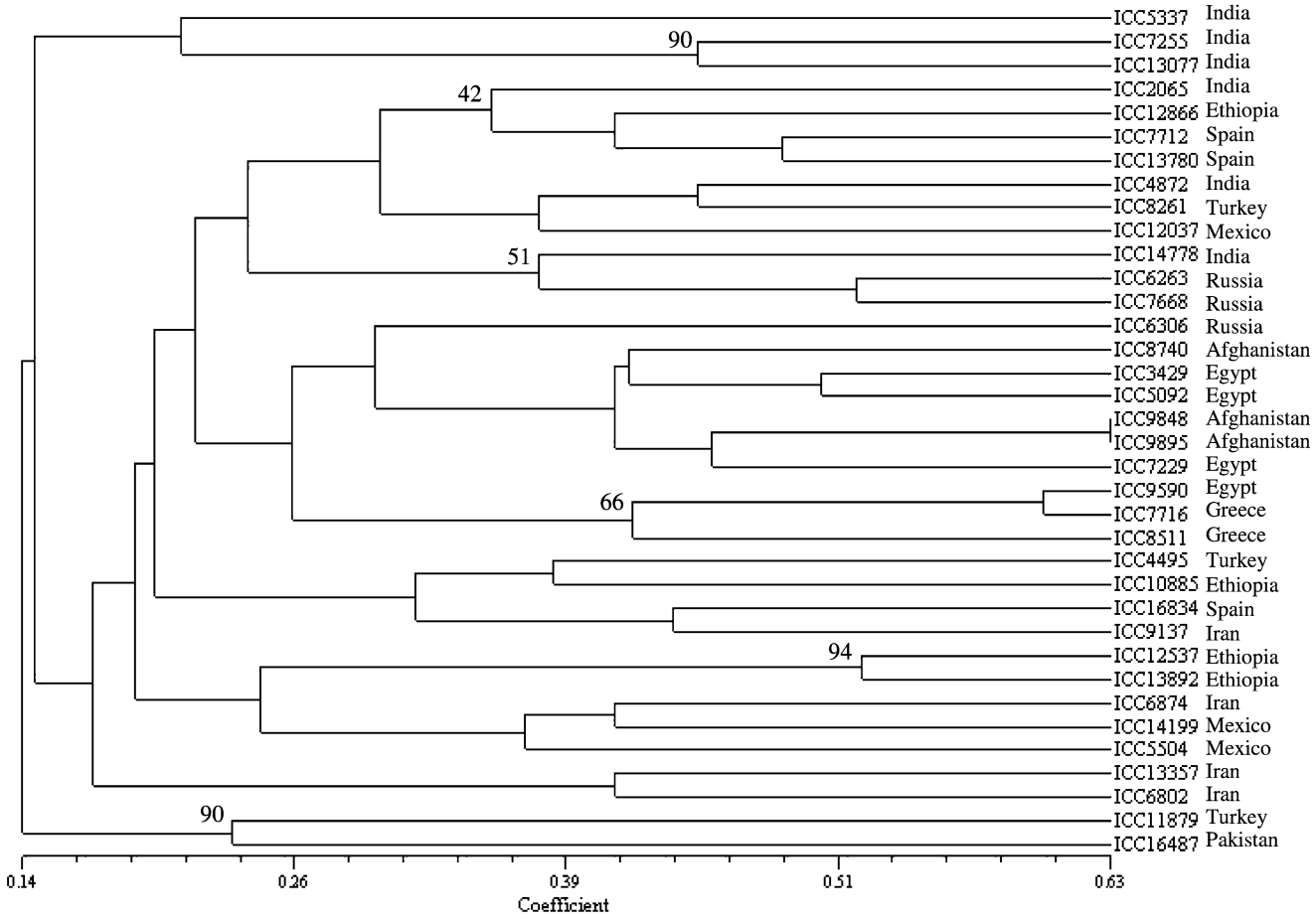

Fig. 1 The UPGMA dendrogram obtained using the Jaccard's coefficient. Thirty-six chickpea accessions were analysed using 25 STMS primer pairs. Names of cultivars and source country are mentioned. Bootstrap values (> 40) are indicated at major nodes

the linkage map (Pfaff and Kahl 2003; Lichtenzveig et al. 2005). Therefore, the present study was undertaken with the aim to develop STMS markers from microsatellites isolated from an enriched library and use them for analysis of genetic diversity and genome mapping.

