Appl Biochem Biotechnol (2015) 175:93–118
DOI 10.1007/s12010-014-1249-8

Development of Novel Simple Sequence Repeat Markers in Bitter Gourd (*Momordica charantia* L.) Through Enriched Genomic Libraries and Their Utilization in Analysis of Genetic Diversity and Cross-Species Transferability

Swati Saxena · Archana Singh · Sunil Archak · Tushar K. Behera · Joseph K. John · Sudhir U. Meshram · Ambika B. Gaikwad

Received: 30 June 2014 / Accepted: 10 September 2014 /
Published online: 21 September 2014

Abstract Microsatellite or simple sequence repeat (SSR) markers are the preferred markers for genetic analyses of crop plants. The availability of a limited number of such markers in bitter gourd (*Momordica charantia* L.) necessitates the development and characterization of more SSR markers. These were developed from genomic libraries enriched for three dinucleotide, five trinucleotide, and two tetranucleotide core repeat motifs. Employing the strategy of polymerase chain reaction-based screening, the number of clones to be sequenced was reduced by 81 % and 93.7 % of the sequenced clones contained in microsatellite repeats. Unique primer-pairs were designed for 160 microsatellite loci, and amplicons of expected length were obtained for 151 loci (94.4 %). Evaluation of diversity in 54 bitter gourd accessions at 51 loci indicated that 20 % of the loci were polymorphic with the polymorphic information content values ranging from 0.13 to 0.77. Fifteen Indian varieties were clearly distinguished indicative of the usefulness of the developed markers. Markers at 40 loci (78.4 %) were transferable to six species, viz. *Momordica cymbalaria*, *Momordica subangulata subsp. renigera*, *Momordica balsamina*, *Momordica dioca*, *Momordica cochinchinesis*, and *Momordica sahyadrica*. The

S. Saxena · S. Archak · A. B. Gaikwad (✉)
Division of Genomic Resources, National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi 110012, India
e-mail: ambika@nbpgr.ernet.in

A. Singh
Division of Biochemistry, Indian Agricultural Research Institute, New Delhi 110012, India

T. K. Behera
Division of Vegetable Science, Indian Agricultural Research Institute, New Delhi 110012, India

J. K. John
National Bureau of Plant Genetic Resources, Thrissur 680656 Kerela, India

S. U. Meshram
Rajiv Gandhi Biotechnology Centre, RTM Nagpur University, Nagpur 440001 Maharashtra, India

microsatellite markers reported will be useful in various genetic and molecular genetic studies in bitter gourd, a cucurbit of immense nutritive, medicinal, and economic importance.

Keywords *Momordica charantia* · Microsatellite markers · SSRs · Genomic library enrichment · Genetic diversity · Transferability

Introduction

Momordica charantia L. ($2x=2n=22$), commonly known as bitter gourd or bitter melon, is an economically important cucurbitaceous vegetable cultivated in India, China, Malaysia, Africa, and South America [1]. The place of its origin is probably India with secondary center of diversity in China [2]. Compared with other cucurbits, it is highly valuable for its nutritive content of proteins, carbohydrates, vitamins, and minerals mainly ascorbic acid and iron [3, 1], as well as varied medicinal properties [4]. The fruit and juice have been used as traditional medicine for treating diabetes [5], as a remedy for curing anemia, jaundice, malaria, and cholera [6] and for inhibiting breast cancer cell proliferation [7].

Indian bitter gourd provides immense phenotypic variation based on various characters such as growth habit, maturity, fruit shape, size, color, and surface texture [8] and sex expression [9]. Compared with morphological markers, DNA-based molecular markers are more useful for genetic characterization and diversity assessments. Various multi-locus dominant DNA markers such as RAPD [10, 11], ISSR [12], and AFLP [13] have been reported for genetic analyses of bitter gourd. Microsatellite or simple sequence repeat (SSR) markers have gained considerable importance in plant genetics and breeding owing to many desirable attributes like their multi-allelic nature, co-dominant transmission, extensive genome coverage, small amount of starting DNA, and ease of detection by polymerase chain reaction (PCR) [14]. However, the number of microsatellite markers available in bitter gourd is few. Of the 70 SSR markers reported, 16 have been developed using FIASCO technique [15, 16], 11 through genomic library enrichment [17], and 43 through cross-species transferability from other cucurbits [18–20]. SSRs are known to have high hetrozygosity values and are more informative than dominant DNA markers [14]. Although the initial cost of SSR marker development is high, once developed, they are highly repeatable and, consequently, easily transferred across laboratories [21]. Notwithstanding their suitability in various genetic analyses, microsatellite markers provide accurate results with a minimum number of loci/alleles employed in the study. For instance, 350 to 400 alleles were proposed to be needed to detect genetic relationships among common wheat varieties [22]. It is established that greater number of markers are necessary for the development of genetic map and marker-assisted selection [23]. Considering the decisive advantages of having greater number of SSR markers in bitter gourd, we report in this study the development and characterization of 160 new microsatellite markers in *M. charantia* through a cost-effective strategy of genomic library enrichment. The utility of these markers in analysis of intra-specific genetic diversity, as well as cross-species transferability to six species of *Momordica* is also demonstrated.

Materials and Methods

Construction of SSR-Enriched Genomic Libraries

Genomic DNA was extracted from fresh leaves of the variety Pusa Vishesh using a modified CTAB method [24]. A small insert genomic library enriched for ten different repeat motifs was

constructed with minor modifications [25]. The 10 μg of genomic DNA was nebulized and size-fractionated on 1 % agarose gel to elute fragments in the range of 0.5–1 kb. These were end-polished and ligated to adaptors OG-1 (5′-CTCTTGCTTAGATCTGGACTA-3′) and OG-2 (5′-TAGTCCAGATCTAAGCAAGAGCACA- 3′) using T_4 DNA ligase. Adaptor-ligated DNA was PCR-amplified at 94 °C for 5 min, 35 cycles of 94 °C for 45 s, 55 °C for 45 s, and then 72 °C for 1 min with a final extension of 72 °C for 5 min. The amplified DNA was denatured and hybridized to 5'-biotinylated repeat oligonucleotides: $(GA)_{20}$, $(CA)_{20}$, $(AT)_{20}$, $(CAT)_{15}$, $(AGA)_{15}$, $(GAT)_{15}$, $(AAT)_{15}$, $(AAC)_{15}$, $(AAAT)_{15}$, and $(AAAG)_{15}$. Microsatellite-containing fragments captured by streptavidin-coated magnetic beads (New England Biolabs, Inc., USA) were washed twice with 6× SSC and 0.1 % SDS at 65 °C and room temperature. The various fragments enriched for different repeats were pooled and ligated into pGEM-T Easy vector (Promega, USA) and transformed into DH5α *Escherichia coli* electro-competent cells. Subsequent to selection on IXA (IPTG, X-gal, and ampicillin) plates, 3,072 white colonies were picked and inoculated in 1 ml LB broth with ampicillin and incubated overnight at 200 rpm and 37 °C. A 50-μl aliquot of each clone was stored as a glycerated culture at −80 °C. Plasmid DNA isolation was followed as per manufacturer's protocol (Wizard miniprep, Promega, USA).

PCR Screening and Sequencing

Identification of clones containing microsatellite inserts was performed as described previously with minor modifications [26, 27]. One primer complementary to adaptor OG-1 and one non-biotinylated repeat primer used for library enrichment were used for PCR amplification. PCR reaction was carried out in a total volume of 10 μl with 1× *Taq* buffer, 2 mM $MgCl_2$, 0.25 mM dNTPs, 0.25 μM of each primer, and 0.5 U *Taq* DNA polymerase (MBI, Fermentas), cycled at 94 °C for 5 min, 35 cycles of 94 °C for 1 min, 55 °C for 1 min, 72 °C for 1 min, and final extension of 72 °C for 5 min. Clones containing two or more amplified fragments were identified by agarose gel (1.5 %) electrophoresis and were sequenced employing M13 universal primers.

Microsatellite Mining, Primer Design, and Validation

The sequence data were mined for microsatellites using the software Simple Sequence Repeat Identification tool [28]. The parameters for detection were a minimum total array size of eight nucleotides [29]. Thus, a minimum number of four and three repeats were considered for di- and trinucleotide motifs, respectively, and that of two repeats for penta-, hexa-, hepta-, and octanucleotide repeats. The identified microsatellites were characterized as perfect, compound, and interrupted (both simple interrupted and compound interrupted). Primers were designed flanking the microsatellite regions using Primer3 software [30] based on the following parameters: G+C content of ~50 %, primer length (18–24 bases), annealing temperature of the primer in 50–65 °C range, and product size in the range of 150–300 bp. The primer pairs were validated for amplification of the expected size of allele using the total genomic DNA of Pusa Vishesh by optimizing for the annealing temperatures (Table 1).

Genetic Diversity Analysis and Cross-Species Amplification

To demonstrate the efficacy of the newly developed primer pairs, 51 of these were employed for genetic analysis of 54 bitter gourd genotypes (Table 2). Genomic DNA (10 ng) was amplified in 20-μl reaction mixture consisting 1× *Taq* buffer, 2 mM $MgCl_2$, 0.25 mM

Table 1 Primer sequences and associated information for the novel 160 SSR loci developed in the study

S. no.	Primer ID	Repeat motif	Left primer sequence (5′-3′)	Right primer sequence (5′-3′)	T_a °C	Product size (bp)
1	McSSR 1*	$(GA)_5$	GACAAAAACAACAACCAGAGGC	CTCCTCCTTCTTCTCTCTGCG	60	207
2	McSSR 2*	$(GA)_4(GAA)_5GGAGA(GAA)_4$	AGGGGAATAACAGAGAGGTGG	TGCTAATTTGCCTCTCGTCG	55	221
3	McSSR 3*	$(AAAGG)_2(GGAGA)_2G(AGAGGA)_2$	TTTTGTCAATTTTCCCGACG	TTTCATCTTCCTCTCGATCTCC	55	187
4	McSSR 4*	$(AT)_{12}$	TCCCGCTTCCTCACATCTGC	GGGGTTGAAACACGAGAGTGC	55	204
5	McSSR 5*	$(CTT)_3GA(TCT)_4TTTTC(TCTT)_2$	CTTTAACTCACCTTCCACACCC	ACGATATGATCGAATGTCCACC	55	176
6	McSSR 6*	$(AG)_4C(GGAG)_2(AAG)_3(AG)_6$	CGTGATTTTGTTTCGCCACC	TAAAACCGAAACCGAAACCC	65	181
7	McSSR 7*	$(GA)_4CAA(GAAAG)_2$	AGAGAGGGAGAACGAGACGG	TTTATATGATGGGTCACTTGGC	60	185
8	McSSR 8*	$(AG)_4AAA(AATT)_2$	TGTAGGCGTGGAGCGAGAGG	CCCTTCCTCGAATCATTCACC	64	187
9	McSSR 9	$(GT)_{12}$	GAAAATGGTCAGTGTGTGAGCG	GCACACGCACACTCACTGGC	NA	170
10	McSSR 10	$(AAG)_3$	CAATTGAGCCACCTTTTGGG	TAGCATCGATCCATGGCTCC	NA	197
11	McSSR 11*	$(TATGTG)_2$	TCGTTGTTTCTCCCTCTCTCG	GCATAACACAGAATTGAGGGACC	48	151
12	McSSR 12*	$(CTT)_5$	CGATCTGCGAATCTTGCAGG	TCCTTCGAGGGAGAAGCACC	58	171
13	McSSR 13*	$(GA)_6$	GTTCGGGATCTTCTTGCTCG	TCCCTTCTCCCCATCTCTCC	57	190
14	McSSR 14*	$(TA)_6$	TTGCATGCTTTTTGGTAGAGC	GACTCATCTACCGAATCAACGG	48	153
15	McSSR 15*	$(TC)_4G(CT)_5$	GGAGGCGTCGTAAGATTCCG	ACATTTGCCAAGCGGAGAGG	57	160
16	McSSR 16*	$(TTC)_5$	GGCTTCCTTCAGTGAGTGCG	GTCTGTCGATGCGTCTTCGG	63	151
17	McSSR 17*	$(GAA)_6$	ACGAAGGCTCTCTTTCGTCG	ACGCCATGTCTGAAGAAGCG	50	159
18	McSSR 18*	$(AT)_8$	TAAAGAATCGGCCAGTTCGG	GGGGTTAGAGAAAATGAGAGGC	58	134
19	McSSR 19*	$(AAG)_5$	GAATAGCTTTCGTCGCCTGC	CGGATATCTCCGCTTCTCTCC	62	194
20	McSSR 20*	$(TCT)_9$	GGAATTCAGGTGAACCTGACG	CCAGGAGGAAGAGGAACTGC	60	214
21	McSSR 21	$(GA)_5GGG(ATGA)_2$	GAAGTTGAGGGAGGGAGAGG	TCTCTCCTCCCTCATCCTCG	NA	163
22	McSSR 22*	$(GGTTC)_3$	CCATGACCGATGTAGCACTCC	TCGAACCAACCTAAACCAG	58	170
23	McSSR 23*	$(GCAT)_2$, $(AATG)_2$	AGGTGGCCCTCTCTCAATCT	TATGTCGGCAGTCTCCCTCT	57	201
24	McSSR 24*	$(AT)_4$	TCGGGAATTGGATTTTATGATT	GGCCTAATGTTGCAAAACCT	50	250

