Theor Appl Genet (2009) 118:591–608
DOI 10.1007/s00122-008-0923-z

ORIGINAL PAPER

Development of chickpea EST-SSR markers and analysis of allelic variation across related species

Shalu Choudhary · Niroj Kumar Sethy · Bhumika Shokeen · Sabhyata Bhatia

Received: 4 January 2008 / Accepted: 24 October 2008 / Published online: 20 November 2008

Abstract Despite chickpea being the third important grain legume, there is a limited availability of genomic resources, especially of the expressed sequence tag (EST)-based markers. In this study, we generated 822 chickpea ESTs from immature seeds as well as exploited 1,309 ESTs from the chickpea database, thus utilizing a total of 2,131 EST sequences for development of functional EST-SSR markers. Two hundred and forty-six simple sequence repeat (SSR) motifs were identified from which 183 primer pairs were designed and 60 validated as functional markers. Genetic diversity analysis across 30 chickpea accessions revealed ten markers to be polymorphic producing a total of 29 alleles and an observed heterozygosity average of 0.16 thereby exhibiting low levels of intra-specific polymorphism. However, the markers exhibited high cross-species transferability ranging from 68.3 to 96.6% across the six annual *Cicer* species and from 29.4 to 61.7% across the seven legume genera. Sequence analysis of size variant amplicons from various species revealed that size polymorphism was due to multiple events such as copy number variation, point mutations and insertions/deletions in the microsatellite repeat as well as in the flanking regions. Interestingly, a wide prevalence of crossability-group-specific sequence variations were observed among *Cicer* species that were phylogenetically informative. The neighbor joining dendrogram clearly separated the chickpea cultivars from the wild *Cicer* and validated the proximity of *C. judaicum* with *C. pinnatifidum*. Hence, this study for the first time provides an insight into the distribution of SSRs in the chickpea transcribed regions and also demonstrates the development and utilization of genic-SSRs. In addition to proving their suitability for genetic diversity analysis, their high rates of transferability also proved their potential for comparative genomic studies and for following gene introgressions and evolution in wild species, which constitute the valuable secondary genepool in chickpea.

Communicated by C. Gebhardt.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-008-0923-z) contains supplementary material, which is available to authorized users.

S. Choudhary · B. Shokeen · S. Bhatia (✉)
National Institute of Plant Genome Research, Post Box Number 10531, Aruna Asaf Ali Marg, Jawaharlal Nehru University Campus, New Delhi 110067, India
e-mail: sabhyata_bhatia@nipgr.res.in; sabhyatab@yahoo.com

Present Address:
N. K. Sethy
Defence Institute of Physiology and Allied Sciences (DIPAS), Defence Research and Development Organization (DRDO), Timarpur, Delhi 110 054, India

Introduction

Extensive efforts at sequencing of expressed genomic regions obtained from tissues under different conditions and developmental stages have led to a large number of EST sequences being deposited in the public database for a number of model species as well as economically important plants. Besides providing an effective approach for gene discovery and transcript pattern characterization, these ESTs emerge as a cost-effective, valuable source for molecular marker generation. These easily accessible sequences provide the advantage of in silico analysis and broaden the field of comparative studies in species where limited or no sequence information is available.

Microsatellites or simple sequence repeats (SSRs) are 1–6 bp iterations of DNA sequences that were earlier known

to occur only in the non-coding regions. However, the occurrence of microsatellites in transcribed sequences is now well established and are commonly known as EST-SSRs or genic SSRs (Morgante et al. 2002; Li et al. 2002, 2004). These have been reported from a number of plant species such as *Oryza* (Cho et al. 2000), *Saccharum* (Cordeiro et al. 2001), *Triticum* (Gupta et al. 2003), *Hordeum* (Thiel et al. 2003), *Medicago* (Eujayl et al. 2004), *Coffea* (Poncet et al. 2006; Aggarwal et al. 2007), *Capsicum* (Yi et al. 2006) and *Citrus* (Chen et al. 2006). Similar to genomic SSRs, the EST-SSRs are useful for many applications in plant genetics and breeding such as molecular mapping, genetic diversity analysis and cross-transferability across related species and genera (Varshney et al. 2005a). Moreover, as a result of their association with coding sequences, they provide the possibility of direct gene tagging for QTL mapping of agronomically important traits. The EST-SSRs find higher levels of cross-species transferability than genomic microsatellite markers (Scott et al. 2000; Eujayl et al. 2004; Zhang et al. 2005) aiding in gene introgression programs, identification of conserved gene order across orthologous linkage groups (Varshney et al. 2005a, b), depiction of gene evolution associated with microsatellites and phylogenetic studies.

In chickpea (*Cicer arietinum* L.), the third most important grain legume crop, research efforts worldwide have led to identification and characterization of a number of microsatellite markers (Hüttel et al. 1999; Sethy et al. 2003; Lichtenzveig et al. 2005; Choudhary et al. 2006, Sethy et al. 2006a) and their utilization for genome mapping (Winter et al. 2000) and phylogenetic analysis of *Cicer* (Sethy et al. 2006a, b). More recently, EST sequences from chickpea have been reported (Boominathan et al. 2004; Romo et al. 2004; Buhariwalla et al. 2005; Coram and Pang 2005). Among these, only the study of Buhariwalla et al. 2005 investigated the use of ESTs as a source of genic markers. But even in this study, of the 106 EST markers developed by them, only 14 contained SSR motifs and these are the only chickpea EST-SSRs reported till date. Hence the need for large scale development of chickpea EST-SSRs was imminent. This would not only help in molecular mapping but would also be of significance in comparative genome analysis in legumes since a high degree of conservation among the genomes of cultivated species and model legumes has been revealed (Weeden et al. 1992; Choi et al. 2004). A recent study on transferability of both genomic and EST-SSR markers of *M. truncatula* to pea, chickpea and lentil revealed a high degree of cross-transferability (Gutierrez et al. 2005). However, to be fully effective, genomic information from one species must be transferred in both directions, i.e. from model species to cultivars and vice versa (Gepts et al. 2005). Thus, there was an urgent need to develop EST-SSR markers in chickpea and assess their transferability to the model as well as to other non-model important legumes and for tracking the introgression of genes from the wild or elite species of chickpea.

The present study was aimed at: (1) development and characterization of chickpea EST-SSRs, (2) assessing the utility of EST-SSRs for genetic diversity analysis, (3) evaluating the cross-transferability of chickpea EST-SSRs among the *Cicer* species and other legumes, and (4) establishing the molecular basis of variation in alleles from related species and genera.

Materials and methods

Plant material and DNA isolation

Chickpea and the wild annual species are classified into various crossability groups. This study includes members of Crossability group I (*C. arietinum, C. reticulatum* and *C. echinospermum*) and Crossability group II (*C. bijugum, C. judaicum and C. pinnatifidum*). Thirty accessions of cultivated chickpea (*C. arietinum*) were used for the analysis of genetic diversity within species (Table 1a). For inter-specific transferability studies, nine accessions belonging to the five wild annual *Cicer* species were used which included a single accession of *C. echinospermum* (ICC17159) and two accessions each of *C. reticulatum* (ICC17121, ICC17164), *C. bijugum* (ICC17125, ICC17122), *C. judaicum* (ICC17148, ICC17150) and *C. pinnatifidum* (ICC17126, ICC17200). For cross-genera studies across legumes, 28 accessions belonging to seven legume genera were used (Table 1b). All accessions used in this study were grown at the field site of NIPGR.

DNA was isolated from fresh, young leaf tissue of chickpea and legume accessions using the CTAB method (Doyle and Doyle 1987). Genomic DNA from the wild *Cicer* accessions was isolated using GenElute genomic DNA miniprep kit (SIGMA Aldrich). The quality and final concentration was estimated by agarose gel electrophoresis using known concentration of uncut λ DNA as a standard.

Construction of a cDNA library and identification of EST-SSRs

Total RNA was isolated from 20 DAF seeds of *C. arietinum* ICCV2; 0.8 g of frozen seeds was ground to a fine powder in liquid nitrogen and transferred to an Eppendorf tube containing 500 µl of extraction buffer (200 mM NaOAc pH −5.2, 1% SDS, 10 mM EDTA pH 8.0) and 500 µl of phenol. This was centrifuged at 14,000×*g* for 10 min at RT. The aqueous phase was separated and extracted twice with phenol:chloroform (1:1) followed by O/N precipitation at 4°C with 0.3 vol of 10 M LiCl for RNA precipitation. The RNA pellet was recovered by centrifugation at 10,000 rpm for 10 min at 4°C, and was washed twice with 2.5 M LiCl

Table 1 A: list of chickpea accessions alongwith their sources, B: list of legume accessions used in this study

S. no.	Acc. no./name	Source
A		
1	ICCV2[a]	India
2	JG62[a]	-do-
3	ICC10945[a]	-do-
4	ICC15406[a]	-do-
5	ICC283[a]	-do-
6	ICC12947[a]	-do-
7	ICC13124[a]	-do-
8	ICC791[a]	-do-
9	ICC5383[a]	-do-
10	ICC11378[a]	-do-
11	ICC5477[a]	-do-
12	ICC15802[a]	Syria
13	ICC156947[a]	-do-
14	ICC16976[a]	Portugal
15	ICC7676[a]	-do-
16	ICC16800[a]	-do-
17	ICC16761[a]	-do-
18	ICC12866[a]	Ethiopia
19	ICC12726[a]	-do-
20	ICC3485[a]	Jordan
21	ICC6293[a]	Italy
22	ICC3631[a]	Iran
23	ICC16487[a]	Pakistan
24	ICC8195[a]	-do-
25	ICC7272[a]	Algeria
26	ICC13780[a]	Spain
27	ICC8444[a]	Tunisia
28	ICC15518[a]	Morocco
29	ICC15407[a]	-do-
30	Pusa362[b]	India