The conventional protocols used for the isolation of microsatellites are cost, time and labor intensive and the efficiency of microsatellite isolation is low ranging from 0.045% to a maximum of 12% (Zane et al. 2002). The isolation of microsatellites from plants becomes technically even more demanding as their frequency in plant genomes is low as compared to animal genomes (Maguire et al. 2000; Squirrell et al. 2003). To overcome these key obstacles in microsatellite development, several enrichment methods have been developed (Zane et al. 2002). The enrichment techniques, which are based on the principle of capturing microsatellites from genomic DNA by hybridization with synthetic oligonucleotides bound to Nylon membranes or magnetic particles (Zane et al. 2002) increase the efficiency of microsatellite identification by upto 90% thereby considerably reducing the cost of microsatellite marker development. Microsatellites from diverse plant species have been isolated efficiently (Gaitan-Solis et al. 2002; Riaz et al. 2004; Lowe et al. 2004) using the membrane based enrichment protocol developed by Edwards et al. (1996). This method, which has the advantage of specifically targeting certain classes of microsatellites, enabled us to identify 74 CT/GA and CA/GT microsatellite motifs. The occurrence and frequency of these two dinucleotide repeats had been reported in the chickpea genome (Huttel et al. 1999; Lichtenzveig et al. 2005). Our studies further established that the abundance of the CT motifs was double that of the CA motifs. The relative abundance of CT microsatellites have been shown to be more abundant than CA motifs in most plants such as *Arabidopsis*, rice, wheat, maize and *Brassica* (Gupta and Varshney 2000). Moreover, it has been shown in *A. thaliana* that the GA/CT microsatellite motifs occur more frequently in the 5′ flanking regions of genes as compared with the whole genome (Zhang et al. 2004), thereby increasing the likelihood of their association with desirable traits.

Intraspecific polymorphism with STMS markers

A set of 74 functional STMS primer pairs were developed and validated in chickpea as well as *C. reticulatum*.