Table 1 (continued)

S. no.	Primer ID	Repeat motif	Left primer sequence (5′-3′)	Right primer sequence (5′-3′)	T_a °C	Product size (bp)
25	McSSR 25*	$(GA)_4$	CCTTGAGGAGCCTACGTTGA	AATGGGCTCACCTTTGAGAA	52	183
26	McSSR 26*	$(GCG)_4ACAC(TGAG)_2$, $(TGAA)_2$	TCCATTTTCTTTCGCAATCC	TGTTATTGGCTCCCTCTGCT	53	206
27	McSSR 27*	$(TCTCGA)_2$	ATTTCCATTTCGCGATTCAG	GCCTTGTTTTCCGAAAGAGAT	50	155
28	McSSR 28*	$(CT)_4$	GGAACTTTTGCTCGCATTGT	TGCCATCCACACCAGAATAA	50	188
29	McSSR 29*	(CT)5, $(GA)_4$, $(TTTTCAT)_2$	TGCCATTTCGGGTTAAGAAG	CTGCGGAAAAATAGCTCGAC	50	231
30	McSSR 30*	$(AG)_4$	ATTCCTAAAACGGCAGGTGA	CTTTGCTCTCTCCCGTTCC	53	171
31	McSSR 31*	$(AG)_4$	CCTTGACCCTGAGATTGAGC	GTCTCTGTTGTCCGCCATCT	54	247
32	McSSR 32*	$(TC)_4$	CCGATCCTTGTTTACCAACC	TCTCGAGAAACAAGTGGGCTA	52	179
33	McSSR 33*	$(GT)_4$, $(AAGA)_2$, $(AGAA)_2AAT(GCAA)_2$, $(AAGA)_2$, $(GTGA)_2$	CCCCAGTGAGGACACTGTTT	TTTTTCTTTCCCCCACTCTT	48	236
34	McSSR 34*	$(GC)_4$, $(CACGCG)_2$	ACGCCAACGATATACCACCT	CCCATGGTTTGAGGTCATTC	50	171
35	McSSR 35*	$(AG)_4$	TTAGCTGCTCGCTTGAGGAT	CAAGGATTCTCACATTTCCACA	43	225
36	McSSR 36*	$(AGAGA)_2$	AACGGTTGTTTTCACTCCAAA	AAGCAAAAAGATGGGGGAAA	48	231
37	McSSR 37*	$(CTTTT)_3CG(TTTCC)_2(TC)_5$	CGCGAGGAGTTTTCTTCAAC	CTGCTGTGGTTCCTCCCTAC	57	198
38	McSSR 38*	$(TTTC)_2C(TTAA)_2(AAAAAG)_3$	CACCAGAACCGGAAGAAGAG	CAGAAGGCAGTGTTTGGTGA	61	210
39	McSSR 39*	$(CT)_4$	GGAACTTTTGCTCGCATTGT	TGCCATCCACACCAGAATAA	50	188
40	McSSR 40*	$(GA)_4$	AAATCTTAAGGCGCATGGAA	GGAACACACCTAAGGAGATGTCA	60	171
41	McSSR 41	$(TTAGGGT)_3(TTAGGG)_{15}$	ATTCGATCGATGCTTCACTG	TTAATGATAATTACCCTGAC	NA	212
42	McSSR 42*	$(TA)_4$	TCCAATAAACTAAACATCCAAGG	GGGCCGTATCCATAATGTTG	46	228
43	McSSR 43*	$(GAAAT)_2$	TCACTTGGAGGAAACACAAAAA	CCCACCTCATAAAGGCATTC	55	242
44	McSSR 44*	$(AGGG)_2$, $(GAA)_3$, $(AATA)_2$	TGGCTAGGTAAGCGTCCTGT	ACTACGGCGACGAAGAATCA	58	189
45	McSSR 45*	$(TGA)_3G(GAT)_3(TAT)_3$	TGTTTCTATTCGGATCATGGTT	GAACCCTTTGTGCTGGTGTT	56	250
46	McSSR 46*	$(GGTA)_2$, $(CT)_4$	ATACCTCGAGCCAATGTTCG	ACCCCTTTCTCCCGAAGTTA	57	208
47	McSSR 47*	$(TA)_4$	TTGATTTTGAATCAGCGTTGT	ATTTTGCACAAGGCCTACCA	51	246
48	McSSR 48*	$(ATTA)_3(AATTA)_2$	TCCATTGGAATTGTTGTAACG	GGCTTTTTGGCCCTTAATCT	54	234

Table 1 (continued)

S. no.	Primer ID	Repeat motif	Left primer sequence (5′-3′)	Right primer sequence (5′-3′)	T_a °C	Product size (bp)
49	McSSR 49*	$(TA)_4$	AACCTTTACAGAGCGGGTCA	TGCATTGTCCAAAATCCAAT	52	226
50	McSSR 50	$(CTC)_4$, $(CTTCGA)_2$, $(GGGA)_2$	TCTTGCTTAGATCTGGACTACCG	CGATTCCCTTTTCACTCTGC	NA	162
51	McSSR 51*	$(AG)_5$	CCATCCACCGTTTTTGTTCT	TCTGCCATTGATGTGCTTGT	52	230
52	McSSR 52*	$(CT)_5$, $(GA)_4$, $(TTTTCAT)_2$	TGCCATTTCGGGTTAAGAAG	CTGCGGAAAAATAGCTCGAC	46	231
53	McSSR 53*	$(TC)_4$	TCTGCAAAACCCAAGAAAGG	AAGTTCCCCTCAAACACCAC	54	178
54	McSSR 54*	$(ATCAT)_2$	CCATCCATATCCCAATTCCA	TCATCACAAACCTCCCTTTTTC	46	173
55	McSSR 55*	$(TC)_4$	ATCCAACCAATAACCGGAAG	CTACCATTTTGGGGACGAGA	46	179
56	McSSR 56*	$(CT)_4$	TGCCATACTCCCAGGAAAAG	CGGAGACCTGTGTTTTTGGT	47	222
57	McSSR 57	$(GAA)_4$, $(AATA)_2$, $(AAATT)_2$	TTCAGAATCCCAATCCAAGG	TGACAACCTCGTTTTCCTCTC	55	192
58	McSSR 58	$(CTCCTA)_2$	CTTGAAAGGCGCTCAAAAAG	AAGGACCCATGACGATGAAG	55	152
59	McSSR 59	$(TTTTGAGA)_2$	ATTCTCCGGAACCACAAGAA	GTTGGAGATAAGCGGACTCG	55	232
60	McSSR 60	$(AGCTTG)_3$	TAGTTGATGGCACGTTGCTC	GACACCCGACCTAGGAGTTG	55	179
61	McSSR 61	$(AGGCGCC)_2$	TTAGGACCATTTGGGAGTGC	ACCAAAACGCATTGGAAGAC	55	150
62	McSSR 62	$(ATGACAA)_2$	GAGCTTCGAAACGACTTTCA	AAACCCAAGACCACCAACAC	55	156
63	McSSR 63	$(CGCAAA)_2$, $(TCA)_3$	TATGCTCAAAACCCCGATTC	ATCGGGACTAGACCAGCAAC	55	179
64	McSSR 64	$(CTT)_5$	TCTGGACTACCTCAGGATCG	GGAGTCTTATGGGGGTCCTT	55	212
65	McSSR 65	$(AAGC)_2$, $(AGTTCGG)_2$, $(AGC)_3$	AGCACAAGGTCAGAGGGAAA	GGACTAGGAAGGTCGGAACC	55	158
66	McSSR 66	$(AAG)_4$, $(GAA)_7$	TTCAGAATCCCAATCCAAGG	TTTCTGCCATTTTTCTTATTATT	52	244
67	McSSR 67	$(AGCT)_2(CTCCTA)_2$, $(TC)_4(TTTA)_2$	TCCGCCCCTACTCAACTAAA	ATATCTCGTTACCCCCATGC	55	265
68	McSSR 68	$(AT)_4AG(AT)_2AG(AT)_4$, $(GAAA)_2(GGGC)_2$	CTTCTCTTTGCCCCTTACGA	CAGTGCCCCACAACTATGAA	55	289
69	McSSR 69	$(TTAT)_2TGAT(TTCAG)_2TTTG(AGTT)_2$, $(AGTA)_2$, $(TTATT)_2$	TGGACTAATGGTTCAAGGACCTA	GCAATCACACCATATCACATCA	NA	226
70	McSSR 70	$(CATC)_2C(ACCCG)_2$	AGATCTGGACTAGGGTAGCAAA	GCCCCTTCACTTTGTTCAAT	NA	291
71	McSSR 71	$(AG)_8GAGAG(GA)_4$	AAATAAATTAGCCGATCTTTGCAT	TCATTTCTGATCTGGAAAACCA	52	178
72	McSSR 72	$(ATTTGA)_2$, $(AGTAGA)_2$, $(GTGA)_2$, $(AACT)_2$	TGCAGCATCCATAGCCATAC	GGCAGTGTGATGTGATTCTGA	55	247

Table 1 (continued)