S. no.	Acc. no.	Species	Common name
B			
31	NRC37[c]	*Glycine max*	Soybean
32	MAUS47[c]	-do-	Soybean
33	PRATAP[c]	-do-	Soybean
34	BRAGG[c]	-do-	Soybean
35	IC381277[d]	*Lens esculenta*	Lentil
36	IC334282[d]	-do-	Lentil
37	IC384444[d]	-do-	Lentil
38	IC383609[d]	-do-	Lentil
39	IC411188[d]	*Trifolium alexandrinum*	Berseem Clover
40	IC411189[d]	-do-	Berseem Clover
41	IC508311[d]	-do-	Berseem Clover
42	IC411183[d]	-do-	Berseem Clover
43	IC347150[d]	*Cajanus cajan*	Pigeonpea
44	IC339040[d]	-do-	Pigeonpea
45	IC337447[d]	-do-	Pigeonpea
46	IC396014[d]	-do-	Pigeonpea
47	IC342955[d]	*Phaseolus mungo*	Blackgram
48	IC328538[d]	-do-	Blackgram
49	IC397612[d]	-do-	Blackgram
50	IC362567[d]	-do-	Blackgram
51	IC279013[d]	*Pisum sativum*	Field pea
52	IC356344[d]	-do-	Field pea
53	RFP-19[c]	-do-	Field pea
54	RFP-18[c]	-do-	Field pea
55	SA27783[e]	*Medicago truncatula*	Barrel Medic
56	SA11959[e]	-do-	Barrel Medic
57	SA3235[e]	-do-	Barrel Medic
58	SA3780[e]	-do-	Barrel Medic

[a] International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

[b] Indian Agriculture Research Institute (IARI), India

[c] Maharana Pratap Agriculture University, India (MPAU)

[d] National Bureau of Plant Genetic Resources (NBPGR), India

[e] Australian Medicago Genetic Resource Centre, SARDI, Australia

and once with 70% ethanol. The pellet was air-dried and dissolved in DEPC-treated ddH_2O. One microgram of total RNA was used to construct the cDNA library using the CLONTECH Smart PCR-cDNA synthesis kit according to the manufacturer's protocol. Double-stranded cDNA was introduced into the pCR2.1-TOPO vector for sequencing using the TOPO TA cloning kit (Invitrogen).

Random 5′ sequencing of cDNAs was done using the BigDye Terminator technology (Applied Biosystems) in an ABI Prism 3700 automated DNA sequencer. After sequence trimming (removal of low quality sequences, vector regions and sequences <100 bp), the EST sequences were mined for microsatellites consisting of ≥5 dinucleotides and ≥4 trinucleotides repeats using the TROLL program (Castelo et al. 2002). To reduce redundancy, cluster analysis was performed on microsatellite containing sequences (EST-SSRs) using the CAP3 program (Huang and Madan 1999). The identified EST-SSR sequences were deposited in the GenBank to obtain the accession numbers (see Table 2). The putative function of the developed chickpea functional markers was found by the BLASTX tool of NCBI, assuming a threshold of <1e − 05.

Generation of EST-SSRs from database

The 1,309 chickpea EST sequences available in the NCBI nucleotide database (up to January 2007) were screened for

Table 2 Characteristics of the chickpea EST-SSR markers

S. no.	Locus name	Primer sequence (5′ → 3′)	Motif	T_m (°C)	Expected size (bp)	GenBank no.	Putative function
1	CESSRDB2	CGGGCAGGTATTGAATTGTAA/ GAAAGGTTTACAGCCGTTGG	$(CT)_{17}$	59.4	169	CD051322	No homology
2	CESSRDB3	TTATCACTTGTTATTGTCCTCTAAG/ AATTTATGGACCCCATGTAA	$(TAA)_6$	60.5	197	AJ609280	No homology
3	CESSRDB4	GAAGAGGTAGCGGAGGAG/ CAAGCAACAGTTTTCACTCA	$(GGT)_3CGT(GGT)_2N_6(GGT)_3$	61.4	274	AJ609279	RNA and export binding protein
4	CESSRDB5	CCGACATCTCTTCTCAATTC/ CTTTAGGTGGTGGTTGTTGT	$(TCA)_{14}$	61.4	177	AY370650	SAT5 gene
5	CESSRDB7	AAGTGGTGTCGGTAATGGT/ TAATACCAAAAGCATGCACA	$(GGT)_5N_6(GGT)_2$	60.2	196	AJ487469	Glycine-rich protein
6	CESSRDB10	CCCTTAATCAATTCA CCTCA/ TTATCCAAACCAATGATTCC	$(TAA)GAA(TAA)_6$	59.8	192	AJ005947	No homology
7	CESSRDB11	AATCTAACAGCAACGACGAT/ ATCAAGCTTCTTCTGCACAT	$(CCA)CAA(CCA)_3$	58.4	298	AJ006048	Unknown protein
8	CESSRDB13	ATCTGGGAGCTTGTGAGTTA/ TTGTATCTCCTTCAGATGGC	$(AT)_5$	60.0	260	AJ012683	Hypothetical protein
9	CESSRDB15	CTTACGATTTCTCCTCCCTT/ TTTCTCATACCGAATCCTTG	$(GCT)N_6(GCT)_4$	61.7	276	AJ012681	Hypothetical protein
10	CESSRDB16	ATGCTATGCATGATGTTTCA/ GTTCCAAACAAACACAACAA	$(TA)_{10}$, $(TTA)_4G(TTA)A(TTA)_2$,	57.0	295	AJ487472	Invertase inhibitor
11	CESSRDB18	TGCAAATAAAGCCTTCAAGT/ GAAAGTGGGAAAATGCAATA	$(TA)_5CA(TG)_2T(TG)_3$	57.4	242	AJ487042	No homology
12	CESSRDB21	GTGTATCGGTCAGGAAAAGA/ GGTACACACCACAATTCACA	$(ATT)_2AT\ (AAT)_3$	60.0	259	AJ012693	Plantacyanin
13	CESSRDB23	GTGTGGACCTGAAATTGAGT/ GAATATGGGAACAAGTGCAT	$(TA)_5$	59.4	221	AJ012689	Ribonuclease T2
14	CESSRDB24	TGTGCTTGACTTGTTCACAT/ TATGCATCCTCATTTTCTCC	$(GGC)_4$	59.6	283	AJ006763	β-amylase
15	CESSRDB26	GGTGCATTCTCTTCCATAAG/ TGCAAATCT TTAACCAAACA	$(GT)_5$	57.4	273	AJ004959	Expansin
16	CESSRDB27	GGTGAGATTAGGAAGCAATG/ TATCCAATCCCCATAAGATG	$(TAT)_7$	58.7	215	AJ271660	Cationic peroxidase
17	CESSRDB29	TTTAGTTGCACAACAACAGC/ AAATCCACATCCAAAAAGGT	$(TGA)_5$	57.4	176	AJ299064	GTP-binding protein

Table 2 continued

S. no.	Locus name	Primer sequence (5′ → 3′)	Motif	T_m (°C)	Expected size (bp)	GenBank no.	Putative function
18	CESSRDB33	GCTGCACAAAAAGTACATGA/ ATCCATCGAAACACCAATAG	(GA)T(GA)TT$(GA)_5$	59.3	234	AJ250836	Pal gene
19	CESSRDB34	AACCTAAAGCCGAAAAGAAG/ CTCCCGTGAAGTAATAGTCG	(AAG)$N_4$$(AAG)_4$	61.4	251	AJ400863	Histone H2B
20	CESSRDB35	TCTAGAGCTAGCCAAAGGAA/ GCATCGTAATCATCGGTACT	$(GAT)_7$	61.7	272	AJ400861	UDP-glycose
21	CESSRDB38	GAGTAAGATGGCACAGTGGT/ GTATCTATTAGCGAAGCGGA	$(CCG)_4$	60.4	197	X93220	Cysteine proteinase
22	CESSRDB39	CTGAGGTTAATGTGAAAGGC/ GTCAACATCACATGCTCAAC	$(GGT)_4$	61.0	257	AJ275314	Glycine-rich protein
23	CESSRDB40	GAAATTAGGAAGCATTGTGC/ AATTGATTGAACCCACTTGT	$(TTAT)_4$	57.0	188	AJ275313	Peroxidase
24	CESSRDB41	GAACCAATAAAGCCTTGAAA/ TGACCAATTGATACAATCCA	$(GCT)_4$, $(TTTA)_4$	57.4	247	AJ275307	PM intrinsic polypeptide
25	CESSRDB42	GAGACAAAGATAGTGGCTGG/ TATTAATCACTCGCACGACA	$(TAAAT)_4$ $(GTTT)_3$	61.3	235	AJ275304	ABA-responsive protein
26	CESSRDB44	ATCCTTTCCTTGTTGTGCTA/ TTTAGTGAAGCATTGTTGGA	$(CTTT)_3$, $(TTG)_4$	57.8	267	AJ012581	Cytochrome P450
27	CESSRDB45	AGATGGTTTGAATGTTGAGG/ CACTTGACCCTTTGATTGTT	$(AT)_7(AG)_5$	59.5	295	AJ249802	Cytochrome P450
28	CESSRDB47	ACGAAGAAAGTTCCTGTGAA/ ACCGAAAACCTGATTCATTA	$(TTA)_3TAA(TTA)_4$	57.6	240	AJ006767	Histone H1
29	CESSRDB51	ACTATTACAAGAGCCCACCC/ CATAATGGTAAGGAGGTGGA	$(CAA)_4$	62.4	297	AJ006770	Extensin
30	CESSRDB53	CCCTTAATCAATTCACCTCA/ GCTTCTTATCCAAACCAATG	$(TAA)_6$	59.5	197	AJ005947	No homology
31	CESSRDB54	AGTGTTGTGGGTTTCATTTC/ TTGATTTGCCAAAGTACACA	$(TTA)_5$	59.6	221	AJ005869	Trans-membrane channel protein
32	CESSRDB55	CGATTATCTCAACTTTTGGC/ ACATGCACACGACAAATAAA	$(TA)_5$, $(ACT)_5$	59.0	136	AJ005000	Transcriptional regulator
33	CESSRDB56	TGTCTGGAACAACAAGTGAG/ GCCAATCAGATTTCCTCTTA	$(ATG)_4$	58.4	247	CK149113	Myb family transcriptional factor
34	CESSRDB61	GCAGAATGGGAGATAATGAA/ TGCTGATTCTGATGTCTACG	$(CTT)_7$	60.0	233	CK149116	bZIP transcriptional factor