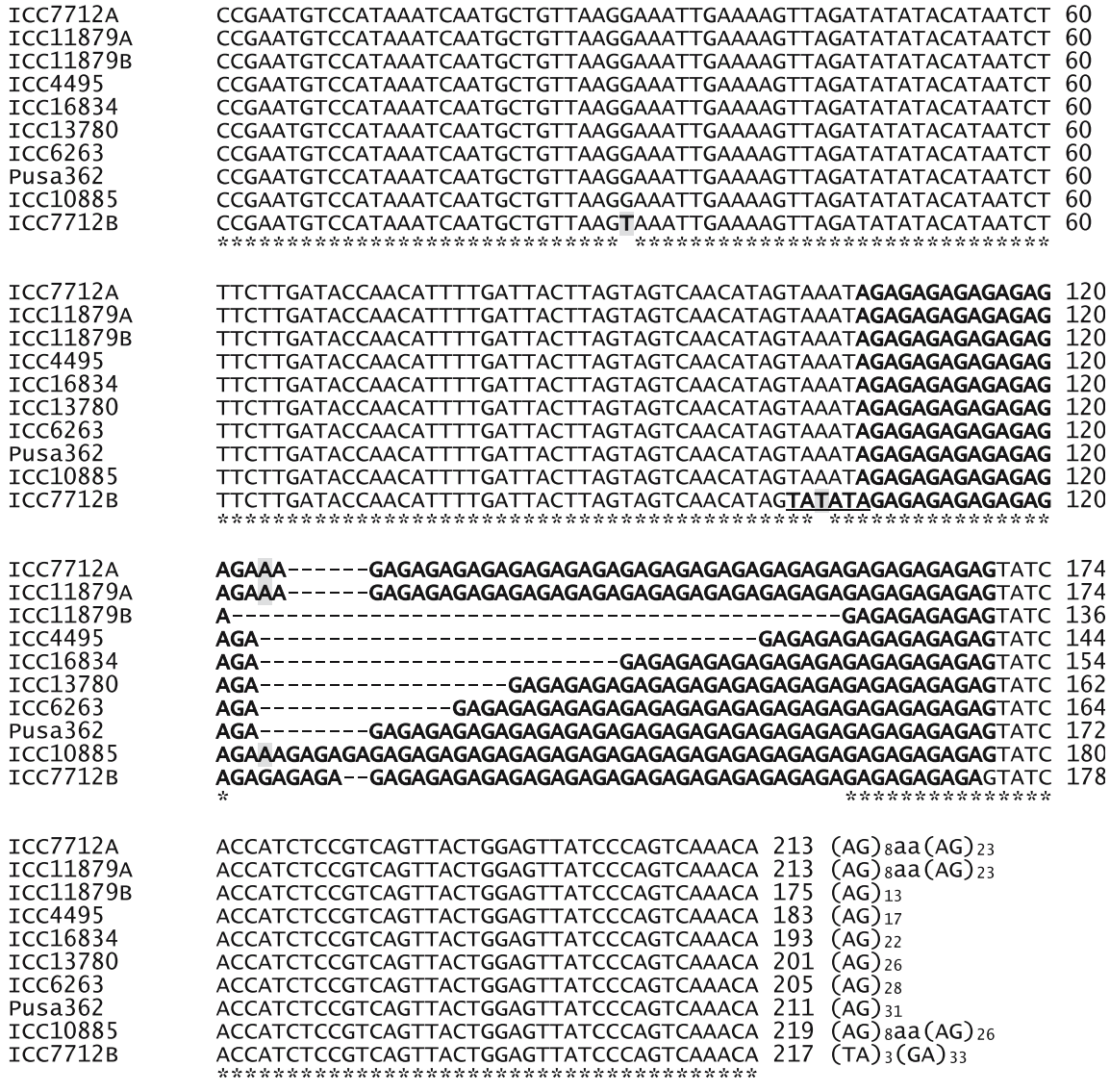

Fig. 2 Multiple sequence alignment of the size variant alleles from Pusa362 and other chickpea accessions (AY973983–AY973991) at the NCPGR81 locus. The accession numbers and variant alleles (A, B) are mentioned as capital roman alphabets. The *asterisks* indicate similar sequences and *dash* indicate alignment gaps. The repeat region is indicated in *boldface*, compound motifs are *underlined* and *shadowed boxes* indicate point mutations

This analysis revealed 66% intraspecific polymorphism within chickpea in comparison to only 40% reported while using the microsatellites developed earlier (Winter et al. 1999; Litchtenzveig et al. 2005). The higher levels of polymorphism revealed in our studies might be attributed to our preferential isolation of the GA/CT repeat motifs which have been reported to be highly polymorphic in other plant systems like rice, bean, tomato and peanut (Cho et al. 2000; Gaitan-Solis et al. 2002; He et al. 2003; Ferguson et al. 2004). A total of 69 chickpea STMS markers (93%) also amplified allelic microsatellite loci in *C. reticulatum* and generated a high interspecific polymorphism (53 markers, 77%) between chickpea and *C. reticulatum*. These primers will aid the interspecific chickpea mapping projects arising from crosses between *C. arietinum* × *C. reticulatum*.

The high levels of intraspecific genetic polymorphism in chickpea was clearly evident from the dendrogram. Moreover, the high average number of alleles amplified per locus (6.4) combined with the average observed heterozygosity values of 0.3297 suggest that considerable polymorphism is present at these microsatellite loci in chickpea. Extensive polymorphism at the chickpea TAA and TTA loci has been well demonstrated earlier (Huttel et al. 1999; Udupa et al. 1999,2001; Lichtenzveig et al. 2005). Our studies along with other studies in chickpea, clearly establish that microsatellite markers could be reliably used for assessing genetic diversity and cultivar identification. It was also observed in the dendrogram that most nodes were supported by low bootstrap values and there was no congruence between the clustering pattern and the geographical origin of the cultivars. This