S. no.	Primer ID	Repeat motif	Left primer sequence (5′-3′)	Right primer sequence (5′-3′)	T_a °C	Product size (bp)
73	McSSR 73	$(CAAT)_2$, $(ATCG)_2$, $(TGAAAA)_2$, $(GAAAT)_2$	AATGGGGATATTCCCGAAAC	AATGGGAGCAAGAATTTCCA	52	249
74	McSSR 74	$(GT)_{13}$	GCCAAGGGAAATTGTAATACG	AAACAACGTTGATGGCAAGA	52	200
75	McSSR 75	$(AAAAAT)_2$, $(TGAA)_2$	GAGTCCAGGTCTTGGGATTG	TCAGAGAGCACCCTTGCTAA	55	208
76	McSSR 76	(GAA)5, $(CCAAC)_3TTA(AAAC)_2$	AAATTTGGGAGAGGGTAGGC	TGGGATGGGCTTATTGTTTT	52	172
77	McSSR 77	$(TCAT)_2$, $(TTG)_4$	GCTTGTGGAGCCTTTCCTAA	TGGATCAAAAACGTGGTCAA	52	215
78	McSSR 78	$(AGACGG)_2$	AGCTGTTGGGTGGTTAGGAC	CATTGAGTTCACCGCCATTA	55	217
79	McSSR 79	$(TTCTT)_2$, $(GAATA)_2$	TGTGCTCGGGGTAGAAGTTT	CCGGGAAAGGGTAGAAGAAT	55	174
80	McSSR 80	$(TCCA)_2$, $(TATGA)_2$	GAAGAGTTCGACCCAATGCT	CGATGGAATCTCATCATCCA	55	242
81	McSSR 81	$(GAAGGA)_2$	CGAGTGACATTGCTTCTTCG	TTCATTGGGCCTTTCGATAC	55	245
82	McSSR 82	(CAAC)2, $(TTTATGAA)_2TTT(TTTAAA)_2$	CGAGGAGTCACTCGATCAAA	CGCTGCCCACAGAAAATTA	55	216
83	McSSR 83	$(GCGAAA)_2$, $(TATT)_2$	CAAGATTTTACCATGACTGCAA	TACTGGAGGAGCAGCAATGA	55	186
84	McSSR 84	$(GT)_{10}(GTGTTA)_2$, $(TGA)_3$, $(AGTG)_2$	AGAGAAAATGGTCAGTGTGTGA	CTGGACTAGCACACGCACA	55	187
85	McSSR 85	(ATCC)2, $(TTCAATA)_2TGGAAA(AAAAAATG)_2$	TCCTAGGCGTAGAGGAACCA	AGTGGGAGAGAAGGGGTTTC	55	236
86	McSSR 86	$(TCTA)_2$	ACTCGTATGGGTGCCTTTTG	ATGTTGATTGGGCAGGAAGT	55	223
87	McSSR 87	$(GTG)_3$	CCTCGGCCCTCATACTTAGA	CCCTATGCTCACGAACCAAT	55	207
88	McSSR 88	$(GGGTCGG)_2$	GTTGTATGGCTCGGGTAGGA	CCCACCCCGTATAAAATCAA	55	219
89	McSSR 89	$(TTAT)_2$, $(ATGTTG)_2$	CAAATTCCGGTCTCCAATGT	AACGCAGGTCGGATCTATCT	55	151
90	McSSR 90	$(TAAAG)_2$, $(AAACA)_2$	ACGTGCTCTTTCCTCCAAAA	AAATCCCGAGCACTTTACATTC	52	169
91	McSSR 91	$(AGA)_3AA(GA)_4$, $(GAAA)_2$, $(AAGAGAGA)_3$	TGTTGATCGTCACCGAAATC	CCCATTCTTTGTTTGTTTTCTCTT	52	196
92	McSSR 92	$(TTTA)_2$, $(CTTAAG)_2$	AGGCTCTCCAGAGCTTTCCT	TTGGAACTGAACACCCTGTG	55	171
93	McSSR 93	$(ACTGTT)_2$	TGGACTAGGAGAAGTCGTTTGA	CCCCAGTAAAAATCCCATCTT	55	186
94	McSSR 94	$(GTTTGAA)_2(GT)_7$	CCTACATTCGACGGGACACT	TACCCCAAACACAGCAACAC	55	224
95	McSSR 95	TGTT)2, $(CGATATC)_2$	GTGTTGCTGTGTTTGGGGTA	GGATTATTTCCAGAACGGACA	55	227
96	McSSR 96	$(GTG)_3$, $(TTGG)_2$, $(AGTG)_2$	GCATGCTGAATTGTGTTGGT	GTGTAACAGCCCTCGACCAT	55	210
97	McSSR 97	$(AAAG)_2$, $(GAG)_4$	CACATAAGCCGACATTACCC	TGCCTCTAAGGGTTCTTTCC	55	212

Table 1 (continued)

S. no.	Primer ID	Repeat motif	Left primer sequence (5′-3′)	Right primer sequence (5′-3′)	T_a °C	Product size (bp)
98	McSSR 98	$(CCACC)_3$	CCTTAGTGGCTAGGAGGAACC	GCTTTTGGACCTTCACATCC	55	250
99	McSSR 99	$(GAGAA)_2$	TATCTGATGGTGGCGAGATG	GCACTCCCAAATGGTCCTAA	NA	247
100	McSSR 100	$(AGGCGCC)_2$	TTAGGACCATTTGGGAGTGC	CCAAATCGTGCTCAAACTGA	55	212
101	McSSR 101	$(GGTC)_2$, $(AGA)_3$, $(AAAG)_2$	CTCTACGATTCCGACCGTCT	TCTTATTCTCCCCCTTCCTTTT	55	166
102	McSSR 102	$(GAG)_4G(GAAA)_3$	GAGGGAGAAGTGGAAGGGATA	CAATGGGATGGGGATTTTATT	55	152
103	McSSR 103	$(GA)_6$	TTCTTGCTCGGAGACAAATG	GGCCAATTCTTTCCCTTTTC	52	160
104	McSSR 104	$(GCCG)_2$, $(GACC)_3$	ACAGAGCGTAGGCTTGCTTT	ATTGGAGGGCAAGTCTGGT	55	245
105	McSSR 105	$(GA)_5$	TCGATCAGTTTTGGTCGAAAT	CCGACATTCTTTCTTGCACA	52	201
106	McSSR 106	$(AAGA)_2$, $(TTAGT)_2$, $(GAT)_3$	AAGAGCTGCTGGTGGAGAAC	CCGATGCTACATCATCAACAA	55	250
107	McSSR 107	$(TCACA)_2$, $(GATTTAGG)_2$	GAAGCACAATCACTCGTTGC	GAACGGGTGTTACCTGAGGA	55	236
108	McSSR 108	$(CAAAA)_2$, $(TGAA)_2$	GCAAATTTCTCATTTCCTCTTGA	ACCCACCCAGATGAATGAAT	52	187
109	McSSR 109	$(GGTGACG)_2$, $(AAGG)_2$	GGGAATTCGATTCTCTCTCG	CCGTGTCAGGATTGGGTAAT	55	183
110	McSSR 110	$(ATACAG)_2$, $(GGGATA)_2$, $(AATC)_2$	CGGGAAGGAATTGGAATGTA	TCATTGAGCGAAAGGTACGA	52	180
111	McSSR 111	(TTTC)2, $(TTTA)_2$	TACTATTGGCTTGGGCATGA	GAGAGGAAAAGAGGGGGAAA	55	170
112	McSSR 112	$(TTAATAAG)_2$, $(AGTT)_2$, $(AAAATG)_2$, $(ATTGAC)_2$	ACCCATAGTCCAGGCTTCAA	TGTCGGCATCTACAATGGTC	55	227
113	McSSR 113	$(TCCTT)_2$, $(AATG)_2$	CACGGAAACATCCGACCTAT	TTTTGGGGAATATGGGTTGA	52	249
114	McSSR 114	$(ATCTTTT)_2$	TTGGTGCATTTGAAAGTTCG	CGCCCCTAAAATCATCAGAC	55	241
115	McSSR 115	$(AGGA)_2$	GCAATGACCCTGTTTGTTCT	CAAAGGAAGAGTGCACTTGTGT	55	250
116	McSSR 116	$(CAAA)_2$, $(ACATA)_2$	TGTTTGAATGTAATGAGCCTATCC	TCCAATGCTGAATCGATGAC	55	207
117	McSSR 117	$(AATAA)_2$, $(TAA)_3$	GTCATCGATTCAGCATTGGA	GACGCAGCATGGTACTCTTTC	55	211
118	McSSR 118	$(AAG)_2$, $(AATA)_2$	TGGCTAGGTAAGCGTCCTGT	CTACGGCGACGAAGAATCA	55	188
119	McSSR 119	$(TTTTC)_2$	CGATAGGGCCTCATTGGTAA	ATTCCACAACAACGAAAGCA	52	150
120	McSSR 120	$(TTTAGC)_2$	AATGGGATGCCCTAATACGTT	TGTGGTCACAACCAGAAAGG	55	233
121	McSSR 121	(TTTG)2, $(TGAA)_2(ACAA)_2$	TGAAATTTTGAGGTTATGTTCTCG	TCTTTTTCTTATGCATGCCTTTT	52	240
122	McSSR 122	$(TTATAAA)_7$, $(AATC)_2$, $(GTTG)_2$,$(AAAT)_2$	TATCCAGGCTCCGCTTAGAA	GACAAATGCCCCAATAGCAT	55	221

Table 1 (continued)

S. no.	Primer ID	Repeat motif	Left primer sequence (5′-3′)	Right primer sequence (5′-3′)	T_a °C	Product size (bp)
123	McSSR 123	$(GAAGCG)_2$	TGGGATGTAAAAATGCATCG	GTCCATCGACTACGCCTTTC	55	217
124	McSSR 124	$(TTTC)_2$, $(GGTTTGCG)_2$	GCTACCCCCTCATTTTCCTC	TCGATCACTGAGGCTGGAT	55	194
125	McSSR 125	$(AT)_5$, $(AT)_4(CATA)_2$, $(AT)_4CTG(AT)_4(CATA)_2$	TGCAATTTTTATTATTCCAG	TTCGATGTAACTTTGATATACT	52	240
126	McSSR 126	$(AGTTTG)_2$	TGGACTACCTCGCACCTCTT	CATTCCAGCAGTTGGTTCAA	55	227
127	McSSR 127	(AGCCC)2, $(CACCAG)_2$, $(AGTAG)_2$	CAGAACCATCCTGTGGAACA	TGGAGCCCCTCAAGTTTTT	55	230
128	McSSR 128	$(TTTCAA)_2$, $(TAAT)_2$	TCTGGTTCACCGCTTTAGGT	AGGGAAGTTGTGAGCATTACG	55	173
129	McSSR 129	$(TTGGG)_2(TCGG)_2$, $(CCGG)_2$	GATCAATTGGAGGGCAAGTC	AGGCTTGCTTTGAGCACTCT	55	241
130	McSSR 130	$(GAAGCA)_3$, $(CAGAT)_2$	TCTTTTTCATTCCCCCTTTG	GAACTGCACGGAGTTGATGA	52	215
131	McSSR 131	$(ACAAGC)_2$	GGGGGCAATGGAATACACTA	GGCGTGAATGCAAATAAAAA	52	245
132	McSSR 132	$(TTCAAC)_2$, $(TAAT)_2$	TCTGGTTCACCGCTTTAGGT	AGGGAAGTTGTGAGCATTACG	55	173
133	McSSR 133	$(ATTATAT)_2$	CGCGTTTGTAATTCCATCAA	GCCCGCTTATTCATCTTTACA	52	226
134	McSSR 134	$(AAAG)_2AACTTGA(ATTT)_2$, (TATC)2	GGTATCAAACCAATAACGATTCA	GCCCCTAGAGGTCGTAGAGA	55	243
135	McSSR 135	$(ACTAAT)_2(CTAATA)_2$	AGGACTCACTGAGCCGAGAT	GATTCTGGCTTTCGTGCTTT	55	203
136	McSSR 136	$(GGGATA)_2$, $(AATC)_2$	CGGGAAGGAATTGGAATGTA	TCATTGAGCGAAAGGTACGA	52	180
137	McSSR 137	$(AAGG)_2$, $(TCCTGC)_2$	CCGAAATGGGTTCCTTACAA	TTTGGCAGCTAATCCTCTTGA	52	167
138	McSSR 138	$(TTTGTG)_2$, $(AAAT)_2$	TGTGCTCCAAGAACTTCAACA	CTTATCATATTTGTCGCAAGCA	52	156
139	McSSR 139	$(AGCA)_2$, $(TTTG)_2$, $(TGAA)_3$, $(CTTC)_2$	CCTACCCTTCTCGAGCCTAC	AGTTGTTTTTGGGTGGGATG	55	228
140	McSSR 140	(ATAA)2, $(ATAGATAA)_2T(ATGG)_2$, $(ATCC)_2$, $(GGTT)_2$	AGGACCAATGAGATGCAAAAA	TTGGTACCGTCCAATCGAA	52	230
141	McSSR 141	$(TTTCC)_2TC(TTTTGTT)_2$	TTGGTGGATAAGCACGTCAG	GAGAGCAGAGCCAAGGCTTA	55	212
142	McSSR 142	$(TTCGTA)_2$	TCCGAAGGTCTAAAGGATCG	ATTGTCAGGTGGGGAGTTTG	55	185
143	McSSR 143	$(ACATCT)_2$	TGTTTACAGCAGCAATTCAACA	TTTTGATGGGTCCTTTTTGC	52	187
144	McSSR 144	$(TCAC)_2$, $(TTTG)_2$	AGCAAACAATAGCAGCGAAA	CGTTCCACTACTAATTCAAGGAAA	52	213
145	McSSR 145	$(AATCTA)_2$	TTACAGGCTGCCGTATTCTG	TTGATTCATTGACAGGTGCAT	52	240
146	McSSR 146	$(GCTGCC)_2AAAA(AAGGC)_2$	AAGAAGGGGAGGCAAATGTT	CCAAATTGCAGTGGAAACAC	52	207
147	McSSR 147	$(ACTAAT)_2CTAATACT(AATAG)_2$	GAGCCCTCTTTCTCCTCGAT	CGAGATCCTTTCGATGACCT	55	242

Table 1 (continued)