Table 2 continued

S. no.	Locus name	Primer sequence (5′ → 3′)	Motif	T_m (°C)	Expected size (bp)	GenBank no.	Putative function
35	CESSR14	GGCACAAGGTATCTCCACAA/ ATGCTTGCCTCAACCTCAGA	$(TGC)_6$	56.8	300	ES544474	Unknown protein
36	CESSR15	CATGACATCCTCAATCCTTGG/ TAGCGACAAATCTTAGCCGTAG	$(TGC)_3AGC(TGC)_2$	59.4	300	ES544475	Unknown protein
37	CESSR20	CGA AACTCGAACGTGCAAT/ TTTGGCGAATTTGAAAGGAG	$(GATTC)_6$, $(ATTTA)_5$	58.5	386	ES544478	Unknown protein
38	CESSR21	CCTCAACGCTCATTCTTCTTCT/ CCCCAAGGAACCATTCTAAGAT	$(CTT)_6$	60.5	233	ES544479	OSH1 related protein
39	CESSR23	CGCGTAAACGTTATTCTCTTCA/ CATCATTTCCCTTAGCATCCTT	$(TTC)_7$, $(CTT)_3$	58.5	399	EX151810	Glutathione peroxidase
40	CESSR25	CTATGGCAAAAAGCATCACAAG/ ATAGCCATGGCCACATTAAACT	$(CCG)_6$	60.3	363	ES544482	No homology
41	CESSR26	GGCAAAATCGAAATTCAACC/ TGATCAATGACAGTGTAGAAGG	$(CTT)_4$, $(CT)_{10}$	59.3	275	ES544483	No homology
42	CESSR30	TCGGACCACAAGAGCATCTA/ CGTGGAAGAAAGGAATGTTG	$(CT)_6TT(CT)_2TT(CT)_3$, $(CAC)_6$	57.8	388	ES544486	Actin-binding potein
43	CESSR31	ACGTAGGTTAAGGTTGCTGGTC/ TTCAACGTGTTCGAAAGCTC	$(AAG)_8T(AAG)$	58.0	113	ES544487	Unknown protein
44	CESSR34	CATTGCTCAAAGCCAATTCA/ TCGATGAATCGGAACAAACA	$(TAT)_4$, $(TGT)_6$	56.9	294	ES544488	No homology
45	CESSR42	TGGTTGAAGAAAAGAAGGTAGTG/ CGGTTCACTAATGCAAAAACCT	$(ACC)_5$	59.5	298	ES544489	Hsr203J homolog protein
46	CESSR43	CATTAAAGCTAGGAGTTTGTGCTG/ ACGGTACCATACCCGACTACAT	$(CTA)_4$	55.6	386	EX567535	Bimodular protein
47	CESSR47	GAGTTCCACATTGTCACAGGAA/ AATGCAACAGTCCTTGTGGATA	$(TTC)_5$	57.4	541	EX567643	Germin like protein
48	CESSR51	CACATGAACAGAAAAAGGGACA/ GCATGTTGAGCCAAAGCTAAAT	$(TTTG)_5$	58.9	205	EX567864	Pleckstrin domain related
49	CEESR61	CACTCTTCCCTCCCTTTCTTTA/ GAATCAGGGTAGGTTTGTTTGC	$(CT)_7T(CT)_2$	60.0	257	EX151660	Armadillo like helical protein
50	CESSR62	ACCAGCTGCTAGACCTGATGTT/ GCAATAAAACAAAATCCTCACACC	$(TGA)_5$, $(TAT)_3$	62.5	245	EX567512	RHG1protein
51	CESSR65	CTCCTCCACTCATCT TCATCTTC GAGAAGGTGTTTCCGGTAAAAGT	$(CTT)_2CTC(CTT)_2GG(CTT)$	58.9	352	EX567847	Hypothetical protein

Table 2 continued

S. no.	Locus name	Primer sequence (5′ → 3′)	Motif	T_m (°C)	Expected size (bp)	GenBank no.	Putative function
52	CESSR68	AATGGCCACCATTTTCTCATC/ AAACGTTCTTTCCATCCTTCTG	$(ACC)_6$	57.8	330	EX151762	Hypothetical protein
53	CESSR71	TTGTAGTTCTCCTCTCTCTCTCTC/ CATCAAAACCAAACCTATGGAG	$(CT)C(CT)_{8,}$ $(CT)_6$, $(CT)_{11}$	59.2	295	EX567905	Unknown protein
54	CESSR72	ATTTCACTCCTCACTTCTCACC/ CACGAAAATCGGATGATTCAG	$(CT)_7$	58.5	345	EX151914	Unknown protein
55	CESSR73	TCTTCTCCCATTCGTTGTTGAT/ GATCTTCTGTTCCTCAGCCAAC	$(CTT)_3ATT(CTT)_6(GTT)_2(AT)_4$	57.0	363	EX151922	No homology
56	CESSR77	CCAACTTAAACTCATTTCGTCTCA/ CCAAGATGTGTTTTTGATGATG	$(GA)_2AA(GA)_4$, $(CAT)_4$	56.2	173	EX567970	DUF647 protein
57	CESSR78	ATTGCTGAGGCTGTGAATTGTA/ CCCAATACATCAAAGATAGATCG	$(TTAAA)_3$	55.0	373	EX567577	Lipid transfer protein
58	CESSR80	TCACCCTTTCTTCTTCAACTTC/ GAACGCATAAAATAGTCGCTGA	$(GA)_7$	61.2	260	EX151949	No homology
59	CESSR85	ATGTACTTGGTCTGGTCCGTCT/ ACCTTTTCGGCGTTTCTTTTAC	(GCT)4	62.5	279	EX151920	Hypothetical Protein
60	CESSR93	ACGAGAAGAGCATTGCATTTG/ TAACGGCCTCTTTAGTCTGCTC	$(AAC)_3AAG$	56.4	353	EX151839	Pex19 related protein

The designed primer pairs, microsatellite repeat motifs, annealing temperature, expected allele size (bp), Genbank accession numbers and their putative functions based on BLASTX results are mentioned. Serial nos 1–34 represent markers designed using the database whereas serial nos 35–60 represent markers designed based on the in-house generated ESTs

the presence of microsatellite motifs using the TROLL program (Castelo et al. 2002) and the same criteria as mentioned above was used.

Designing EST-SSR primers

Primers were designed using the Primer3.0 software (Rozen and Skaletsky 1997) and designated as CESSR (for in-house generated ESTs) and CESSRDB (for EST-SSRs generated using the database). The parameters used for primer design were: (1) primer length 18–24 bp with an optimum of 20 bp, (2) annealing temperature 50–60°C with an optimum of 55°C, (3) percentage GC in the range of 40–50, and (4) product size in the range of 100–400 bp. All the oligonucleotides were synthesized from Illumina Inc. (USA).

Amplification and detection of microsatellite alleles

All PCR amplifications of genomic DNA (including *Cicer* and legume species) were carried out in a 15 µl reaction volume in a BIORAD thermal cycler (Icycler) containing 40–50 ng of genomic DNA, PCR buffer (20 mM Tris–HCl, 50 mM KCl), 0.75 µM of each primer, 0.125 mM of each dNTP, 1.5 mM $MgCl_2$ and 0.5 U of *Taq* DNA polymerase (Life Technology, India). The following touchdown amplification profile was used: (1) initial denaturation 94°C 3 min, (2) 18 cycles of 94°C 50 s, 65°C 50 s, decreasing annealing temperature 0.5°C/cycle, 72°C 50 s, (3) 20 cycles of 94°C 50 s, 55°C 50 s, 72°C 50 s, and (4) final extension 72°C 7 min. The amplification products were separated on 6% polyacrylamide gels or 3% Metaphor agarose gels (Cambrex, USA) depending upon the size range, stained with ethidium bromide and analyzed using the gel documentation system AlphaImager 2200 (Alpha Innotech Corp., USA). Fragment sizes for each locus were evaluated using standard size markers. All *Cicer* species were scored in a binary matrix and analyzed using POPGENE version 1.32 (Yeh and Boyle 1997) and the UPGMA based dendrogram was constructed using NTSYS-pc Version 2.1 (Rohlf 1994).

Sequence analysis of amplified fragments

Size variant alleles from different *Cicer* and legume accessions were amplified and resolved on 6% PAGE gels. The bands were eluted, cloned into pGEM-T Vector (Promega) and transformed to DH5α cells. After blue-white selection, plasmids from putative recombinants were isolated using the alkaline lysis method (Sambrook et al. 1989). Sequencing reactions were performed as above. Four to six recombinants from each allele were sequenced and the sequences were aligned with the original chickpea sequence using CLUSTAL W (1.83).

Results

Development of functional EST-SSR markers

ESTs from two sources were used for development of the EST-SSR markers in this study. First, using the 1,309 EST sequences reported in the NCBI database till January 2007, representing approximately 0.76 Mb, 133 microsatellite motifs were identified. Second, a cDNA library from immature seed was constructed and used as a source of EST-SSRs. From the 822 seed ESTs generated in-house, 159 EST-SSRs (19%) were identified that clustered into a total of 99 consensus sequences possessing a total of 113 microsatellite motifs. Sequence analysis of the 246 SSR motifs from the two sources (133 + 113) revealed that 207 (84.1%) were perfect repeats, 29 (11.7%) were imperfect and ten (4.0%) were compound. The copy number of the dinucleotide repeat motifs at the perfect loci varied from 5 to 17 and the trinucleotide motifs from 4 to 14. A diverse range of SSR motifs was present which varied widely with trinucleotide repeats (51.2%) being the most abundant followed by di- (37.3%), tetra- (6.9%) and pentanucleotide (4.4%) motifs. The most frequently occurring dinucleotide motifs were GA followed by TA and GT, whereas among trinucleotides AAG was predominant followed by ATT.

Using the flanking regions of 246 SSR motifs, 183 primers were designed of which 94 could be validated in chickpea acc. Pusa362 and ICCV2. Of these, 34 primers either did not amplify or produced anomalous-sized fragments. Therefore, finally 60 functional EST-SSR markers were developed (described in Table 2) of which 49 primer pairs produced single expected sized alleles, whereas 11 primers amplified 2–4 alleles. Based on the BLASTX analysis, putative functions could be assigned to majority (65.0%) of the EST-SSRs that showed significant homology to reported proteins, whereas 18.4% of them represented unknown/hypothetical proteins and 16.6% to novel sequences (Table 2).