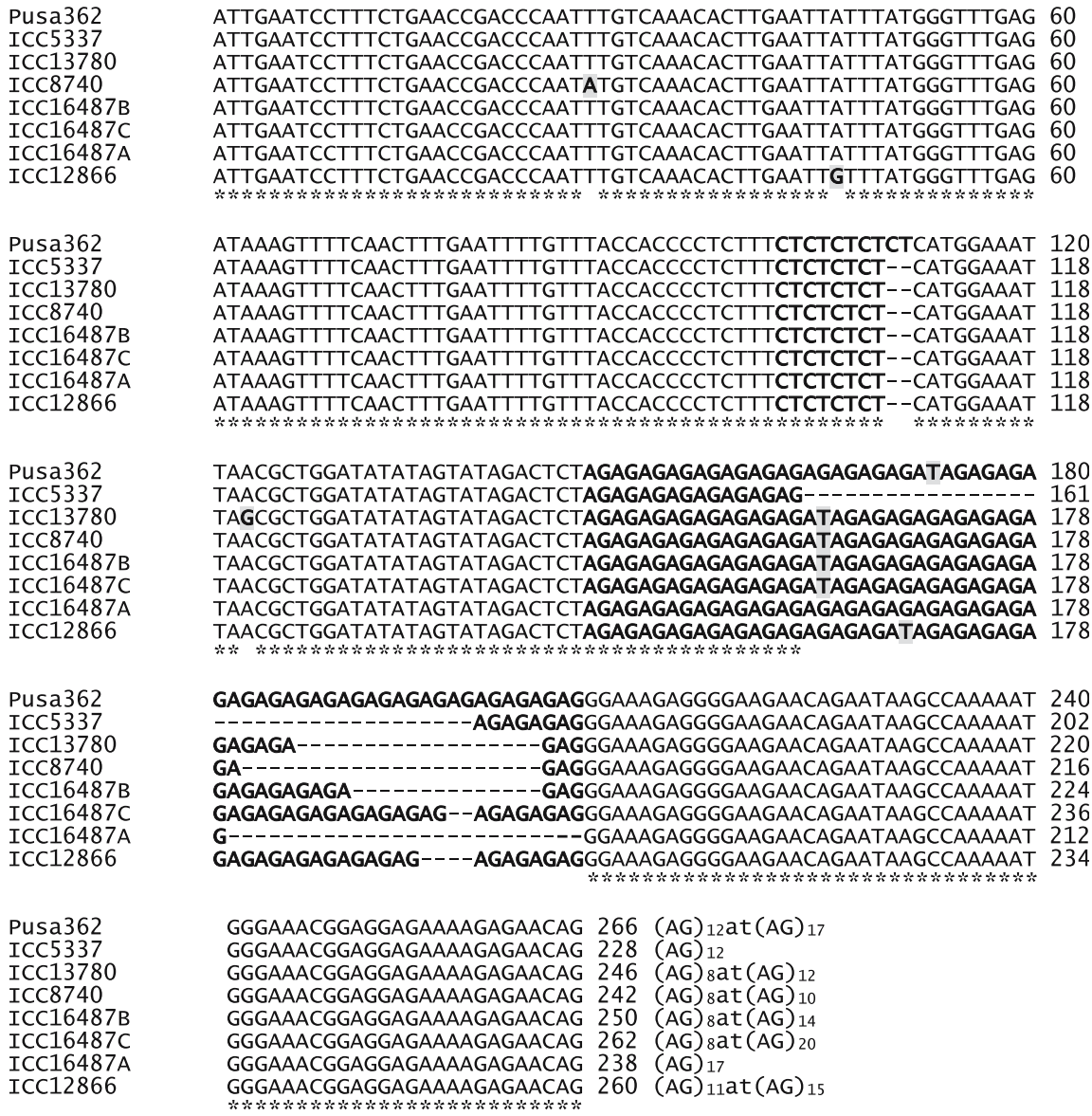

Fig. 3 Multiple sequence alignment of size-variant alleles from Pusa362 and other chickpea accessions (DQ192589-DQ192595) at NCPGR91 locus. The accession numbers along with variant alleles (A–C) are mentioned. Repetitive sequences are indicated in *bold*, *asterisks* mark conserved sequences and *dash* denote alignment gaps. The *shadowed boxes* indicate point mutations

maybe attributed to the fact that the present day chickpea has only one center of origin and a recent worldwide dispersal (Abbo et al. 2003) thereby leading to genetically similar backgrounds. Moreover, the genetic variability due to repeat number variation at microsatellite loci leading to low bootstrap support has been reported earlier by Udupa et al. (1999). To depict the patterns of variation we have further cloned and sequenced the variant alleles at two microsatellite loci.