S. no.	Primer ID	Repeat motif	Left primer sequence (5′-3′)	Right primer sequence (5′-3′)	T_a °C	Product size (bp)
148	McSSR 148	$(ATACAG)_2$, $(GGGATA)_2$, $(AATC)_2$	CGGGAAGGAATTGGAATGTA	TCATTGAGCGAAAGGTACGA	52	180
149	McSSR 149	$(AGAA)_2$	TTCATTTTGAGGGGTTCAGG	TCGTGGATTTGAACTTTTATGG	52	242
150	McSSR 150	$(TGTGAA)_2$	AAGACTTGAGATTGAATCCACCA	AGAGAGGAAAACGCACCAAC	55	230
151	McSSR 151	$(GGAGCT)_2$	GACGATATCGACCGTGACCT	CATCTTTCACAATCCCTGGAG	55	239
152	McSSR 152	$(TAAAT)_2$, $(GAAAA)_2$	CCATATTCCCCAAAAAGTGG	CGATAGGGCCTCATTGGTAA	55	229
153	McSSR 153	$(ACAAGC)_2$	GGGGGCAATGGAATACACTA	GGCGTGAATGCAAATAAAAA	52	245
154	McSSR 154	$(CTTTCG)_2$	TGCGGAAGAAAGGAAAAAGA	GTTTAGGTTCGGCCTCAATG	52	240
155	McSSR 155	$(CCAAAG)_2$, $(ATTTTT)_2$, $(TTTAGT)_2$	GTTGGCCATGGAATAAAGGA	GGAGATCCAAACCAAGAAGC	NA	173
156	McSSR 156	$(GTACTA)_2$, $(TACTTG)_2$	TGTAGGTCCGGGATAATCCTT	TTTACGCCCCGTAATTCTTC	55	246
157	McSSR 157	$(GGTC)_3$	GATCAATTGGAGGGCAAGTC	AGGCTTGCTTTGAGCACTCT	55	241
158	McSSR 158	$(CAG)_{12}$	TCATCAACAACAACAATTCCA	TCTTGAATTGCACCGAACAC	52	150
159	McSSR 159	$(TCGATT)_2GGTT(TCTG)_2$	ATCACGGTTGAGGGCTAATG	GTTCGATCGGCCAGAATATC	55	231
160	McSSR 160	$(TGAATC)_2$	GATTGGAAATCGATGGAGGA	TCTTATCTTGCCCCTGCTTC	55	205

*Primers used for evaluating genetic diversity and cross-species transferability among bitter gourd genotypes and six *Momordica* species

NA not amplified, T_a annealing temperature

Table 2 List of 54 bitter gourd genotypes used for genetic diversity analysis

S. no.	Genotypes	Place of collection
1	DBG-3	Baud, Odisha, India
2	DBG-7	Bhopal, Madhya Pradesh,India
3	DBG-8	Shilong, Meghalaya,India
4	DBG-9	Mayurbhanj, Odisha,India
5	DBG-33	West Bengal, India
6	DBG-34	West Bengal, India
7	DBG-38	West Bengal, India
8	DBG-41	West Bengal, India
9	DBG-42	West Bengal, India
10	DBG-46	West Bengal, India
11	DBG-51	West Bengal, India
12	DBG −52	West Bengal, India
13	EC620325	Thailand
14	EC620326	Thailand
15	EC620327	AVRDC Taiwan
16	EC620329	Philippines
17	EC620330	Philippines
18	EC620331	Philippines
19	EC620332	Philippines
20	EC620333	Philippines
21	EC620335	Indonesia
22	EC620337	Pakistan
23	EC620338	Lao People's Democratic Republic
24	EC620341	AVRDC Taiwan
25	EC620346	AVRDC Taiwan
26	EC620351	AVRDC Taiwan
27	EC620352	AVRDC Taiwan
28	EC620353	AVRDC Taiwan
29	EC620354	P.R. of China
30	EC620355	P.R. of China
31	EC620356	P.R. of China
32	EC620357	P.R. of China
33	EC620358	P.R. of China
34	EC620359	AVRDC Taiwan
35	EC620360	AVRDC Taiwan
36	Pusa do Mausami	IARI, New Delhi, India
37	Pusa Vishesh	IARI, New Delhi, India
38	Priya	KAU, Kerala, India
39	DBGy-201	IARI, New Delhi, India
40	DBGy-202	IARI, New Delhi, India
41	NDBT-9	NDAUT, Faizabad, India
42	Arka harit	IIHR, Bangalore, India
43	Kalyanpur Baramasi	Kanpur, U.P, India
44	Phule green gold	MPKV, Rahuri, India

Table 2 (continued)

S. no.	Genotypes	Place of collection
45	Selection-5	IARI, New Delhi, India
46	Hirakani	MPKV, Rahuri, India
47	NDBT-7	NDAUT, Faizabad, India
48	Kashi Urvashi	IIVR, Varanasi, India
49	Meghana–2	IIVR, Varanasi, India
50	Selection–1	IARI, New Delhi, India
51	Preeti	KAU, Kerala, India
52	Punjab bittergourd-14	PAU, Ludhiana, India
53	Phule Ujwala	MPKV, Rahuri, India
54	HABG–22	IIVR, Varanasi, India

dNTPs, 0.25 μM of each primer, and 0.5 U *Taq* DNA polymerase (MBI, Fermentas) in a thermocycler (Biometra). The PCR conditions were as follows: initial denaturation at 94 °C for 4 min, 30 cycles of denaturation at 94 °C for 40 s, annealing at experimentally determined annealing temperature of each primer pair (Table 1) for 40 s, and then extension at 72 °C for 40 s followed by final extension at 72 °C for 5 min. The amplified fragments were separated on 4 % metaphor gel and visualized by ethidium bromide staining. These primer pairs were also tested for their cross-species transferability potential on a representative panel including one accession each of *Momordica cymbalaria*, *Momordica subangulata subsp. renigera*, *Momordica balsamina*, *Momordica dioca*, *Momordica cochinchinesis*, *Momordica sahyadrica*, and *M. charantia* as control. PCR and electrophoresis conditions were the same as described above. The data for diversity analysis of 54 bitter gourd genotypes at each locus were scored as present (1) or absent as (0). The Free Tree software version 0.9.1.50 [31] was used to calculate the Jaccard's similarity coefficients [32], and the resulting matrix was subjected to UPGMA cluster analysis. Similar analysis was done for a subset of 15 released Indian varieties of bitter gourd. Parameters of genetic diversity including expected heterozygosity (H_e) and the observed heterozygosity (H_o) were calculated using POPGENE version 1.3 [33]. The polymorphic information content (PIC) was calculated for each locus using the formula, $PIC = 1 - \sum p_i^2$ [14], where p_i is the frequency of the i^{th} allele.

Sequencing for Validation of Alleles

Amplicons obtained in seven bitter gourd genotypes (including Pusa Vishesh) and seven *Momordica* species showing allelic variation for the locus McSSR_54 $(ATCAT)_2$ were gel-eluted, purified, and sequenced. The sequences were compared using ClustalW multiple sequence alignment tool [34]. Sequence analysis was also done for the locus McSSR_20 $(TCT)_9$ that showed allelic variation in another set of nine bitter gourd genotypes.

Results

Microsatellite Isolation and Characterization

A small insert genomic library of bitter gourd enriched for ten different repeat motifs was constructed. The repeats employed in the study were three di-repeats: $(GA)_{20}$, $(CA)_{20}$, $(AT)_{20}$;

five tri-repeats: $(CAT)_{15}$, $(AGA)_{15}$, $(GAT)_{15}$, $(AAT)_{15}$, $(AAC)_{15}$, and two tetra-repeats: $(AAAT)_{15}$, and $(AAAG)_{15}$. The library comprised of 3,072 clones, and upon PCR-based screening, 582 clones putatively containing microsatellite repeats were identified. The presence of a microsatellite repeat was indicated by the presence of two bands when amplified using one primer as the repeat used for library enrichment and the other as OG-1. High-quality sequence data were obtained for 466 clones (80 % sequencing efficiency). The sequence information is available for these in NCBI Genbank (accession numbers KC905911 to KC906159 and KC918559 to KC918775).

Sequence analysis with the software SSRIT showed that 437 of the clones sequenced contained in all 1,939 microsatellite repeat stretches of a minimum of eight nucleotides (dinucleotide core repeated four times, trinucleotide core repeated three times, and tetra-, penta-, hexa-, hepta-, and octanucleotide cores repeated twice). These simple sequence repeat stretches comprised 1,814 (93.55 %) simple, three (0.15 %) simple interrupted, 38 (1.95 %) compound, and 84 (4.33 %) compound interrupted repeats. Tetranucleotide repeats were the most abundant class of microsatellites accounting for 1,087 (56 %) of the repeats. These were followed by 335 pentanucleotides (17.2 %), 182 hexanucleotides (9.38 %), 164 trinucleotides (8.45 %), 113 dinucleotides (5.82 %), 47 heptanucleotides (2.4 %), and 11 octanucleotide (0.56 %) repeat containing microsatellites, respectively.

Six dinucleotide repeat motifs, GA/CT, AG/TC, AT/TA GT/CA, AC/TG, and CG/GC were found 40, 35, 24, 6, 4, and 4 times, respectively. Twenty-seven trinucleotide repeat motifs, AGA/TCT, CTT/GAA, AAC/TTG, AAG/TTC, and CAA/GTT were found 26, 19, 13, 12, and 10 times, respectively. As many as 118 tetranucleotide motifs were identified of which AAAT/TTTA, ATTT/TAAA, AAAG/TTTC, AATA/TTAT, CTTT/GAAA, and AGAA/TCTT were found 40, 38, 37, 37, 31, and 30 times, respectively. An assorted penta-, hexa-, hepta-, and octanucleotide repeat motifs were identified and found 175, 123, 34, and 11times, respectively.

Primer Designing and Validation

Based on the availability of sufficient length of sequences flanking the repeat region in the clone and compatible melting temperatures, 160 primer pairs (Table 1) were designed from 437 of the microsatellite containing clones (primer design efficiency was 36.6 %) in the enriched genomic library of bitter gourd. The designed primers were validated for PCR amplification using the genomic DNA of the variety Pusa Vishesh, which was used for library construction. The PCR amplification efficiency of the designed primers was 94.4 %, as sharp discrete fragments of expected size were obtained for 151 primer pairs.

Genetic Diversity Analysis with Newly Developed Microsatellite Markers

Genetic diversity among 54 bitter gourd accessions was evaluated by amplifying markers using 51 of the above developed primer pairs. A representative amplification profile is given in Fig. 1. A total of 68 alleles were amplified, ranging in size from 127 to 300 bp of which 28 (41.17 %) that were generated by ten primer pairs were polymorphic. The number of alleles at these loci ranged from two to five with an average of 2.80 alleles per locus. The indices of utility of these loci are given in Table 3. The observed heterozygosity (H_o) ranged from 0.042 to 0.587 (average=0.106), and the expected heterozygosity (H_e) ranged from 0.140 to 0.783 (average=0.373). PIC values ranged from 0.139 to 0.775 (average=0.369).

Jaccard's similarity coefficient-based UPGMA cluster analysis of the SSR marker data revealed three major clusters and three small groups of two to three accessions each (Fig. 2). First major cluster of 26 accessions comprised eleven landraces and nine released varieties.

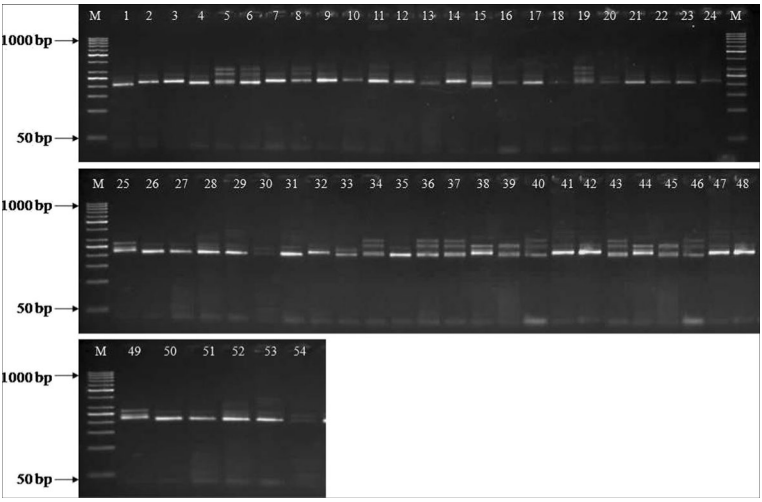

Fig. 1 Amplification of alleles of 214 and 205 base pairs at the locus McSSR_20 in 54 bitter gourd genotypes. *M* is the 50 bp DNA ladder, *lanes 1–54* bitter gourd genotypes as appearing in Table 1

The second major cluster comprised mostly of the exotic Asian germplasm, i.e., from China (two accessions), Taiwan (five accessions), Philippines (three accessions), Pakistan, and Indonesia (one accession each) and two released varieties, Meghana-2 and NDBT-9. The third cluster contained five varieties, three landraces, and three exotic accessions. The genotypes EC620354, Selection-5, EC620330, and EC620332 were out-grouped from the rest of the accessions. The clustering pattern was not in consonance with the geographical distribution of the genotypes used in the study. The maximum similarity value (0.9818) was observed between indigenous genotypes, DBG-8 and DBG-34, and minimum similarity was seen between DBGy-201 and EC620330. A UPGMA dendrogram based on Jaccard's similarity coefficient was constructed for a subset of 15 released varieties (Fig. 3). All the 15 varieties were discriminated with the maximum similarity (0.96) between the varieties Hirakani and Phule Ujjwala, and, the varieties HABG-22 and Kalyanpur Baramasi were the most distant varieties (0.75).