Intra-specific diversity within chickpea revealed by EST-SSR markers and sequence analysis

To elucidate intra-specific variability, 60 chickpea EST-SSR primers (Table 2) were used to amplify the genomic DNA from 30 chickpea accessions listed in Table 1a. Of these 60 markers, ten produced polymorphic amplification profiles in the 30 accessions, amplifying a total of 29 alleles with a maximum of five alleles with the primer pair CESSR73 in the chickpea cultivars (Table 3). The observed heterozygosity

Table 3 Number of alleles (N_a), size range of amplified fragments, observed (H_O) and expected heterozygosity (H_E), Shannon's informative index (I) and fixation index (F_{IS}) values calculated for 30 chickpea accessions at ten polymorphic EST-SSR loci

Locus	N_a	Size range (bp)	H_O	H_E	I	F_{IS}
CESSR23	2	399–402	0.00	0.44	0.63	0.08
CESSR42	3	295–303	0.00	0.59	0.94	1.00
CESSR43	2	386–390	0.00	0.43	0.62	1.00
CESSR47	4	539–650	0.60	0.66	1.20	1.00
CESSR61	3	254–262	0.00	0.57	0.92	1.00
CESSR62	3	243–295	0.46	0.66	1.06	0.28
CESSR71	2	295–301	0.00	0.44	0.62	1.00
CESSR72	3	342–348	0.00	0.59	0.95	1.00
CESSR73	5	359–445	0.60	0.76	1.47	0.19
CESSR77	2	173–176	0.00	0.50	0.68	1.00
Average	2.9	–	0.16	0.56	0.91	–
SD	0.99	–	0.2711	0.1110	0.28	–

values ranged from 0 to 0.6 and expected heterozygosity ranged from 0.43 to 0.76 with an average of 0.56 (Table 3).

To assess the basis of length variation across chickpea, homologous alleles amplified by primer pair CESSSR73 were cloned and sequenced (Fig. 1). Sequence comparisons revealed the presence of a $(CTT)_n$ compound repeat motif. Allelic size variations were mainly due to the presence of an additional repeat motif (CTT) in some of the cultivars accompanied by single base insertions/deletions in the MFR regions. Cultivar specific insertions were observed at position 301 in ICCV2 and deletions at 96, 102 in ICC15406 and 286 in ICC7676. In addition, isolated point mutations were also observed in the MFRs such as at position 195 in ICC7676 (Fig. 1). However, when similar sized alleles from monomorphic loci such as CESSRDB13, CESSRDB27, and CESSRDB44 were sequenced, even though point mutations were observed at these loci, no variation in the copy number of repeats was found (data not shown).

Cross-species transferability across Cicer

In order to assess the transferability rates of the chickpea EST-SSR markers across related *Cicer* species, genomic DNA from nine accessions belonging to five wild, annual *Cicer* species (listed in "Materials and methods") were amplified with the same 60 EST-SSR primers used for intra-specific analysis. The transferability rates of chickpea EST-SSRs varied from a high of 96.6% in *C. reticulatum* to a low of 68.3% in *C. judaicum* with an average of 82.6% (Table 4). Forty-one markers (68.3%) amplified in all the annual species, of which, 27 were polymorphic across the wild *Cicer* species. Allelic data generated using the 60 EST-SSRs revealed a minimum of one and maximum of nine alleles (CESSRDB47) with a total of 156 alleles at 60 loci leading to an average of 2.6 alleles per locus. Observed heterozygosity (H_o) ranged 0.15–0.83 with an average of 0.22 (Table 4). Of the 60 primer pairs, two (CESSRDB3 and CESSRDB5) amplified only in chickpea (and in no other wild species) and were therefore specific to *C. arietinum*. The EST-SSR markers also exhibited crossability group-specific transferability. Among the first crossability group members (*C. arietinum*, *C. reticulatum*, and *C. echinospermum*), 55 markers (91.6%) were amplified and 24 markers produced polymorphic amplification profiles. Similarly, 41 markers (58.8%) successfully amplified in all the second crossability group members (*C. bijugum*, *C. judaicum*, and *C. pinnatifidum*) and 23 primers detected variation between the three species. Five markers (CESSRDB7, CESSRDB23, CESSRDB24, CESSRDB41, and CESSR26) exhibited crossability-group-specific transferability to only first crossability group members.

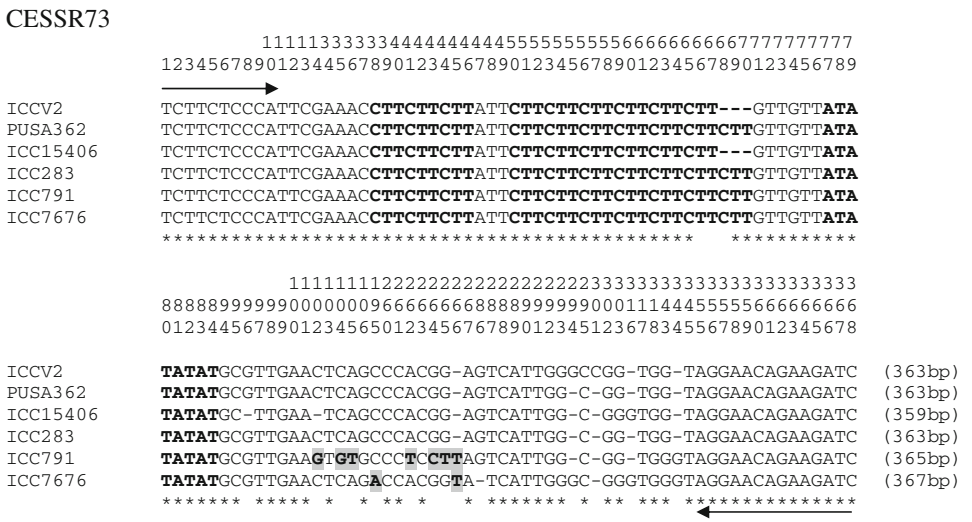

Fig. 1 Partial sequence alignment of size variant alleles amplified using primer pair CESSR73 across six chickpea accessions (EX151922, EU332161-EU332163, EU332165, EU332166). The *asterisks* represent similar sequences and *dash* indicates alignment gaps. Repeat regions are in *boldface*, primer-binding sites are represented by *arrows* and characters in *bold shaded boxes* indicate point mutations

Table 4 Number of alleles and their sizes (bp) obtained in six annual species of *Cicer* at 60 genic microsatellite loci are mentioned

Locus	*C. arietinum* No. of alleles (size in bp)	*C. reticulatum* No. of alleles (size in bp)	*C. echinospermum* No. of alleles (size in bp)	*C. bijugum* No. of alleles (size in bp)	*C. judaicum* No. of alleles (size in bp)	*C. pinnatifidum* No. of alleles (size in bp)	Transferability of each marker (%)	H_o	H_e
CESSRDB2	2 (137, 169)	1(169)	1 (169)	3 (129, 169, 222)	1 (129)	1 (129)	100	0.3333	0.7424
CESSRDB3	1 (197)	–	–	–	–	–	0		
CESSRDB4	1 (274)	1 (274)	1 (274)	1 (262)	1 (262)	1 (262)	100	0.0000	0.5455
CESSRDB5	1 (177)	–	–	–	–	–	0		
CESSRDB7	1 (196)	1 (196)	1 (196)	–	–	–	40	0.0000	0.0000
CESSRDB10	1 (192)	1 (192)	1 (192)	1 (179)	1 (182)	1 (173)	100	0.0000	0.7273
CESSRDB11	1 (298)	1 (298)	1 (298)	2 (298, 507)	–	–	60	0.2500	0.2500
CESSRDB13	1 (260)	1 (260)	2 (260, 334)	2 (227, 334)	3 (277, 334 343)	3 (277, 334, 343)	100	0.6667	0.7576
CESSRDB15	2 (261, 276)	2 (261, 276)	2 (261, 276)	1 (276)	1 (276)	1 (276)	100	0.5000	0.4091
CESSRDB16	2 (295, 346)	2 (323, 374)	2 (323, 374)	2 (323, 374)	–	1 (295)	80	0.8000	0.8000
CESSRDB18	1 (242)	3 (242, 279 315)	3 (252, 308 315)	3 (252, 300 324)	2 (252, 291)	2 (252, 291)	100	0.8333	0.8485
CESSRDB21	3 (259, 284, 296)	2 (259, 296)	3 (259, 284, 296)	1 (259)	2 (270, 296)	2 (270, 296)	100	0.8333	0.6818
CESSRDB23	1 (221)	1 (221)	1 (221)	–	–	–	40	0.0000	0.0000
CESSRDB24	1 (283)	1 (283)	1 (283)	–	–	–	40	0.0000	0.0000
CESSRDB26	1 (273)	1 (273)	1 (273)	1 (298)	1 (291)	1 (291)	100	0.0000	0.6667
CESSRDB27	1 (215)	1 (215)	1 (215)	1 (209)	1 (205)	1 (205)	100	0.0000	0.6667
CESSRDB29	1 (176)	1 (176)	1 (176)	1 (176)	1 (176)	1 (176)	100	0.0000	0.0000
CESSRDB33	1 (234)	1 (234)	1 (234)	1 (234)	1 (234)	1 (234)	100	0.0000	0.0000
CESSRDB34	1 (251)	1 (251)	1 (251)	2 (235, 251)	1 (251)	1 (251)	100	0.1667	0.1667
CESSRDB35	1 (272)	1 (272)	1 (272)	1 (272)	1 (272)	1 (272)	100	0.0000	0.0000
CESSRDB38	2 (197, 234)	2 (197, 234)	2 (197, 234)	2 (197, 234)	–	–	60	0.0000	0.0000
CESSRDB39	1 (257)	1 (257)	1 (257)	1 (257)	1 (257)	1 (257)	100	0.0000	0.0000
CESSRDB40	1 (188)	3 (188, 255 266)	2 (212, 272)	3 (199, 255 266)	–	3 (208, 266 272)	80	0.8000	0.9111
CESSRDB41	3 (247, 256, 278)	3 (251, 260, 281)	2 (251, 260)	–	–	–	40	1.0000	0.8667
CESSRDB42	2 (235, 280)	1 (235)	1 (235)	1 (235)	1 (235)	1 (235)	100	0.1667	0.1667
CESSRDB44	1 (267)	1 (267)	–	–	–	–	20	0.0000	0.0000
CESSRDB45	3 (295, 426, 449)	3 (295, 426, 449)	3 (295, 426, 449)	3 (295, 426, 449)	–	–	60	1.0000	0.5714
CESSRDB47	4 (240, 246 253. 276)	3 (240, 253 276)	4 (240, 246 253. 276)	4 (235, 248 258, 281)	3 (248, 268 281)	3 (248, 268 281)	100	1.0000	0.8182
CESSRDB51	1 (297)	1 (297)	1 (297)	1 (297)	1 (297)	1 (297)	100	0.0000	0.0000
CESSRDB53	1 (197)	1 (197)	1 (197)	1 (190)	1 (186)	1 (186)	100	0.0000	0.6667
CESSRDB54	2 (221, 305)	2 (221, 305)	2 (221, 305)	2 (221, 305)	–	–	60	1.0000	0.5714