Sequence variability and evolution of microsatellite loci

The length variant alleles at two microsatellite loci, NCPGR81 and NCPGR91, were cloned and sequenced in order to elucidate the mechanisms underlying the intra-specific microsatellite variation within and across accessions. It was clearly demonstrated that the basis of all allelic length variation was primarily due to the variation in the copy number of the repeat motifs whereas isolated point mutations within the repetitive regions or in the MFR (microsatellite flanking region) maybe responsible for sequence polymorphism. Large variation in the copy number of repeats was observed in the alleles at the two chickpea loci (Figs. 2, 3). Similar results demonstrating exceptional levels of length polymorphism at microsatellite loci have been reported in Arabidopsis (Symonds and Llyod 2003), soybean (Peakall et al. 1998), rice (Cho et al. 2000), maize (Matsuoka et al. 2002) and chickpea (Udupa and Baum

1999, 2001). In the chickpea NCPGR81 locus, all the longer repeat motifs containing more than 32 repeat units were punctuated by G → A conversion at position 124 demonstrating that the accumulation of point mutations in the longer (GA) motifs added imperfections within the long repeat stretches. Similar G → T conversions interrupting the long perfect repeat motifs were also observed at the NCPGR91 locus in accessions Pusa362, ICC8740, ICC12866 and ICC16487 B and C alleles (Fig. 3). It has been demonstrated earlier in humans, barn swallows and yeast that there exists a mutational bias towards an increase in repeat length of microsatellite loci at long perfect repeat tracts (Amos et al. 1996; Primmer et al. 1998; Wierdl et al. 1997). On the other hand, it has also been demonstrated that mutations accumulate more frequently in long perfect repeats punctuating them with imperfections (Zhu et al. 2000) thereby limiting the infinite expansion of microsatellite repeats. This maybe an important mechanism but necessary in the life cycle of microsatellites, which is essentially a balance between expansion by slippage and degradation by introduction of imperfections (Kruglyak et al. 1998). The point mutations within the repeat motif tend to decrease the slippage frequency (Richards and Sutherland 1994; Jin et al. 1996; Zhu et al. 2000) by adding imperfections to the long perfect repeat, which is the basis of microsatellite life cycle (Taylor et al. 1999; Zhu et al. 2000) and has been termed as "slippage/point-mutation" theory (Sibly et al. 2003). Studies of microsatellite evolution in human (Xu et al. 2000), Drosophila (Harr and Schlötterer 2000) and bacteria (Metzgar et al. 2002) have shown that repeat contraction also contribute in bringing about a shortening of the microsatellite alleles. Repeat contractions may also occur in plant microsatellites, however, no such evidence is as yet available in plants where only repeat disruptions have been implicated in stabilizing plant microsatellites (Symonds and Lloyd 2003).

Another interesting observation was the A → T point mutation immediately upstream of the repeat which resulted in the formation of a compound microsatellite $(TA)_3(GA)_{33}$ in the B allele of ICC7712 (Fig. 2). Random point mutations followed by the subsequent expansion by replication slippage are a well-documented mechanism for the birth of new microsatellites in animals (Messier et al. 1996; Schlötterer 2000). It has been shown in primates that a single G → A point mutation in a $(GT)_n$ microsatellite motif resulted in the formation and expansion of a (ATGT) repeat in the hominoids while a transition of A → G resulted in the formation of a (GT) repeat in Owl monkey (Messier et al. 1996). Such information in plants is only limited to a study in rice where a simple A → T base substitution mutation in a poly-(A) stretch of a microsatellite motif in wild rice, *Oryza longistaminta*, gave rise to a (TAA) microsatellite motif in the cultivated *O. sativa* (Akagi et al. 1998). Hence we may speculate that in chickpea also, point mutations upstream/downstream of perfect long microsatellites and subsequent expansion can give rise to new compound microsatellites. This process may also be helpful in bypassing the accumulation of point mutations/substitutions targeted at long microsatellite motifs for repeat interruption. Sequencing of this allele and its variants from diverse chickpea accessions and its wild relatives will depict this evolutionary process more accurately.