Table 3 Characteristics of ten polymorphic SSR loci

Primer ID	Core motif	A_T	H_o	H_e	PIC
McSSR_11	$(TATGTG)_2$	2	0.042	0.157	0.155
McSSR_17	$(GAA)_6$	2	0	0.14	0.139
McSSR_18	$(AT)_8$	2	0	0.14	0.139
McSSR_20	$(TCT)_9$	5	0.211	0.783	0.775
McSSR_22	$(GGTTC)_7$	2	0.226	0.202	0.2
McSSR_27	$(TCTCGA)_2$	2	0	0.258	0.256
McSSR_47	$(TA)_4$	3	0.587	0.668	0.661
McSSR_54	$(ATCAT)_2$	2	0	0.542	0.537
McSSR_55	$(TC)_4$	3	0	0.447	0.443
McSSR_56	$(CT)_4$	3	0	0.398	0.394

A_T total number of alleles, H_o observed heterozygosity, H_E expected heterozygosity, *PIC* polymorphic information content

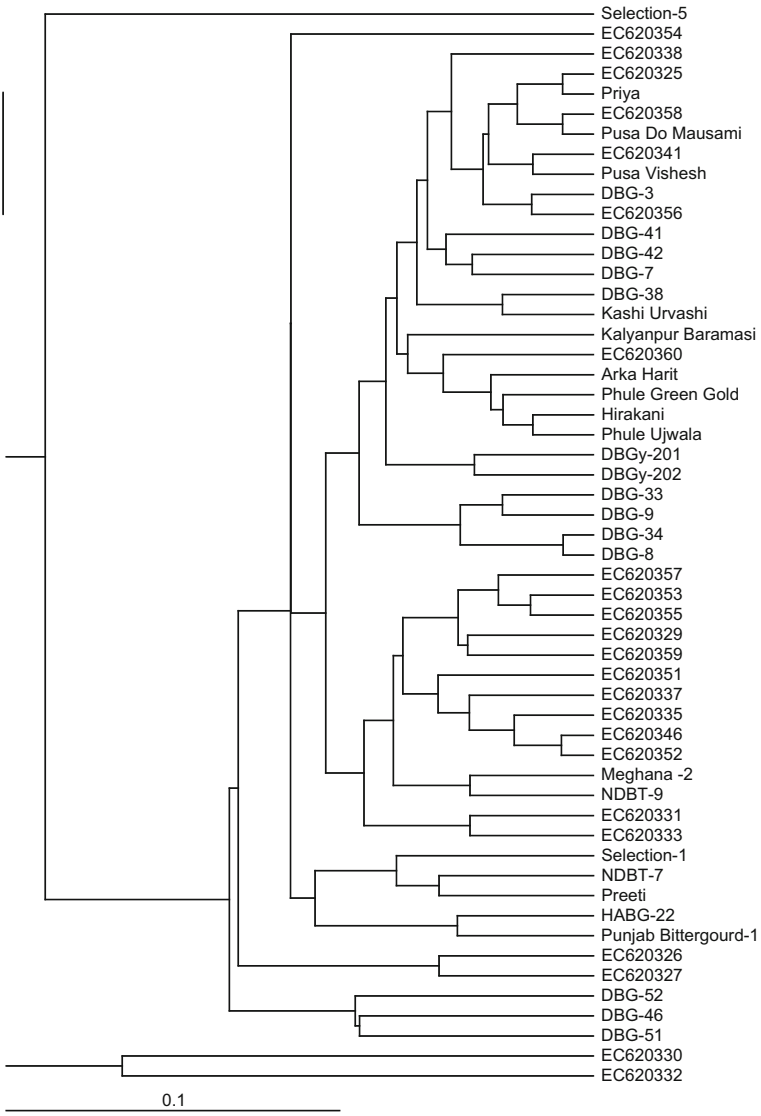

Fig. 2 UPGMA-based dendrogram of 54 bitter gourd genotypes based on 51 STMS loci

Marker Transferability to Other *Momordica* Species

Cross-species amplification was tested in 51 microsatellite markers in six different *Momordica* species: *M. cymbalaria*, *M. subangulata subsp. renigera*, *M. balsamina*, *M. dioca*, *M. cochinchinesis*, and *M. sahyadrica*. One accession of each species (including that of *M. charantia* as control) was used for detecting cross-species transferability potential of these markers (Fig. 4). Forty SSR primer pairs (78.43 %) designed from *M. charantia* sequences were transferable to other species while 11 primer pairs failed to amplify in any of the six *Momordica* species. Whereas 12 of these 40 primer pairs gave amplification products in all the six *Momordica* species, six primer pairs amplified in five species, four primer pairs in four

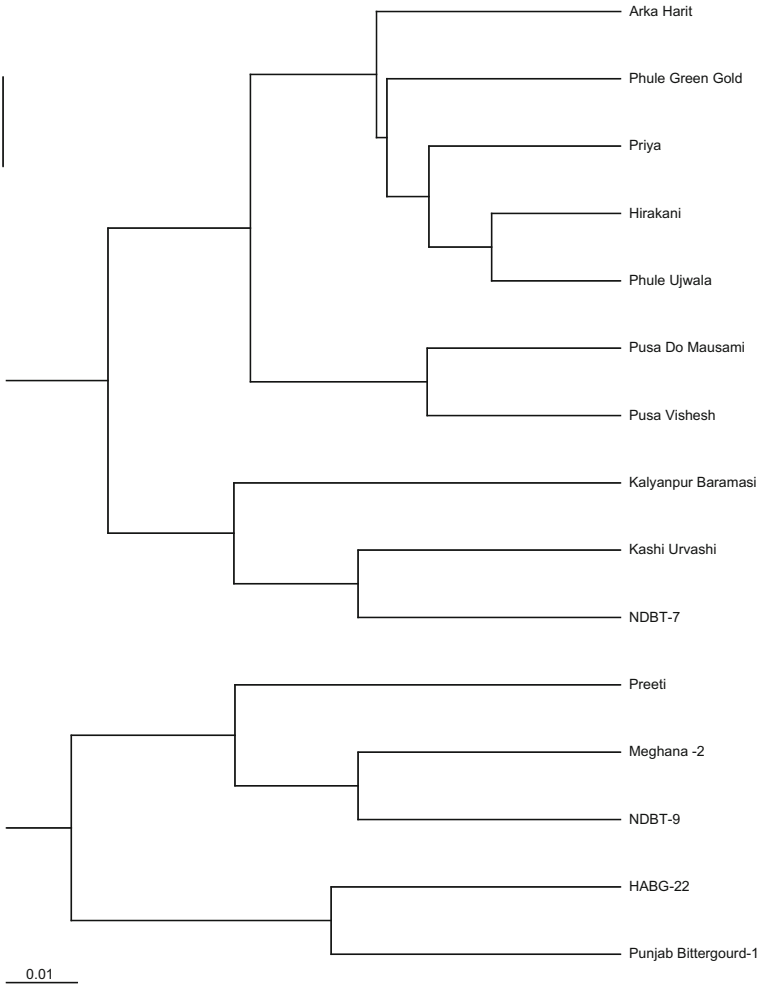

Fig. 3 UPGMA-based dendrogram of 15 bitter gourd varieties based on 28 polymorphic STMS markers

species, five primers pairs in three species, four primer pairs in two species, and nine primer pairs in only one of the six species included in the study (Table 4). Transferability of the 40 SSR markers was the most to *M. balsamina* (30), followed by *M. sahyadrica* (29), *M. cymbalaria*, and *M. dioca* (27), *M. subangulata subsp. renigera* (24), and least in *M. cochinchinesis* (13). UPGMA analyses based on Nei's genetic distance values (Fig. 5) generated a dendrogram with two major clusters with *M. cymbalaria* as an outlier. Maximum similarity (0.5483) was observed between *M. subangulata subsp. renigera* and *M. dioca* and minimum (0.2321) between *M. cymbalaria* and *M. charantia*.

Validation of Allelic Variation by Sequencing

Amplicons of sizes 173 and 169 bp were amplified with the primer pair McSSR_54. Sequence comparison of the 173-bp fragment eluted by amplification of the genotypes Pusa Vishesh, Arka Harit, HABG-22, Phule Green Gold, and Kalyanpur Baramasi, and the169-bp fragment

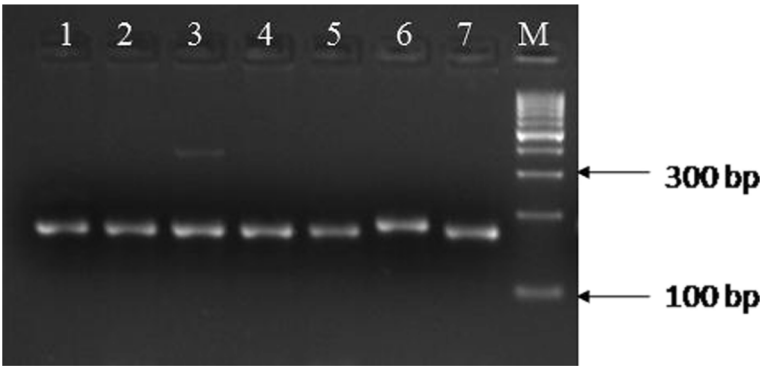

Fig. 4 Amplification of alleles of 173 and 165 bp at the locus McSSR_54 in seven *Momordica* species. *M* is the 100-bp DNA ladder, *lanes 1 M. cymbalaria, 2 M. subangulata subsp. renigera, 3 M. balsamina, 4 M. dioca, 5 M. cochinchinesis*, *6 M. charantia*, *7 M. sahyadrica*

eluted from the genotypes Priya and Kashi Urvashi revealed that the polymorphism was due to a 4 bp (CATA) indel falling in the repeat motif $(ATCAT)_2$ at nucleotide position 126 (Fig. 6a).