Table 4 continued

Locus	*C. arietinum* No. of alleles (size in bp)	*C. reticulatum* No. of alleles (size in bp)	*C. echinospermum* No. of alleles (size in bp)	*C. bijugum* No. of alleles (size in bp)	*C. judaicum* No. of alleles (size in bp)	*C. pinnatifidum* No. of alleles (size in bp)	Transferability of each marker (%)	H_o	H_e
CESSRDB55	2 (136, 154)	2 (136, 154)	2 (143, 154)	2 (140, 168)	2 (140, 168)	2 (140, 155)	100	0.1515	0.7778
CESSRDB56	1 (247)	1 (247)	1 (247)	1 (247)	1 (247)	1 (247)	100	1.0000	0.8485
CESSRDB61	1 (233)	1 (233)	1 (233)	1 (233)	–	–	60	0.0000	0.000
CESSR14	1 (300)	1 (300)	1 (300)	1 (300)	1 (300)	1 (300)	100	0.0000	0.0000
CESSR15	1 (300)	1 (300)	1 (300)	1 (300)	1 (292)	1 (292)	100	0.0000	0.4848
CESSR20	1 (386)	1 (386)	1 (386)	–	–	1 (390)	60	0.0000	0.5333
CESSR21	1 (233)	1 (233)	3 (233, 300, 324)	3 (233, 300, 324)	1 (245)	1 (245)	100	0.3333	0.6667
CESSR23	1 (399)	1 (391)	1 (391)	1 (395)	1 (391)	1 (395)	100	0.0000	0.6667
CESSR25	1 (363)	1 (363)	1 (363)	1 (363)	1 (363)	1 (363)	100	0.0000	0.0000
CESSR26	1 (275)	1 (275)	1 (275)	–	–	–	40	0.0000	0.0000
CESSR30	1(388)	1(388)	1(388)	1(388)	1(388)	1(388)	100	0.0000	0.0000
CESSR31	1 (113)	1 (113)	2 (113, 108)	1 (113)	1 (113)	1 (113)	100	0.1667	0.16607
CESSR34	1 (294)	1 (287)	2 (294, 280)	1 (294)	2 (294, 280)	1 (294)	100	0.3333	0.4545
CESSR42	1 (298)	1 (305)	1 (298)	1 (305)	1 (305)	1 (298)	100	0.0000	0.5455
CESSR43	1 (386)	1 (395)	1 (395)	1 (395)	1 (395)	1 (395)	100	0.0000	0.3030
CESSR47	1 (541)	1 (541)	1 (552)	1 (552)	–	–	60	0.0000	0.5714
CESSR51	1 (205)	1 (209)	1 (205)	1 (209)	1 (205)	1 (209)	100	0.0000	0.5455
CESSR61	1 (257)	1 (257)	1 (257)	1 (262)	1 (257)	1 (252)	100	0.0000	0.5455
CESSR62	1 (245)	1(249)	1 (249)	1 (254)	1 (254)	1 (260)	100	0.0000	0.7879
CESSR65	1 (352)	1 (352)	1 (370)	1 (366)	1 (366)	1(366)	100	0.0000	0.6667
CESSR68	3 (330, 342, 350)	3 (330, 342, 350)	3 (322, 342, 350))	2 (322, 342)	2 (322, 342)	1 (322)	100	0.8333	0.6898
CESSR71	1 (295)	1 (286)	–	–	–	–	20	0.0000	0.6667
CESSR72	1 (345)	1 (348)	1 (345)	1 (350)	1 (350)	1 (350)	100	0.0000	0.6667
CESSR 73	1 (363)	3 (363, 370, 375)	–	1 (370)	–	–	40	0.3333	0.7333
CESSR 77	1 (173)	1 (170)	1 (170)	1 (170)	1 (170)	1 (170)	100	0.0000	0.3030
CESSR 78	1 (373)	2 (373, 500)	2 (366, 500)	2 (366, 500)	2 (370, 500)	2 (370, 500)	100	0.8333	0.1970
CESSR 80	1 (260)	1 (260)	1 (260)	1 (260)	1 (260)	1 (260)	100	0.0000	0.0000
CESSR85	1 (279)	1 (279)	1 (274)	1 (274)	1 (279)	1 (274)	100	0.0000	0.5455
CESSR93	1 (353)	1 (353)	1 (353)	1 (353)	1 (353)	1 (353)	100	0.0000	0.0000
Transferable markers (%)	100	96.6	91.6	83.3	68.3	73.3		Avg – 0.22	Avg – 0.35

Observed (H_o) and expected (H_e) heterozygosity values were estimated. Hyphen represents no amplification

Sequence analysis of size variant alleles from *Cicer* species

To investigate the basis of variation among size variant alleles of six *Cicer* species, fragments amplified in various annual species at the five loci, CESSRDB4, CESSRDB10, CESSRDB26, CESSRDB27 and CESSRDB34, were cloned and sequenced. Multiple alignments of nucleotide sequences from each locus were done and the results for three loci (CESSRDB4, CESSRDB10 and CESSRDB26) are shown in Fig. 2. Sequence comparisons at loci revealed that even though there was overall sequence conservation in the internal microsatellite structure and the primer-binding sites, variations such as differences in the copy number of repeat motifs and repeat interruptions accompanied by indels and point mutations in the microsatellite flanking regions (MFR) frequently occurred, all of which contributed to the allelic length variation (Fig. 2; Table 4). For example, at the CESSRDB4 locus, an addition of three repeat motifs (48–56 bp) accompanied by mutations at positions 26, 44 and 47 resulted in motifs specific for the first and second crossability group members (Fig. 2a). Similarly at the CESSRDB10 locus, a mutation (T → G) at nucleotide position 50 and a TAG repeat expansion among members of first crossability group were observed (Fig. 2b). At the CESSRDB26 locus the second crossability group alleles were much longer due to insertions in the MFR regions (Fig. 2c).

Another interesting feature revealed by sequence comparisons was the crossability-group-specific point mutations and indels. Point mutations (nucleotide positions 26, 44, 47 and 184 in CESSRDB4; positions 34, 45, 50, 115, 121, 143 and 164 in CESSRDB10; positions 92, 145, 193, 205, 210, 219, 223, 224, 234, 252, 266, 267 and 275 in CESSRDB26) and indels [12 bp (48–59) in CESSRDB4], [9 bp (68–76) and 4 bp (106–109) in CESSRDB10] and [14 bp (95–108), 1 bp (161) and (5 bp (255–259) in CESSRDB26] were highly crossability group specific. Additionally in the flanking regions, there were species-specific point mutations for example at position 103 in *C. arietinum* and 198 in *C. judaicum* at CESSRDB4 locus (Fig. 2).

Phylogenetic analysis

The allelic data obtained from the 60 chickpea EST-SSR markers were used to visualize the genetic relationships among the 30 chickpea accessions and the six annual *Cicer* species. After scoring and computing the allelic data, a dendrogram was constructed that clearly separated the members of the first and second crossability groups into clusters I and II (Fig. 3). Cluster I corresponded with the first crossability group members grouping all chickpea accessions into Cluster IA and the *C. reticulatum* and *C. echinospermum* into Cluster IB. The ClusterIA clearly distinguished all the chickpea accessions except ICC15518 and ICC8195; however, no correlation between the clustering pattern and geographical location was obtained. Cluster II represented the second crossability group species with *C. judaicum* and *C. pinnatifidum* being closely placed together.

Cross-genera transferability and sequence variation of chickpea EST-SSRs across legumes

Thirty-four of the chickpea functional markers were also utilized to assess their cross-genera transferability across 32 accessions spanning eight legume genera (ESM S1). This analysis revealed varied levels of marker transferability across legumes ranging from 29.4% in *P. mungo*, 35.2% in *P. sativum*, 41.1% in *G. max* and *T. alexandrinum*, 47.0% in *L. esculenta*, 50.0% in *C. cajan and* 61.7% in *M. truncatula* with an average of 43.6%. Eight markers (23.5%) amplified in all the legume genera though all the accessions of each species were not amplified. Twenty-five markers (74%) amplified in at least one legume species other than chickpea, whereas nine primers (26%) amplified only in the chickpea accessions and no other legume indicating the uniqueness of these loci to the chickpea genome. Twelve markers produced polymorphic amplification profiles across legumes even though intra-specific polymorphism was not observed.

Different sized alleles amplified at various loci across legumes were cloned and sequenced. For example, sequence analysis at CESSRDB56 (Fig. 4) demonstrated that although the same-sized alleles had high sequence conservation, variable alleles such as the 228 bp allele in *Trifolium* revealed polymorphism that was due to both differences of repeat motifs as well as variability of the flanking sequences marked by indels/point mutations. A similar observation was also noted at locus CESSRDB39 amplifying multiple alleles (data not shown).