In conclusion, a set of 74 CT/GA and CA/GT microsatellite markers were identified and characterized from the chickpea nuclear genome. These microsatellites were found to be polymorphic within chickpea as well as between chickpea and *C. reticulatum*, thereby increasing the marker availability and applicability in genotyping and mapping projects. The sequence analysis at two microsatellite loci revealed copy number changes at repeats motifs as the major reason for microsatellite variability in chickpea. The dynamic role of point mutations in microsatellite evolution was also depicted.

Acknowledgements Financial assistance for this work was provided by Department of Biotechnology, Government of India (BT/PR3518/AGR/02/179/2002) and National Centre for Plant Genome Research. The seed materials provided by ICRISAT, and IARI, India are duly acknowledged. The authors acknowledge the support and encouragement provided by Prof. P.K.Gupta, C.C.S University, Meerut, India, for initiating this work. NK and BS are recipients of Senior Research Fellow (SRF) scholarships from Council for Scientific and Industrial Research (CSIR), Govt. of India.

References

Abbo S, Shtienberg D, Lichtenzveig J, Lev-Yadun S, Gopher A (2003) The chickpea, summer cropping, and a new model for pulse domestication in the ancient near east. Q Rev Biol 78:435–448

Ahmad F, Gaur PM, Slinkard AE (1992) Isozyme polymorphism and phylogenetic interpretations in the genus *Cicer* L. Theor Appl Genet 83:620–627

Akagi H, Yokozaki Y, Inagaki A, Fujimura T (1998) Origin and evolution of twin microsatellites in the genus *Oryza*. Heredity 81:187–197

Amos W, Sawcer SJ, Feakes W, Rubinszstein DC (1996) Microsatellite show mutational bias and heterozygote instability. Nat Genet 13:390–391

Billotte N, Lagoda PJL, Risterucci AM, Baurens FC (1999) Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. Fruits 54:277–288

Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, McCouch SR, Park WD, Ayres N, Cartinhour S (2000) Diversity of microsatellite derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100:713–722

Choudhary S, Sethy NK, Shokeen B, Bhatia S (2006) Development of sequence-tagged microsatellite site markers for chickpea (*Cicer arietinum* L.). Mol Ecol Notes 6:93–95

Chowdhury MA, Vandenberg V, Warkentin T (2002) Cultivar identification and genetic relationship among selected breeding lines and cultivars in chickpea (*Cicer arietinum* L.). Euphytica 127:317–325

Croser JS, Ahmad F, Clarke HJ, Siddique KHM (2003) Utilisation of wild *Cicer* in chickpea improvement—progress, constraints and prospects. Aust J Agri Res 54:429–444

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Edwards KJ, Backer JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques 20:758–760

Ferguson ME, Burow MD, Schulze SR, Bramel PJ, Paterson AH, Kresovich S, Mitchell S (2004) Microsatellite identification and characterization in peanut. Theor Appl Genet 108:1064–1070

Flandez-Galvez H, Ford R, Pang EC, Taylor PW (2003) An intraspecific linkage map of the chickpea (*Cicer arietinum* L.) genome based on sequence tagged microsatellite site and resistance gene analog markers. Theor Appl Genet. 106:1447–1456

Gaitan-Solis E, Duque MC, Edwards KJ, Tohme (2002) Microsatellite repeats in common bean (*Phaseolus vulgaris*): isolation, characterization, and cross-species amplification in *Phaseolus* ssp. Crop Sci 42:2128–2136

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

Harr B, Schlötterer C (2000) Long microsatellite alleles in *Drosophila melanogaster* have a downward mutation bias and short persistence times, which cause their genome-wide underrepresentation. Genetics 155:1213–1220

He C, Poysa V, Yu K (2003) Development and characterization of simple sequence repeat (SSR) markers and their use in determining relationships among *Lycopersicon esculentum* cultivars. Theor Appl Genet 106:363–373