Table 4 Cross-species transferability of SSR markers derived from *M. charantia*

Primer ID	*M.cymbalaria*	*M. subangulata subsp. renigera*	*M.balsamina*	*M.dioca*	*M.cochinchinesis*	*M.sahyadrica*
McSSR_3	+	+	+	+	—	+
McSSR_4	−	−	+	−	−	−
McSSR_5	+	+	−	−	−	+
McSSR_6	+	+	+	+	+	+
McSSR_8	−	−	−	−	−	+
McSSR_11	−	−	−	+	−	+
McSSR_12	−	−	−	−	−	+
McSSR_13	+	+	+	+	−	+
McSSR_14	−	+	+	+	−	+
McSSR_15	−	−	−	−	−	+
McSSR_16	+	+	+	+	+	+
McSSR_18	+	+	−	+	−	+
McSSR_19	+	−	+	+	−	−
McSSR_20	+	+	+	−	−	−
McSSR_22	+	−	−	+	−	+
McSSR_23	+	+	−	−	−	−
McSSR_26	−	+	+	−	−	−
McSSR_27	+	+	+	+	+	+
McSSR_29	+	+	+	+	+	+
McSSR_31	−	−	−	+	+	+
McSSR_32	+	+	+	+	+	+
McSSR_33	+	−	+	+	−	+
McSSR_34	+	+	+	+	−	+
McSSR_35	−	−	+	+	−	−

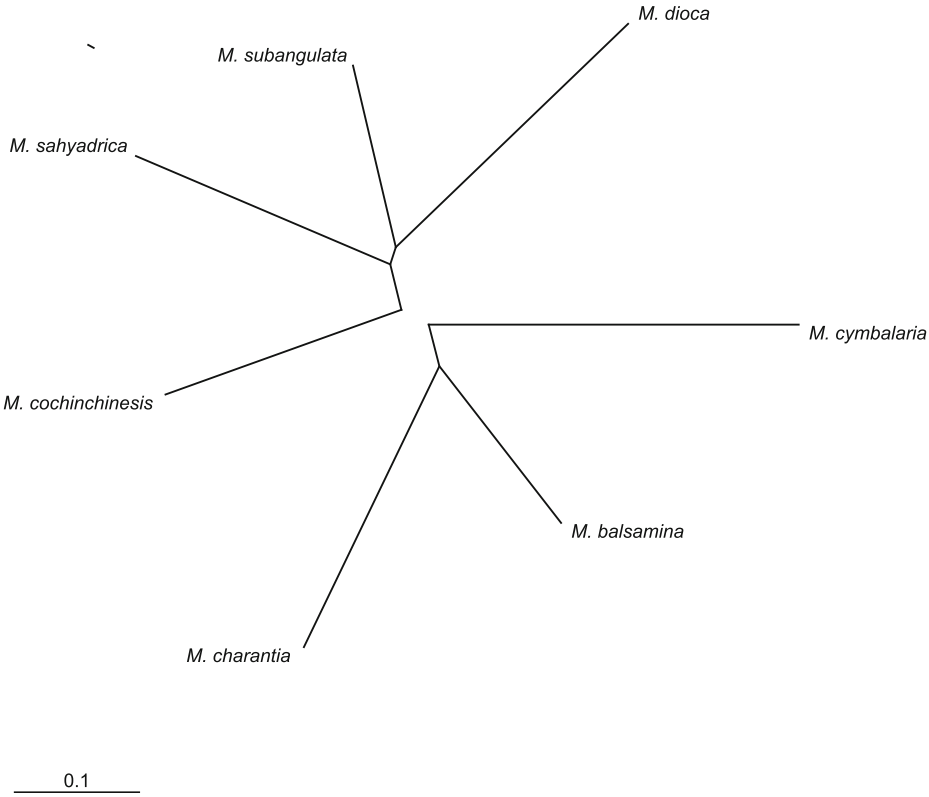

Fig. 5 UPGMA dendrogram based on Nei's genetic distance values for *Momordica* species

Sequence comparison of the fragments amplified at the same locus (McSSR_54) across *Momordica* species showed an 8 bp (CATATCAT) indel at position 121 in *M. charantia* that is a part of the $(ATCAT)_2$ repeat motif that was absent in all the other six species (Fig. 6b). Sequence analysis of the 214 bp fragment from the genotypes Pusa Vishesh, DBG-7, DBGy-201, EC620341, and EC620356, and 205-bp fragment from the genotypes DBG-3, DBG-9, Preeti, and EC620326 at the locus McSSR_20 revealed a 9 bp (TTTCTCCAT) indel in DBG-3, DBG-9, Preeti, and EC620326 from nucleotide positions 87–95 in the flanking region of the $(TCT)_9$ microsatellite repeat (Fig. 6c).

Discussion

The inherent properties of microsatellite markers make them the most preferred type of markers for various biotechnological applications [35]. Efforts to isolate microsatellite markers in bitter gourd have been few. SSR markers have been developed in bitter gourd using different isolation strategies such as Fast Isolation of by AFLP of sequence-containing repeats (FIASCO) technique [15, 16] and genomic library enriched for a single trinucleotide $(GAA)_{10}$ probe [17]. Small insert genomic library enriched for different microsatellite repeats is an effective method for targeting many repeat-containing loci [25]. The large number of clones to be sequenced while generating microsatellite markers through construction of

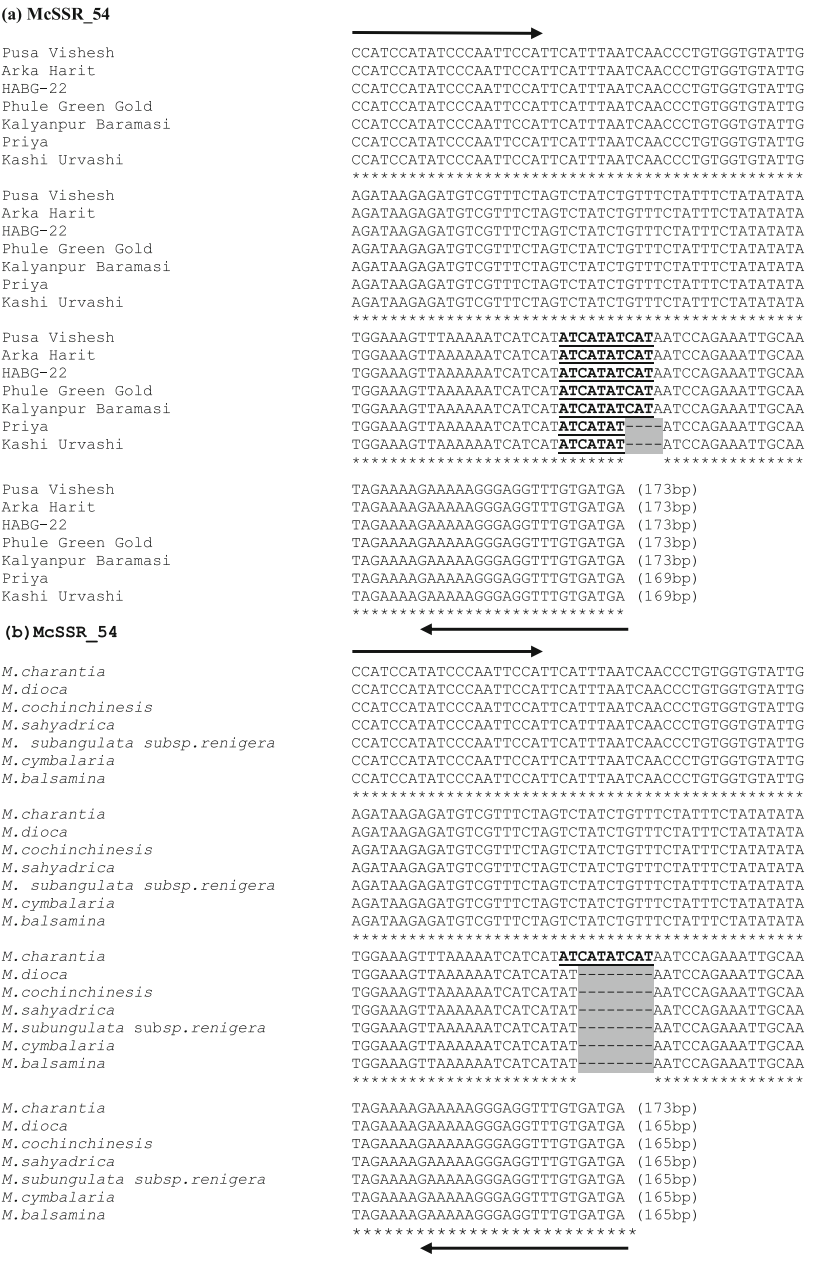

Fig. 6 Sequence comparison between alleles amplified at microsatellite loci McSSR54 and McSSR_20. The repeat motifs are represented as *bold letters* and *underlined. Asterisk* indicates similar sequences, and *dash* indicates alignment gaps. *Shaded box* indicates base insertions/deletions. *Left* and *right arrows* indicate primer sequences. **a** Sequence alignment of alleles amplified using McSSR_54 from bitter gourd varieties: Pusa Vishesh, Arka Harit, Phule Green Gold, Kalyanpur Baramasi, HABG-22, Priya, and Kashi Urvashi. **b** Sequence alignment of alleles amplified using McSSR_54 from *Momordica* species: *M. charantia* var. Pusa Vishesh*, M. dioca, M. cochinchinesis, M. sahyadrica*, *M. subangulata subsp. renigera, M. cymbalaria*, and *M. balsamina*. **c** Sequence alignment of alleles amplified using McSSR_20 from bitter gourd genotypes: Pusa Vishesh, DBG-7, DBGy-201, EC620341, EC620356, DBG-3, DBG-9, Preeti, and EC620326

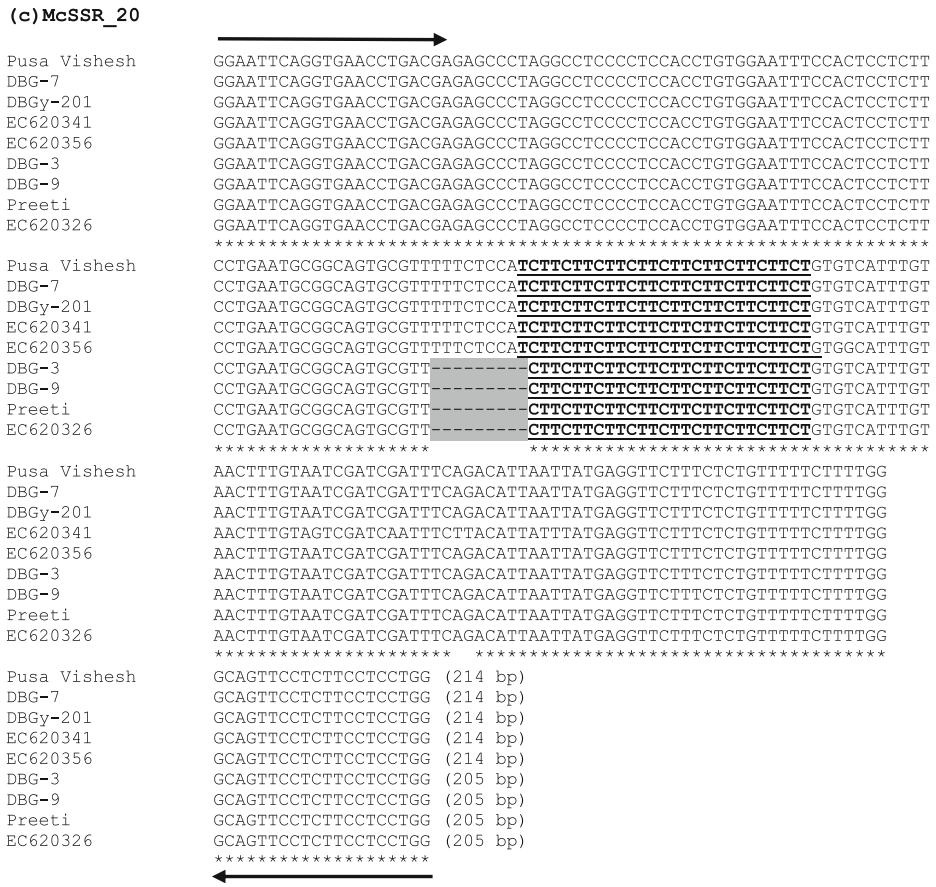

Fig. 6 (continued)

enriched genomic libraries is what makes this method more expensive than mining them from databases or transferring them from related species. PCR-based methods to screen for microsatellite containing clones from enriched genomic libraries [27] are a cost-effective means as the number of clones to be sequenced is greatly reduced. The clones that produce two or more bands when screened with the adaptor primer and a repeat-specific primer are considered likely to contain microsatellites [26]. Based on this strategy, the number of clones to be sequenced from the enriched genomic library of the variety Pusa Vishesh was reduced by 81 %, i.e., from 3072 to 582. Quality sequences were obtained for 466 clones of which 437 (93.7 %) contained microsatellite repeats. The detection of a higher percentage of microsatellite-positive colonies following PCR screening makes this a quick and cost-effective method for microsatellite identification.