Discussion

This study reports for the first time, development of a large number of EST-SSR markers in chickpea and assesses their transferability across a wide-spectrum of related species and genera, thereby establishing that the chickpea EST-SSR markers are a valuable genetic resource for investigating species relationships and comparative mapping in legumes. The easiest way to develop genic markers is by screening of EST sequences for the presence of hypervariable SSR motifs. In the publicly available EST database of legumes (Fabaceae) nearly 1 million EST

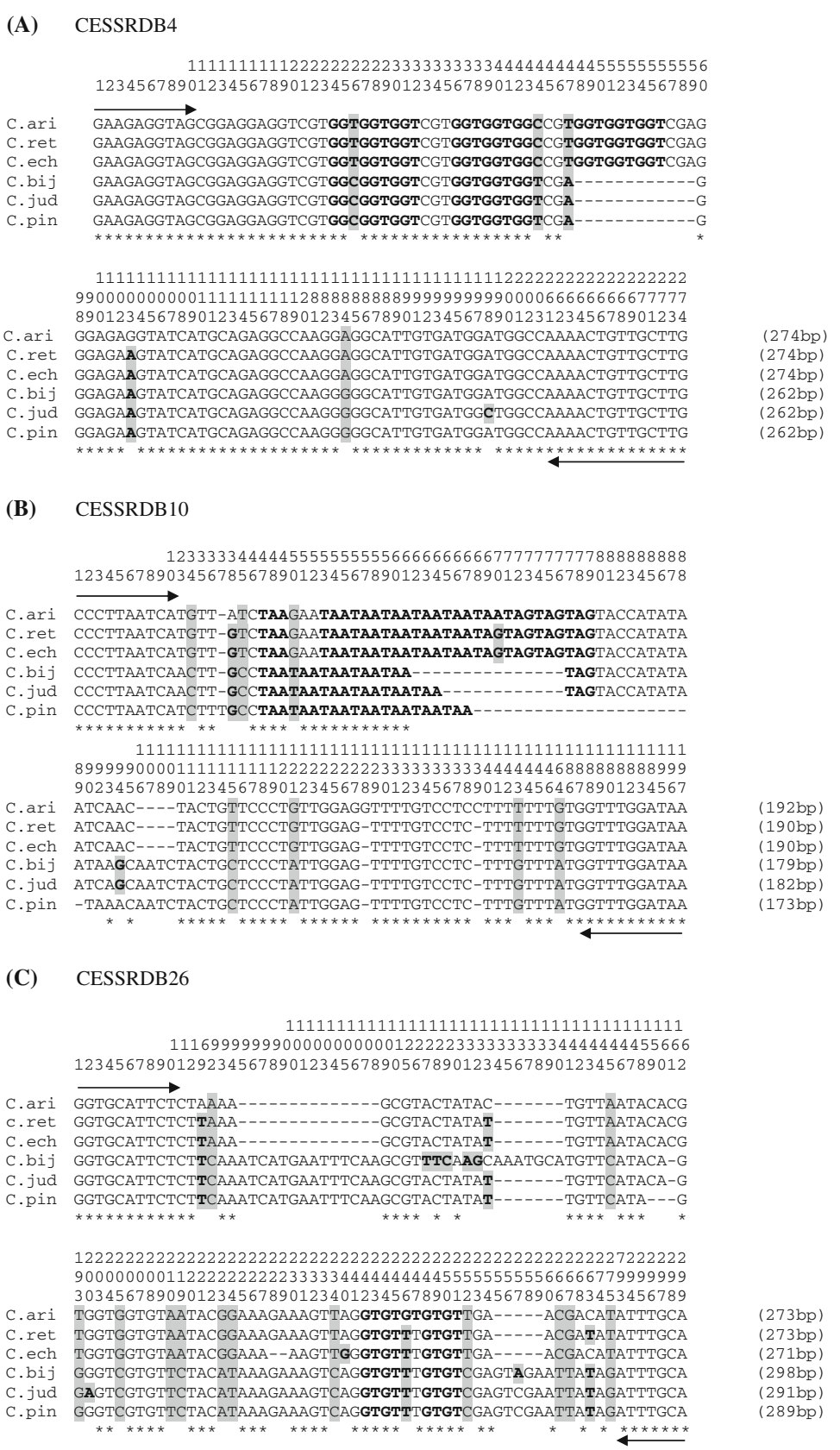

Fig. 2 Partial sequence alignment of size variant alleles amplified using (**a**) CESSRDB4 primer across six annual *Cicer* species (EF595573-EF595577) (**b**) CESSRDB10 primer across six annual *Cicer* species (EF595578-EF595582) and (**c**) CESSRDB26 primer across six annual *Cicer* species (EF595583-EF595587). The repeat motifs are represented as *bold letters*, *arrows* represent primer-binding sites, − Indicates alignment gaps and * represents similar sequences. Characters in *bold shaded boxes* indicate base substitutions and group specific mutations are demarcated by grey background. C.ari (*C. arietinum)*, C.ret (*C. reticulatum)*, C.ech (*C. echinospermum)*, C.bij (*C. bijugum)*, C.jud (*C. judaicum*), C.pin (*C. pinnatifidum*)

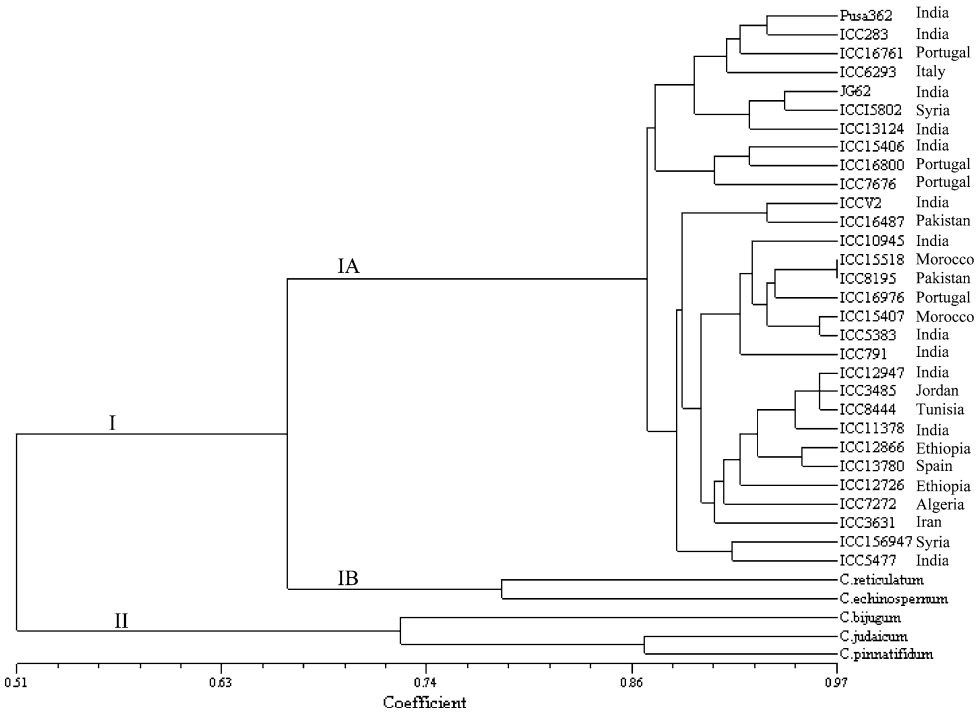

Fig. 3 UPGMA based dendrogram of thirty chickpea accessions and five annual *Cicer* species was obtained using 60 functional EST-SSR markers and Jaccard's coefficient. Name of cultivars, species and source country are mentioned

CESSRDB56

```
C.ari      TGTCTGGAACAACAAGTGAGTATAATTTGAATGGACAA---TCTGAGTGTTCTGAACAGA
M.tru      TGTCTGGAACAACAAGTGAGTATAATTTGAATGGACAA---TCTGAGTGTTCTGAACAGA
T.ale      TGTCTGGAACAACAAGTGAGTATAATTTGAATGGACAACAATCTGAGTGTTCTGAACAGA
           ************************************   ******************

C.ari      CAACATCAACAGAAGGACCTGTTTTGGAAGAGTTTATTCCAATTAAGAAAAGGGCTTCAC
M.tru      CAACATCAACAGAAGGACCTGTTTTGGAAGAGTTTATTCCAATTAAGAAAAGGGCTTCAC
T.ale      CAACATCAACAGATGGACCTGTTTTGGAAGAGTTTATTCCAATTAAGAAAAGGGCTTCAC
           ************ ********************************************

C.ari      CTTATTGTGAACAAGTGTATGATGATGATGA----AAAGAAGATGATGAACAGC-TTCTC
M.tru      CTTATTGTGAACAAGTGTATGATGATGATGA----AAAGAAGATGATGAACAGC-TTCTC
T.ale      CTTATTGTGAACAAGTTTATGATGATGAGGATGGTGATGAAGATGATGAACAACATTCTC
           *************** *********** **     * ************** * *****

C.ari      ATCATAAGCAGCAAAAGATTTCATCAAATGATAATAATAAGAACAGTGATAAGAGGAAAT
M.tru      ATCATAAGCAGCAAAAGATTTCATCAAATGATAATAATAAGAACAGTGATAAGAGGAAAT
T.ale      ATCATAAACAACAAAAGATTTCA----ATG----------------GATAAGAGGAAAT
           ******* ** ************    ***                ************

C.ari      CTGATTGGC        (247bp)
M.tru      CTGATTGGC        (247bp)
T.ale      CTGATTGGC        (228bp)
           *********
```

Fig. 4 Multiple sequence alignment of alleles amplified from chickpea and two legumes namely *M. truncatula* (M.tru) and *T. alexandrinum* (T.ale) at CESSRDB56 locus (EF621420 and EF595632). These primers also amplified alleles from legumes which were of the same size as chickpea and also shared high sequence homology, hence are not shown. Only the size variants are shown. The *asterisks* represent similar sequences, *dash* indicates alignment gaps. Repeat regions are in *boldface*, primer-binding sites are represented by *underlined letters* and characters in *bold shaded* boxes indicate point mutations

sequences are available, of which over 92% represent ESTs derived from *M. truncatula*, *L. japonicus* and *G. max* (Ramírez et al. 2005). For chickpea, only about 1,300 ESTs were publicly available (upto January 2007). Hence, our study utilized this resource for developing EST-SSR markers. Moreover, since only a limited number of ESTs were available, our study also undertook to generating new EST sequences and using them for the development of a novel set of functional markers. These markers will not only be a significant addition to the limited set of SSR markers available in chickpea, but will have the added advantage of marker-trait associations.