Huttel B, Winter P, Weising K, Choumane W, Weigand F, Kahl G (1999) Sequence-tagged microsatellite site markers for chickpea (*Cicer arietinum* L.). Genome 42:210–217

Iruela M, Rubio J, Cubero JI, Gil J, Millan T (2002) Phylogenetic analysis in the genus *Cicer* and cultivated chickpea using RAPD and ISSR markers. Theor Appl Genet 104:643–651

Jin L, Macaubas C, Hallmayer J, Kimura A, Mignot E (1996) Mutation rate varies among alleles at a microsatellite locus: phylogenetic evidence. Proc Natl Acad Sci USA 93:15285–15288

Kruglyak S, Durrett RT, Schug M, Aquadro CF (1998) Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. Proc Natl Acad Sci USA 95:10774–10778

Labdi M, Robertson LD, Singh KB, Charrier A (1996) Genetic diversity and phylogenetic relationships among the annual *Cicer* species as revealed by isozyme polymorphism. Euphytica 88:181–188

Lichtenzveig J, Scheuring C, Dodge J, Abbo S, Zhang HB (2005) Construction of BAC and BIBAC libraries and their applications for generation of SSR markers for genome analysis of chickpea, *Cicer arietinum* L. Theor Appl Genet 110:492–510

Lowe AJ, Moule C, Trick M, Edwards KJ (2004) Efficient large-scale development of microsatellites for marker and mapping applications in *Brassica* crop species. Theor Appl Genet 108:1103–1112

Maguire TL, Edwards KJ, Saenger P, Henry R (2000) Characterisation and analysis of microsatellite loci in a mangrove species, *Avicennia marina* (Forsk.) Vierh. (Avicenniaceae). Theor Appl Genet 101:279–285

Malmberg R, Messing J, Sussex I (1985) Molecular biology of plants: a laboratory course manual, Cold Spring Harbor Laboratory, Cold Spring Harbor

Matsuoka Y, Mitchell SE, Kresovich S, Goodman M, Doebley J (2002) Microsatellites in *Zea*—variability, patterns of mutations, and uses for evolutionary studies. Theor Appl Genet 104:436–450

Messier W, Li SH, Stewart CB (1996) The birth of microsatellites. Nature 381:483

Metzgar D, Liu L, Hansen C, Dybvig K, Willis C (2002) Domain-level differences in microsatellite distribution and content result from different relative rates of insertion and deletion mutations. Genome Res 12:408–413

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within genus and other legume Genera: implications of the transferability of SSRs in plants. Mol Biol Evol 15:1275–1287

Pfaff T, Kahl G (2003) Mapping of gene-specific markers on the genetic map of chickpea (*Cicer arietinum* L.). Mol Genet Genom 269:243–251

Primmer CR, Saino N, Moller AP, Ellegren H (1998) Unraveling the process of microsatellite evolution through analysis of germ line mutation in barn swallows, *Hirundo rustica*. Mol Biol Evol 15:1047–1054

Riaz S, Dangl GS, Edwards KJ, Meredith CP (2004) A microsatellite marker based framework linkage map of *Vitis vinifera* L. Theor Appl Genet 108:864–872

Richards RI, Sutherland GR (1994) Simple repeat DNA is not replicated simply. Nat Genet 6:114–116

Rohlf FJ (1998) NTSYS: numerical taxonomy and multivariate analysis system, Version 2.1. Applied Biostatistics, Setauket

Rozen S, Skaletsky HJ (1998) Primer 3. Code available at http://www.genome.wi.mit.edu/genome_software/other/primer3.html

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd Edn. Cold Spring Harbor Laboratory, New York

Sant VJ, Patankar AG, Sarode ND, Mhase LB, Sainani MN, Deshmukh RB, Ranjekar PK, Gupta VS (1999) Potential of DNA markers in detecting divergence and analysis in heterosis in Indian elite chickpea cultivars. Theor Appl Genet 98:1217–1225

Schlötterer C (2000) Evolutionary dynamics of microsatellite DNA. Chromosoma 109:365–371

Sethy NK, Shokeen B, Bhatia S (2003) Isolation and characterization of sequence tagged-microsatellite markers in chickpea (*Cicer arietinum* L.). Mol Ecol Notes 3:428–430