This is the first report in the genus *Momordica* where an attempt has been made to develop microsatellite markers from a genomic library enriched for ten different repeat motifs. Most simple repeats and their complementary counterparts can be represented by several basic unit patterns [36]. The dinucleotide motifs used for enrichment in the present study were $(CA)_{20}$, $(GA)_{20}$, and $(AT)_{20}$ which represent all the possible combination of two bases, not considering mononucleotide sequences. The 96 % of the 113 detected dinucleotide motifs showed specific

complementarity to these oligoprobes [66 % to (GA) probe, 21 % to (AT) probe, and 9 % to (CA) probe]. Four microsatellites had the motif CG that was not used for enrichment. The relative abundance of $(GA/CT)_n$ repeats in bitter gourd followed by the $(AT/TA)_n$ and $(CA/GT)_n$ repeats is in accordance with that reported for many plant species [37], including related cucurbits [38, 20]. Amongst the 164 trinucleotide repeats detected, 122 (74.4 %) showed sequence complementarity to the five trinucleotide repeat motifs used for library enrichment with those complementary to the $(AGA/TCT)_n$ being predominant as in other plants [39, 40], including *Curbita pepo* and *Cucumis sativus* [41–43]. As in cucumber, tetranucleotide repeat motifs represent the most abundant (56 %) microsatellites found in the present study. The 53 % of the tetranucleotide repeats were AT-rich followed by 31 % having AT=GC and 16 % being GC-rich. The 24.3 % of the 1,087 tetranucleotide motifs detected showed sequence complementarity to the AAAT and AAAG repeats used for enrichment. These were also the most abundant tetranucleotide motifs, as also in cucumber genomic SSRs [43]. Even though three di-, five tri-, and two tetranucleotide repeat motifs were used for library enrichment, surprisingly the second most abundant repeat motifs found in bitter gourd were the pentanucleotide repeat motifs (335), followed by hexa- (182), tri- (164), di-(113), hepta- (47), and octanucleotide (11) repeats. Comparison of microsatellite repeat motifs across eight plant species shows that the tetranucleotide or trinucleotide repeats are most prevalent followed by the dinucleotide repeat motifs [43]. The few numbers of dinucleotide motifs detected in the present study may be attributed to the high-stringency washes during the enrichment with the dinucleotide motifs or PCR bias during screening. The 71 % of the pentanucleotide repeats detected in the present study were AT-rich, and 29 % were GC-rich. Of the 182 hexanucleotides identified in the present study, 56.11 % were AT-rich followed by 27.22 % with AT=GC, and 16.66 % were GC-rich. Prevalence of hexanucleotides repeats in the exons has been reported in different taxa [44]. In cucumber, a higher percentage of these repeats is present in EST database as compared with genomic sequence [43]. Identification of hexanucleotide repeat motifs, therefore, could be of great significance as they may be present in the protein coding region and be responsible for differential expression or regulation of genes which govern important traits in this crop. The heptanucleotides and octanucleotide repeat motifs that were identified amongst the sequenced clones were mostly complementary to the tetranucleotide probes used for enrichment or were a part of compound interrupted repeats identified.

On the basis of taking eight nucleotides as a threshold for screening microsatellite repeats [29], a majority (93.6 %) of the 1,939 microsatellites identified in the present study were simple repeats. This is similar to what has been reported in sugarcane, pumpkin, wheat, and grass pea [41, 45–47] and far greater than 34 % and 41 % reported in *Cucumis melo* [38, 48]. Perfect repeats have higher mutation rates as compared with imperfect repeats and, hence, increase variability and evolutionary rates [49]. These loci are therefore more useful for primer development for diversity analysis studies.

Microsatellites can be easily converted into PCR-based assays promoting their widespread use as molecular markers for genetic diversity analysis [50]. Primer pairs were designed for 160 new STMS loci after considering the factors such as little or no flanking regions surrounding the microsatellite regions and melting temperature (Tm) compatibility (±1 °C). The primer design efficiency of 32.5 % was higher than 21 % and 27 % of the microsatellite containing clones in wheat and sugarcane [51, 52]. Amplification conditions were optimized at 151 loci, resulting in an amplification efficiency of 94.4 %. The 51 primer pairs were used for genetic characterization of 54 bitter gourd genotypes and 20 % (10) primer pairs detected polymorphism with number of alleles ranging between 2 and 5. A total of 28 polymorphic markers were generated resulting in 41.17 % average percentage polymorphism which was

higher than 36.5 % obtained by RAPD markers [10] and lower to that obtained by ISSR (74.7 %) and AFLP (78.5 %) markers in bitter gourd [12, 13]. The average PIC value of STMS markers was 0.369 which is higher than 0.22 obtained by AFLP markers [13] pointing towards the efficiency of these markers in genetic analysis of the bitter gourd germplasm. The PIC values of the polymorphic primers designed in the study ranged from 0.139 to 0.775, respectively. These are similar to previously reported in bitter gourd and melon [17, 19]. These newly developed markers are thus efficient in evaluating genetic diversity in bitter gourd germplasm and highly useful for constructing SSR-based genetic map in bitter gourd.

The informativeness of microsatellites tends to increase with increase in number of repeats [53]. This was observed for the locus McSSR_20 with the core repeat $(TCT)_9$ which had a high PIC value of 0.775. However, the markers at the loci McSSR_47, McSSR_55, and McSSR_56 with repeat number $n=4$ and at McSSR_54 with repeat number $n=2$ were equally capable of detecting polymorphism in bitter gourd accessions, revealing that SSR length is not necessarily a benchmark for detecting polymorphism [54].

The UPGMA-derived dendrogram based on the Jaccard's similarity clustering pattern generated three major clusters and three small clusters with two to three genotypes each. Although the clusters could not be defined on basis of the morphological characters (data not shown), most exotic and indigenous accessions fell in two distinct groups (Fig. 2). The PIC (0.139 to 0.775) and expected heterozygosity (0.140 to 0.783) values of these newly developed co-dominant SSR markers indicate their usefulness in solving issues related to genotype identity, variety protection, and hybrid seed purity testing. An analysis of a subset of 15 Indian-released varieties based on these SSR markers indeed proves the efficiency of the 28 polymorphic markers amplified at ten loci in discriminating these varieties from each other (Fig. 3). These newly developed SSR markers would benefit in discriminating the varieties during plant variety protection processes.

Sequence comparison of the alleles amplified in different varieties at the locus McSSR_54 shows that the allelic variation of 4 bp leading to a 169 bp fragment in the varieties Priya and Kashi Urvashi when compared with Pusa Vishesh was due to a indel of the bases CATA in the repeat motif $(ATCAT)_n$ (Fig. 6a) whereas allelic variation at the same locus across the species showed an indel of 8 bp in all the six *Momordica* species when compared with *M. charantia* var Pusa Vishesh (Fig. 6b). The allelic variation at this locus may be due to indel slippage which is length independent of the repeat array and is thought to operate at repeats with few repetitions at a constant rate [55]. Comparison of the sequences amplified at locus McSSR_20 in nine genotypes (Fig. 6c) shows that the variation between the alleles is due to 9-bp indel in the flanking region of the $(TCT)_9$ repeat, leading to a 205-bp fragment in genotype Preeti, DBG-3, DBG-9, and EC620326. Indels in the flanking regions have been reported to cause interspecific and intraspecific variations at the microsatellite loci in plants and animals [56–59].

DNA sequences flanking microsatellite motifs are often conserved, and SSR markers are highly transferable among species and genera [60, 54]. Transferability of SSR markers to species within the same genus and to different genera of the same family has been reported in many crops such as *Arabidopsis*, rice, wheat, sorghum, soybean, sugarcane, and oat [58, 61–66]. SSR markers developed in bitter gourd have been successfully transferred across species and genera [15–17]. SSR markers at 40 of the 51 tested loci were transferable to six other *Momordica* species, i.e., *M. cymbalaria*, *M. subangulata subsp. renigera*, *M. balsamina*, *M. dioca*, *M. cochinchinesis*, and *M. sahyadrica*. The transferability of the bitter gourd markers to various *Momordica* species ranged from 32.5 % to 75 % which is much more than the across-genera transferability of markers from bottle gourd and cucumber [18, 20]. Cluster analysis of seven species of *Momordica* based on these 40 loci show that the dioecious species:

M. sahyadrica, *M. dioca, M. subangulata subsp. renigera*, and *M. cochinchinesis* formed one major cluster, and while the monoecious species *M. charantia* and *M. balsamina* grouped together, *M. cymbalaria* was distinct from them. These results based on 40 primer pairs are in accordance with the relationship between these Indian species based on morphological [67] as well as molecular (21 RAPD and 12 ISSR primers) markers [68]. The newly developed co-dominant markers, reported here, are thus an important resource not only for intra-specific genetic analyses, but also in discerning species relationship within the genus *Momordica*.

Conclusions

A novel set of informative microsatellite markers for bitter gourd is provided. One hundred sixty novel microsatellite markers have been developed in *M. charantia* through sequencing of small insert genomic library enriched for ten different repeat motifs. Evaluation of a small set of these markers for assessment of genetic diversity among 54 bitter gourd genotypes indicates their potential in genetic analysis of bitter gourd for mapping, variety protection, and hybrid seed purity testing. These markers are highly transferable to six other *Momordica* species and may be used as an efficient tool in phylogenetic and comparative studies among the *Momordica* species. Therefore, the newly developed *M. charantia* microsatellite markers are of great significance as they add up to the few available SSR markers in bitter gourd. These are a valuable resource for genetic characterization of bitter gourd genotypes and accomplish the desire for development of large number of SSR markers for constructing SSR-based genetic map in bitter gourd and related species.

Acknowledgments This work was supported by NBPGR's institutional project on development of genomic tools for enhanced utilization of horticultural crops and a DBT project on marker development in bitter gourd.

References

1. Miniraj, N., Prasanna, K. P., & Peter, K. V. (1993). Bitter gourd *Momordica* spp. In G. Kalloo & B. O. Bergh (Eds.), *Genetic improvement of vegetable plants* (pp. 239–246). Oxford: Pergamon Press.
2. Grubben, G. J. H. (1977). *Tropical vegetable and their genetic resources* (pp. 51–52). Rome: IBPGR.
3. Behera, T. K. (2004). Heterosis in bitter gourd. In P. K. Singh, S. K. Dasgupta, & S. K. Tripathi (Eds.), *Hybrid vegetable development* (pp. 217–221). NY: The Haworth Press.
4. Alam, S., Asad, M., Asdaq, S. M., & Prasad, V. S. (2009). Antiulcer activity of methanolic extract of *Momordica charantia* L. in rats. *Journal of Ethnopharmacology, 123*, 464–469.
5. Baynes, J. W. (1995). *Mechanistic approach to diabetes* (2nd ed., pp. 203–231). Chichester, UK: Ellis Harwood Limited.
6. Ross, I. A. (1999). *Medicinal plants of the world* (pp. 213–219). NJ, USA: Humana Press.
7. Ray, R. B., Raychoudhuri, A., Steele, R., & Nerurkar, P. (2010). Bitter melon (*Momordica charantia*) extract inhibits breast cancer cell proliferation by modulating cell cycle regulatory genes and promotes apoptosis. *Cancer Research, 70*, 1925–1931.
8. Robinson, R. W., & Decker-Walters, D. S. (1997). *Cucurbits*. Wallingford, Oxford, UK: CAB International.
9. Behera, T. K., Dey, S. S., & Sirohi, P. S. (2006). 'DBGy-201' and 'DBGy-202': Two gynoecious lines in bitter gourd (*Momordica charantia* L.) isolated from indigenous source. *Indian Journal of Genetics and Plant Breeding, 66*, 61–62.
10. Dey, S. S., Singh, A. K., Chandel, D., & Behera, T. K. (2006). Genetic diversity of bitter gourd (*Momordica charantia* L.) genotypes revealed by RAPD markers and agronomic traits. *Scientia Horticulturae, 109*, 21–28.
11. Paul, A., & Raychaudhuri, S. S. (2010). Medicinal uses and molecular identification of two *Momordica charantia* varieties—A review. *Electronic Journal of Biology, 6–2*, 43–51.