From the publicly available database and in-house ESTs, a total of 246 SSRs were identified which represented 11.5% of the screened ESTs. This SSR frequency was comparable with those obtained in citrus (10.6%, Chen et al. 2006), pepper (10.7%, Yi et al. 2006) and in other dicot species (Kumpatia and Mukhopadhyay 2005).

However, Kantety et al. 2002 obtained comparatively lower frequency of EST-SSRs ranging from 1.5 to 4.7% in monocots. The abundance of SSRs mined from a sequence database depends on the SSR search criteria, the size of the dataset and the database mining tools (Varshney et al. 2005a). On applying stringent SSR criteria with a minimum of 20 bp, about 5% of ESTs have been shown to contain SSRs in plants (Varshney et al. 2005a) whereas the same when applied to this set of chickpea EST sequences, only 3.1% sequences contained SSRs. The abundance of trinucleotide motifs in the chickpea coding sequences (51.2%) was in close agreement with observations in monocot and dicot plants (Kantety et al. 2002; Tian et al. 2004; Yi et al. 2006) establishing the need of the coding regions to maintain the reading frame (Varshney et al. 2002; Li et al. 2004). The predominance of GA motifs among dinucleotides in the chickpea ESTs was similar to reports in cereals (Varshney et al. 2002) and dicots like *Medicago*, soybean and *Arabidopsis* (Tian et al. 2004). Similarly among trinucleotides, the abundance of AAG motifs in chickpea was quite consistent with the findings of Li et al. (2004) and Kumpatia and Mukhopadhyay (2005). However, earlier studies on chickpea microsatellites have reported the $(TAA)_n$ motif to be most abundant (Udupa et al. 1999). Moreover, it was observed that even though the EST-microsatellites contained less number of repeat motifs than the genomic microsatellites (gSSRs) reported earlier (Sethy et al. 2006a), they proved to be highly informative in the genetic diversity and cross-species transferability studies (Scott et al. 2000; Thiel et al. 2003).

Chickpea has been shown to exhibit overall low levels of polymorphism with the various molecular markers analyzed so far and this has been attributed to its self-pollinating nature as well its recent worldwide dispersal (Udupa et al. 1999). In this study also, the EST-SSR markers displayed a low level of polymorphism (16%) within chickpea accessions in comparison to earlier reports of 40–60% polymorphism detected by gSSRs (Sethy et al. 2006a; Lichtenzveig et al. 2005). This observation is noteworthy as SSRs located in the coding regions are under strong selection pressure and therefore accumulate few mutations (Varshney et al. 2005a; Li et al. 2004). However, despite the lower polymorphism, the genic-SSRs are preferable over gSSRs as these are associated with the coding regions of the genome and therefore represent "true genetic diversity" that would directly assist in "perfect" marker-trait associations (Eujayl et al. 2002; Thiel et al. 2003). Other species such as rice (Cho et al. 2000), sugarcane (Cordeiro et al. 2001) and wheat (Gupta et al. 2003), have also revealed similarly low levels of polymorphism using EST-SSRs compared to genomic SSR markers.

In contrast to the low levels of intra-specific polymorphism with EST-SSRs, the inter-specific polymorphism was significantly higher in the wild relatives of chickpea. Wild genepool is extremely valuable in inter-specific hybridization programs since they serve as sources of resistance/tolerance to many stresses. Our study with EST-SSR markers will potentially facilitate the transfer of traits of agronomic value into cultivated chickpea thereby leading to the broadening of the narrow genetic base and development of superior genotypes of chickpea. The dendrogram obtained with the EST-SSR markers clearly showed the closeness of *C. judaicum* with *C. pinnatifidum* (Fig. 3) which was in agreement with the earlier protein based (Tayyar and Waines 1996) and EST-based studies (Buhariwalla et al. 2005), whereas using DNA-based marker systems like AFLP (Shan et al. 2005) and STMS markers (Sethy et al. 2006b) the closeness of *C. pinnatifidum* with *C. bijugum* has been reported. The resemblances of the dendrograms based on protein markers with genic markers suggest that coding sequences of *C. judaicum* and *C. pinnatifidum* may have followed a common evolutionary pathway.

The chickpea EST-SSRs developed in this study revealed much higher rates of transferability (mean 82.6%) across wild annuals than the chickpea-derived gSSRs (68%; Choumane et al. 2000). Higher inter-specific transferability was in accordance with other studies (Scott et al. 2000; Eujayl et al. 2004; Zhang et al. 2005), establishing that functional markers were more transferable and therefore more useful than gSSR markers with the added potential of being used in allele-mining for identification of useful agronomic traits. It was also observed that the mean transferability rates across the primary and secondary crossability groups were an average of 96.0 and 74.9%, respectively. This difference could be explained on the basis of an earlier study of Decroocq et al. (2003), which said that the level of sequence conservation of microsatellite loci is inversely proportional to the genetic distance. Similar observations have been made in other species such as wheat (McLauchlan et al. 2001) and sugarcane (Cordeiro et al. 2001) where the genic markers displayed low level of polymorphism in cultivated accessions compared to other members of the genus thereby directing the breeders to look into the related species for introgression of novel genetic material into the germplasm.

Results from the sequencing data also provided evidence for limited sequence variability within the chickpea alleles in comparison to much higher levels of variation across the orthologous alleles from annual species. Sequence comparisons of size variant microsatellite alleles within chickpea accessions illustrated approx. 95% overall sequence conservation with few indels in the repeat as well as the MFR region (Fig. 1) suggesting the presence of evolutionary constraints within transcribed regions that limit the mutational events and increase sequence similarity. However, sequence comparison of microsatellite

alleles from various *Cicer* species revealed a wide range of length and sequence variability both in terms of band size and allele number. Similar results have also been obtained by Buhariwalla et al. 2005 thereby establishing that EST based microsatellite markers of chickpea were not only efficient for marker-assisted introgression programs using wild germplasm but also reliable for synteny studies within the genus *Cicer*. Sequence variations occurred both at the repeat motifs and in the flanking regions and were interestingly found to be crossability-group-specific and therefore highly phylogenetically informative that could help in understanding the evolution of microsatellites in a phylogenetic context since it has been shown that such events at the genic loci might play an important role in speciation or gene functionality diversification during the evolutionary process.

Earlier studies have provided evidence, which shows that microsatellites undergo expansion during the course of evolution (Zhu et al. 2000; Peakall et al. 1998). In our study, there was an expansion of the GGT motif at the locus CESSRDB4 (Fig. 2a), resulting in the presence of three additional repeats in members of the first crossability group. At the locus CESSRDB10, expansion of the TAA motif was accompanied by the birth of a new TAG motif (via a A → G mutation) which later expanded in the members of the first crossability group. Such A → G transition was also observed by Messier et al. (1996) in owl monkey. It has been speculated that base substitution allows the birth of new motifs that subsequently expand by replication slippage (Gordon 1997). Recently, the role of microsatellite expansion/deletion in terms of gene regulation is being investigated well in mammals as well as in plants (Li et al. 2004). The presence of SNPs in the sequence of similar sized alleles from different chickpea cultivars apparently indicated the limitation of scoring the accessions simply based on the amplicon size on gel. Also, this clearly highlights the prospects of SNP mapping in chickpea as these represent the most fundamental source of variation for molecular marker development.

In our study it was observed that the microsatellite motifs were long but punctuated by imperfections which are most often regarded as an effective mechanism for prevention of infinite growth of microsatellites (Kruglyak et al. 1998; Palsboll et al. 1999). At the CESSRDB4 and CESSRDB10 loci, base substitutions at positions 44 and 50, respectively, in the first crossability group members implied that such interruptions may have a dramatic impact in the long-term evolution of the microsatellite sequence. On the other hand, the phenomenon of microsatellite purification (loss of interruptions), a mechanism counteracting the accumulation of imperfections is also known to occur (Harr et al. 2000). This was observed at locus CESSRDB26 (Fig. 2c) where all species, except *C. arietinum*, harbor "T" at position 246 indicating that T represents the ancestral character state. The occurrence of longer motifs in the focal species in comparison to the related species may also be explained by the hypothesis of ascertainment bias (Ellegren et al. 1997; Peakall et al. 1998; Vigouroux et al. 2002) which in our study was demonstrated by sequence comparisons at the CESSRDB4 and CESSRDB10 loci (Fig. 2a, b). The cross-transferability of chickpea EST-SSR markers across legume species was high (mean 43.6%) clearly depicting the conservation of primer-binding sites in genomic DNA over a long evolutionary period. The usefulness of EST-SSR markers over genomic SSRs for transferability across distant relatives has been established in species such as *Medicago* (Gutierrez et al. 2005), wheat (Gupta et al. 2003), barley (Thiel et al. 2003) and grapes (Scott et al. 2000). However in chickpea, no extensive study of cross-genera transferability of genomic SSRs was available, except for a small study by Pandian et al. (2000), the transferability rates across distant species of genomic versus EST-SSRs could not be compared. Our study showed that the highest rate of transferability of the chickpea EST-SSR markers was to *Medicago* (61.7%), whereas an earlier study by Gutierrez et al. (2005) showed significant, yet lower levels of transferability of the *Medicago* markers to chickpea (36.3%). The difference in the rates of transferability could be attributed to the choice of loci and the overall number of markers analyzed. Our study also demonstrated that the rate of transferability decreases from within the genus *Cicer* (82.6%) to outside the genus (43.6%) which was in agreement with earlier reports in cereals (Thiel et al. 2003; Gupta et al. 2003), grapes and apricot (Decroocq et al. 2003) and *Medicago* (Gutierrez et al. 2005) suggesting that amplification decreases with increasing evolutionary distance from focal species. Overall, the chickpea markers transferred very efficiently to some members of the galegoid legumes (such as *Medicago* and *Trifolium*) as compared to the phaseoloid legumes (such as *P. mungo*). However, *Pisum* and *Cajanus* were exceptions to this. The variable marker transferability rate obtained across different legume genera indicated the occurrence of genus-specific evolutionary events.