Sharma PC, Huttel B, Winter P, Kahl G, Gardner RC, Weising K. (1995) The potential of microsatellites for hybridization- and polymerase chain reaction-based DNA fingerprinting of chickpea (*Cicer arietinum* L.) and related species. Electrophoresis 16:1755–1761

Sibly RM, Meade A, Boxall N, Wilkinson MJ, Corne DW, Whittaker JC (2003) The structure of interrupted human AC microsatellites. Mol Biol Evol 20:453–459

Singh KB (1987) Chickpea breeding In: Saxena MC, Singh KB (eds) The Chickpea, CAB International, Wallingford pp. 127–162

Squirrell J, Hollingsworth PM, Woodhead M, Russell J, Lowe AJ, Gibby M, Powell W (2003) How much effort is required to isolate nuclear microsatellites from plants? Mol Ecol 12:1339–1348

Sudupak A, Akkaya S, Kence A (2002) Analysis of genetic relationships among perennial and annual *Cicer* species growing in Turkey using RAPD markers. Theor Appl Genet 105:1220–1228

Symonds VV, Lloyd M (2003). An analysis of microsatellite loci in *Arabidopsis thaliana*: Mutational dynamics and application. Genetics 165:1475–1488

Taylor JS, Durkin JMH, Breden F (1999) The death of a microsatellite: a phylogenetic perspective on microsatellite interruptions. Mol Evol Biol 16:567–572

Udupa SM, Baum M (2001) High mutation rate and mutational bias at $(TAA)_n$ microsatellite loci in chickpea (*Cicer arietinum* L.). Mol Gen Genom 265:1097–1103

Udupa SM, Robertson LD, Weigand F, Baum M, Kahl G. (1999) Allelic variation at $(TAA)_n$ microsatellite loci in a world collection of chickpea (*Cicer arietinum* L.) germplasm. Mol Gen Genet 261:354–363

Udupa SM, Sharma A, Sharma RP, Pai RA (1993) Narrow genetic variability in *Cicer arietinum* L. as revealed by RFLP analysis. J Plant Biochem Biotech 2:83–86

Weising K, Kaemmer D, Weigand F, Epplen JT, Kahl G (1992) Oligonucleotide fingerprinting reveals various probe-dependant levels of informativeness in chickpea (*Cicer arietinum*). Genome 35:436–442

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

Winter P, Kahl G (1995) Molecular marker technologies for plant improvement. World J Microbiol Biotechnol 11:438–448

Winter P, Pfaff T, Udupa SM, Huttel B, Sharma PC, Sahi S, Arreguin-Espinoza R, Weigand F, Muehlbauer FJ, Kahl G (1999) Characterization and mapping of sequence-tagged microsatellite sites in the chickpea (*Cicer arietinum* L.) genome. Mol Gen Genet 262:90–101

Winter P, Benko-Iseppon AM, Hüttel B, Ratnaparkhe M, Tullu A, Sonnante G, Pfaff T, Tekeoglu M, Santara D, Sant VJ, Rajesh PN, Kahl G, Muehlbauer FJ (2000) A linkage map of chickpea (*Cicer arietinum* L.) genome based on recombinant inbred lines from a *C. arietinum* × *C. reticulatum* cross: localization of resistance genes for *fusarium* wilt races 4 and 5. Theor Appl Genet 101:1155–1163

Xu X, Peng M, Fang Z, Xu XP (2000) The direction of microsatellite mutations is dependent upon allele length. Nature Genet 24:396–399

Yap IV, Nelson RJ (1996) WinBoot: a program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrograms. IRRI, Philippines

Yeh FC, Boyle TJB (1997) Population genetic analysis of codominant and dominant markers and quantitative traits. Belgian J Bot 129:157

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. Mol Ecol 11:1–16

Zhang L, Yuan D, Yu S, Li Z, CaoY, Miao Z, Qian H, Tang K (2004) Preference of simple sequence repeats in coding and noncoding regions of *Arabidopsis thaliana*. Bioinformatics 20:1081–1086

Zhu Y, Queller DC, Strassmann JE (2000) A phylogenetic perspective on sequence evolution in microsatellite loci. J Mol Evol 50:324–338