12. Singh, A. K., Behera, T. K., Chandel, D., Sharma, P., & Singh, N. K. (2007). Assessing genetic relationships among bitter gourd (*Momordica charantia* L.) accessions using inter-simple sequence repeat (ISSR) markers. *Journal of Horticultural Science and Biotechnology, 82*, 217–222.
13. Gaikwad, A. B., Behra, T. K., Singh, A. K., Chandel, D., Karihaloo, J. L., & Staub, J. E. (2008). Amplified fragment length polymorphism analysis provides strategies for improvement of bitter gourd (*Momordica charantia* L.). *Scientia Horticulturae, 43*, 127–133.
14. Powell, W., Machray, G. C., & Provan, J. (1996). Polymorphism revealed by simple sequence repeats. *Trends in Plant Sciences, 1–7*, 215–221.
15. Wang, S. Z., Pan, L., Hu, K., Chen, C. Y., & Ding, Y. (2010). Development and characterization of polymorphic microsatellite markers in *Momordica charantia* (Cucurbitaceae). *American Journal of Botany, 97*, E75–E78.
16. Guo, D. L., Zhang, J. P., Xue, Y. M., & Hou, X. G. (2012). Isolation and characterization of 10 SSR markers of *Momordica charantia* (Cucurbitaceae). *American Journal of Botany, 99*, E182–E183.
17. Ji, Y., Luo, Y., Hou, B., Wang, W., Zhao, J., Yang, L., Xue, Q., & Ding, X. (2012). Development of polymorphic microsatellite loci in *Momordica charantia* (Cucurbitaceae) and their transferability to other cucurbit species. *Scientia Horticulturae, 140*, 115–118.
18. Xu, P., Wu, X., Luo, J., Wang, B., Liu, Y., Ehlers, J. D., Wang, S., Lu, Z., & Li, G. (2011). Partial sequencing of the bottle gourd genome reveals markers useful for phylogenetic analysis and breeding. *BMC Genomics, 12*, 467.
19. Chiba, N., Suwabe, K., Nunome, T., & Hirai, M. (2003). Development of microsatellite markers in melon (*Cucumis melo* L.) and their application to major Cucurbit crops. *Breeding Science, 53*, 21–27.
20. Watcharawongpaiboon, N., & Chunwongse, J. (2008). Development and characterization of microsatellite markers from an enriched genomic library of cucumber (*Cucumis sativus*). *Plant Breeding, 127*, 74–81.
21. Maughan, P. J., Saghai-Maroof, M. A., & Buss, G. R. (1995). Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean. *Genome, 38*, 715–723.
22. Zhang, X. Y., Li, C. W., Wang, L. F., Wang, H. M., You, G. X., et al. (2002). An estimation of the minimum number of SSR alleles needed to reveal genetic relationships in wheat varieties. I. Information from large-scale planted varieties and corner-stone breeding parents in Chinese wheat improvement and production. *Theoretical and Applied Genetics, 106*, 112–117.
23. Tang, R., Gao, G., He, L., Han, Z., Shan, S., Zhong, R., Zhou, C., Jiang, J., Li, Y., & Zhuang, W. (2007). Genetic diversity in cultivated groundnut based on SSR markers. *Journal of Genetics and Genomics, 34–5*, 449–459.
24. Saghai-Maroof, M. A., Soliman, K. M., Jorgenson, R., & Allard, R. W. (1984). Ribosomal DNA spacer length polymorphism in barley: Mendelian inheritance, chromosomal locations and population dynamics. *Proceedings of the National Academy of Sciences, USA, 81*, 8014–8018.
25. Mottura, M. C., Gailing, O., Verga, A. R., & Finkeldey, R. (2004). Efficiency of microsatellite enrichment in *Prosopis chilensis* using magnetic capture. *Plant Molecular Biology Reporter, 22*, 251–258.
26. Gardner, M. G., Cooper, S. J. B., Bull, C. M., & Grant, W. N. (1999). Isolation of microsatellite loci from a social lizard, *Egernia stokesii*, using a modified enrichment procedure. *Journal of Heredity, 90*, 301–304.
27. Wang, X. W., Trigiano, R. N., Windham, M. T., Devries, R. E., Scheffler, B. E., Rinehart, T. A., & Spires, J. M. (2007). A simple PCR procedure for discovering microsatellites from small insert libraries. *Molecular Ecology Notes, 7*, 558–561.
28. Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartinhour, S., & McCouch, S. (2001). Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential. *Genome Research, 11*, 1441–1452.
29. Chambers, G. K., & MacAvoy, E. S. (2000). Microsatellites: Consensus and controversy. *Comparative Biochemistry and Physiology (Part B), 126*, 455–476.
30. Rozen, S., & Skaletsky, H. J. (2000). Primer3 on the www for general users and for biologist programmers. In S. Krawetz & S. Misener (Eds.), *Bioinformatics methods and protocols: Methods in molecular biology* (pp. 365–386). Totowa, NJ: Humana.
31. Pavlieek, A., Pavlieek, T., & Fvlegr, J. (1999). Free tree version 0.9.1.50. *Folia Biology, 45*, 97–99.
32. Jaccard, P. (1908). Nouvelles recherche' sur la distribution florale. *Bull Soc Vaudoise Sciences Naturelles, 44*, 223–270.
33. Yeh, F. C., & Boyle, T. (1999). POPGENE version 1.3.2: Microsoft Window-based freeware for population genetic analysis. http://www.ualberta.ca/~fyeh/index.htm .
34. Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research, 11–22*, 4637–4680.
35. Wang, M.L; Barkley, N.A. and Jenkins, T.M. (2009). Microsatellite markers in plants and insects. PartI: Applications of biotechnology. *Genes, Genomes and Genomics*.

36. Jurka, J., & Pethiyagoda, C. (1995). Simple repetitive DNA sequences from primates: Compilation and analysis. *Journal of Molecular Evolution, 40*, 120–126.
37. Lagercrantz, U., Ellegren, H., & Andersson, L. (1993). The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acid Research, 21*, 1111–1115.
38. Danin-Poleg, Y., Reis, N., Tzuri, G., & Katzir, N. (2001). Development and characterization of microsatellite markers in *Cucumis*. *Theoretical and Applied Genetics, 102*, 61–72.
39. Morgante, M., & Olivieri, A. M. (1993). PCR-amplified microsatellites as markers in plant genetics. *The Plant Journal, 3–1*, 175–182.
40. Gupta, P. K., Balyan, H. S., Sharma, P. C., & Ramesh, B. (1996). Microsatellites in plants: A new class of molecular markers. *Current Science, 70–1*, 45–54.
41. Gong, L., Stift, G., Kofler, R., Pachner, M., & Lelley, T. (2008). Microsatellites for the genus *Cucurbita* and an SSR-based genetic linkage map of *Cucurbita pepo* L. *Theoretical and Applied Genetics, 117*, 37–48.
42. Fazio, G., Staub, J. E., & Chung, S. M. (2002). Development and characterization of PCR markers in cucumber. *Journal of American Society of Horticultural Science, 127–4*, 545–557.
43. Cavagnaro, P. F., Senalki, A. D., Yang, L., Simon, P. W., Harkins, T. T., Kodira, C. D., Huang, S., & Weng, Y. (2010). Genome-wide characterization of simple sequence repeats in cucumber (*Cucumis sativus* L.). *BMC Genomics, 11*, 569.
44. Toth, G., Gaspari, Z., & Jurka, J. (2000). Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Research, 10*, 967–981.
45. Parida, S. K., Kalia, S. K., Sunita, K., Dalal, V., Hemaprabha, G., Selvi, A., Pandit, A., Singh, A., Gaikwad, K., Sharma, T. R., Srivastava, P. S., Singh, N. K., & Mohapatra, T. (2009). Informative genomic microsatellite markers for efficient genotyping applications in sugarcane. *Theoretical and Applied Genetics, 118*, 327–338.
46. Varshney, R. K., Harindra, A. K., Balyan, S., Roy, J. K., Prasad, M., & Gupta, P. K. (2000). Characterization of microsatellites and development of chromosome specific STMS markers in bread wheat. *Plant Molecular Biology Reporter, 18*, 5–16.
47. Lioi, L., & Galasso, I. (2013). Development of genomic simple sequence repeat markers from an enriched genomic library of grass pea (*Lathyrus sativus* L.). *Plant Breeding, 132*, 649–653.
48. Ritschel, P. S., Lins, T. C. L., Tristan, R. L., Buso, G. S. C., Buso, J. A., & Ferreira, M. E. (2004). Development of microsatellite markers from an enriched genomic library for genetic analysis of melon (*Cucumis melo* L.). *BMC Plant Biology, 4*, 9–23.
49. Jin, L., Macaubas, C., Hallmayer, J., Kimura, A., & Mignot, E. (1996). Mutation rate varies among alleles at a microsatellite locus: Phylogenetic evidence. *Proceedings of the National Academy of Sciences of USA, 93*, 15285–15288.
50. Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acid Research, 17*, 6463–6471.
51. Ma, Z. Q., Roder, M., & Sorrells, M. E. (1996). Frequency and sequence characteristics of di-, tri- and tetranucleotide microsatellites in wheat. *Genome, 39*, 123–130.
52. Cordeiro, G. M., Taylor, G. O., & Henry, R. J. (2000). Characterization of microsatellite markers from sugarcane (*Saccharum* spp.), a highly polyploid species. *Plant Science, 155*, 161–168.
53. Weber, J. L. (1990). Informativeness of human poly $(GT)_n$ polymorphisms. *Genomics, 7*, 524–530.
54. Katzir, N., Danin-Poleg, Y., Tzori, G., Karchi, Z., Lavi, U., & Cregan, P. B. (1996). Length polymorphism and homologies of microsatellites in several Cucurbitaceae. *Theoretical and Applied Genetics, 93*, 1282–1290.
55. Coenye, T., & Vandamme, P. (2005). Characterization of mononucleotide repeats in sequenced prokaryotic genomes. *DNA Research, 12*, 221–233.
56. Angers, B., & Bernatchez, L. (1997). Complex evolution of a salmonid microsatellite locus and its consequences in inferring allelic divergence from size information. *Molecular Biology and Evolution, 14*, 230–238.
57. Grimaldi, M. C., & Crouau-Roy, B. (1997). Microsatellite homoplasy due to variable flanking sequences. *Journal of Molecular Evolution, 44*, 336–340.
58. Peakall, R., Gilmore, S., Keys, W., Morgante, M., & Rafalski, A. (1998). Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: Implications for the transferability of SSRs in plants. *Molecular Biology and Evolution, 15*, 1275–1287.
59. Matsuoka, Y., Mitchell, S. E., Kresovich, S., Goodman, M., & Doebley, J. (2002). Microsatellites in *Zea*-variability, patterns of mutations, and their use for evolutionary studies. *Theoretical and Applied Genetics, 104*, 436–450.
60. Barbara, T., Palma-Silva, C., Paggi, G. M., Bered, F., Fay, M. F., & Lexer, C. (2007). Cross-species transfer of nuclear microsatellite markers: Potential and limitations. *Molecular Ecology, 16*, 3759–3767.
61. Clauss, M. J., Cobban, H., & Mitchell-Olds, T. (2002). Cross-species microsatellite markers for elucidating population genetic structure in *Arabidopsis* and *Arabis* (Brassicaeae). *Molecular Ecology, 11*, 591–601.

62. Zhao, X., & Kochert, G. (1993). Phylogenetic distribution and genetic mapping of a $(GGC)_n$ microsatellite from rice (*Oryza sativa* L.). *Plant Molecular Biology, 21*, 607–614.
63. Roder, M. S., Plaschke, J., Konig, S. U., Borner, A., Sorrells, M. E., Tanksley, S. D., & Ganal, M. W. (1995). Abundance, variability and chromosomal location of microsatellites in wheat. *Molecular and General Genetics, 246*, 327–333.
64. Brown, S. M., Hopkins, M. S., Mitchell, S. E., Senior, M. L., Wang, T. Y., Duncan, R. R., Gonzales-Candelas, F., & Kresovich, S. (1996). Multiple methods for the identification of polymorphic simple sequence repeats (SSRs) in sorghum [*Sorghum bicolor* (L.) Moench]. *Theoretical and Applied Genetics, 93*, 190–198.
65. Selvi, A., Nair, N. V., Balasundaram, N., & Mohapatra, T. (2003). Evaluation of maize microsatellite markers for genetic diversity analysis and fingerprinting in sugarcane. *Genome, 46*, 394–403.
66. Fu, Y. B., Chong, J., Fetch, T., & Wang, M. L. (2007). Microsatellite variation in *Avena sterilis* oat germplasm. *Theoretical and Applied Genetics, 114*, 1029–1038.
67. Bharathi, L. K., Munshi, A. D., Behera, T. K., Joseph John, K., Bhat, K. V., & Sidhu, A. S. (2013). Morphological relationship among the *Momordica* species of Indian occurrence. *Indian Journal of Genetics, 73*, 278–286.
68. Bharathi, L. K., Parida, S. K., Munshi, A. D., Behera, T. K., Raman, K. V., & Mohapatra, T. (2012). Molecular diversity and phonetic relationship of *Momordica* spp. of Indian occurrence. *Genetic Resources and Crop Evolution, 59*, 937–948.