In conclusion, our study was the first attempt at characterization of a large number of SSRs from the coding regions of the chickpea genome. This study not only contributed to strengthening the chickpea EST database but also provided the first set of functional SSR markers for evaluating the chickpea germplasm and molecular mapping. In this study, it was established that the chickpea EST-SSRs were highly transferable across a number of distantly related species thereby providing ample opportunity for mining of superior alleles and development of candidate gene markers for use in gene introgression programs and comparative genomics in legumes. Further, our study also provided the molecular evidence for understanding the basis of allelic variation

within and across species, which demonstrated the presence of complex mutational processes, highlighting the evolution of microsatellites in a phylogenetic context within the genus *Cicer*.

Acknowledgments We are grateful to the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) India, National Bureau for Plant Genetic Resources (NBPGR) India, Maharana Pratap Agriculture University (MPAU) India, and Australian Medicago Genetic Resource Centre, SARDI, Australia for providing us the seed material of different accessions of cultivated chickpea, wild *Cicer* species and other legumes. Financial assistance for this work was provided by National Institute for Plant Genome Research (NIPGR), New Delhi, India and also the Department of Biotechnology (DBT), Government of India by means of a project grant. The fellowship provided to SC by University Grants Commission (UGC), India and to NK and BS by Council for Scientific and Research (CSIR), India is gratefully acknowledged.

References

Aggarwal RK, Hendre PS, Varshney RK, Bhat PR, Krishnakumar V, Singh L (2007) Identification, characterization and utilization of EST-derived genic markers for genome analyses of coffee and related species. Theor Appl Genet 114:359–372

Boominathan P, Shukla R, Kumar A, Manna D, Negi D, Verma PK, Chattopadhyay D (2004) Long term transcript accumulation during the development of dehydration adaptation in *Cicer arietinum*. Plant Physiol 135:1608–1620

Buhariwalla HK, Jayashree B, Eshwar K, Crouch JH (2005) Development of ESTs from chickpea roots and their use in diversity analysis of the *Cicer* genus. BMC Plant Biol 5:16

Castelo AT, Martius W, Gao GR (2002) TROLL—Tandem repeat occurrence locator. Bioinformatics 18:634–636

Chen C, Zhou P, Choi YA, Huang S, Gmitter FG Jr (2006) Mining and characterizing microsatellites from citrus ESTs. Theor Appl Genet 112:1248–1257

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

Choi HK, Mun JH, Kim DJ, Zhu H, Baek JM, Mudge J, Roe B, Ellis N, Doyle J, Kiss GB, Young ND, Cook DR (2004) Estimating genome conservation between crop and model legume species. Proc Natl Acad Sci USA 101:15289–15294

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

Choumane W, Winter P, Weigand F, Kahl G (2000) Conservation and variability of sequence-tagged microsatellite sites (STMSs) from chickpea (*Cicer arietinum* L.) within the genus *Cicer*. Theor Appl Genet 101:269–278

Coram T, Pang E (2005) Isolation and analysis of candidate ascochyta blight defense genes in chickpea, Part I. Generation and analysis of an expressed sequence tag (EST) library. Physiol Mol Plant Pathol 66:192–200

Cordeiro GM, Casu R, McIntyr CL, Manners JM, Henry RJ (2001) Microsatellite markers from sugarcane (*Saccharum* spp.) ESTs cross transferable to erianthus and sorghum. Plant Sci 160:1115–1123

Decroocq V, Favé MG, Hagen L, Bordenave L, Decroocq S (2003) Development and transferability of apricot and grape EST microsatellite markers across taxa. Theor Appl Genet 106:912–922

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

Ellegren H, Moore S, Robinson N, Byrne K, Ward W, Sheldon BC (1997) Microsatellite evolution—a reciprocal study of repeat lengths at homologous loci in cattle and sheep. Mol Bio Evol 14:854–860

Eujayl I, Sorrells ME, Baum M, Wolters P, Powell W (2002) Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theor Appl Genet 104:399–407

Eujayl I, Sledge MK, Wang L, May GD, Chekhovskiy K, Zwonitzer JC, Mian MAR (2004) *Medicago truncatula* EST-SSRs reveal cross species genetic markers for *Medicago* spp. Theor Appl Genet 108:414–422

Gepts P, Beavis WD, Brummer EC, Shoemaker RC, Stalker HT, Weeden NF, Young ND (2005) Legumes as a model plant family. Genomics for food and feed report of the cross-legume advances through Genomics Conference. Plant Physiol 137:1228–1235

Gordon AJE (1997) Microsatellite birth register. J Mol Evol 45:337–338

Gupta PK, Rustgi S, Sharma S, Singh R, Kumar N, Balyan HS (2003) Transferable EST-SSRs markers for the study of polymorphism and genetic diversity in bread wheat. Mol Gen Genet 270:315–323

Gutierrez MV, Vaz Patto MC, Huguet T, Cubero JI, Moreno MT, Torres AM (2005) Cross-species amplification of *Medicago truncatula* microsatellites across three major pulse crops. Theor Appl Genet 110:1210–1217

Harr B, Zangerl B, Schlötterer C (2000) Removal of microsatellite interruptions by DNA replication slippage: phylogenetic evidence from *Drosophila*. Mol Bio Evol 17:1001–1009

Huang X, Madan A (1999) CAP3: a DNA sequence assembly program. Genome Res 9:868–877

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

Kantety RV, Rota ML, Matthews DE, Sorrells ME (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, sorghum and wheat. Plant Mol Biol 48:501–510

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

Kumpatia SP, Mukhopadhyay S (2005) Miming and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species. Genome 48:985–998

Li Y-C, Korol AB, Fahima T, Beiles AV, Eviatar Nevo (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol Eco 11:2453–2465

Li YC, Korol AB, Fahima T, Nevo E (2004) Microsatellites within genes: structure, function, and evolution. Mol Biol Evol 21:991–1007

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

McLauchlan A, Henry RJ, Isaac PG, Edwards KJ (2001) Microsatellites analysis in cultivated wheat and wild relatives. In: Henry RJ (ed) Plant genotyping-the DNA fingerprinting of plants. CABI publishing, Wallingford

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

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat Genet 30:194–200

Palsboll PJ, Berube M, Jorgensen H (1999) Multiple levels of single-strand slippage at cetacean tri-and tetra nucleotide repeat microsatellite loci. Genetics 151:285–296

Pandian A, Ford R, Taylor PWJ (2000) Transferability of sequenced tagged microsatellite site (STMS) primers across four major pulses. PMB Reporter 18:395a–395h

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

Poncet V, Rondeau M, Tranchant C, Cayrel A, Hamon S, Kochko AD, Hamon P (2006) SSR mining in coffee tree EST databases: potential use of EST-SSRs as markers for the *Coffea* genus. Theor Appl Genet 276:436–449

Ramírez M, Graham MA, Blanco-López L, Silvente S, Medrano-Soto A, Blair MW, Hernández G, Vance CP, Lara M (2005) Sequencing and analysis of common bean ESTs. Building a foundation for functional genomics. Plant Physiol 137:1211–1227

Rohlf FJ (1994) NTSYS-pc: numerical taxonomy and multivariate analysis system. Applied Biostatistics, New York

Romo S, Labrador E, Dopico B (2004) Water stress-regulated gene expression in *Cicer arietinum* seedlings and plants. Plant Physiol Biochem 39:1017–1026

Rozen S, Skaletsky HJ (1997) 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

Scott KD, Eggler P, Seaton G, Rossetto M, Ablett EM, Lee LS, Henry RJ (2000) Analysis of SSRs derived from grape ESTs. Theor Appl Genet 100:723–726

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

Sethy NK, Shokeen B, Edwards KJ, Bhatia S (2006a) Development of microsatellite markers and analysis of intra-specific genetic variability in chickpea (*Cicer arietinum* L.). Theor Appl Genet 112:1416–1428

Sethy NK, Choudhary S, Shokeen B, Bhatia S (2006b) Identification of microsatellite markers from *Cicer reticulatum*: molecular variation and phylogenetic analysis. Theor Appl Genet 112:347–357

Shan F, Clarke HC, Plummer JA, Yan G, Siddique KHM (2005) Geographical patterns of genetic variation in the world collections of wild annual *Cicer* characterized by amplified fragment length polymorphisms. Theor Appl Genet 110:381–391

Tayyar RI, Waines JG (1996) Genetic relationships among annual species of *Cicer* (Fabaceae) using isozymes variation. Theor Appl Genet 92:245–254

Thiel T, Michalek W, Varshney RK, 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

Tian AG, Wang J, Cui P, Han YJ, Xu H, Cong LJ, Huang XG, Wang XL, Jiao YZ, Wang BJ, Wang YJ, Zhang JS, Chen SY (2004) Characterization of soybean genomic features by analysis of its expressed sequence tags. Theor Appl Genet 108:903–913

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

Varshney RK, Thiel T, Stein N, Langridge P, Graner A (2002) In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species. Cell Mol Biol Lett 7:537–546

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

Varshney RK, Sigmund R, Börner A, Korzun V, Stein N, Sorrells ME, Langridge P, Graner A (2005b) Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. Plant Sci 168:195–202

Vigouroux Y, Jaqueth JS, Matsuoka Y, Smith OS, Beavis WD, Smith JSC, Doebley J (2002) Rate and pattern of mutation at microsatellite loci in maize. Mol Biol Evol 19:1251–1260

Weeden NF, Muehlbauer FJ, Ladizinsky G (1992) Extensive conservation of linkage relationships between pea and lentil genetic maps. J Hered 83:123–129

Winter P, Benko-Iseppon AM, Hüttel B, Ratnaparkhe M, Tullu A, Sonnante G, Pfaff T, Tekeoglu M, Santra D, Sant VJ, Rajesh PN, Kahl G, Muehlbauer FJ (2000) A linkage map of the 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

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

Yi G, Lee JM, Lee S, Choi D, Kim BD (2006) Exploitation of pepper EST-SSRs and an SSR-based linkage map. Theor Appl Genet 114:113–130

Zhang LY, Bernard M, Leroy P, Feuillet C, Sourdille P (2005) High transferability of bread wheat EST-derived SSRs to other cereals. Theor Appl Genet 111:677–687

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