Turkish Journal of Field Crops, 2011, 16(2): 172-178

NEW NON-REDUNDANT MICROSATELLITE AND CAPS-MICROSATELLITE MARKERS FOR COTTON (*GOSSYPIUM* L.)

Mehmet KARACA* Ayşe Gül İNCE

[1]*Akdeniz University, Faculty of Agriculture, Department of Field Crops, Turkey*
**Corresponding author's email: mkaraca@akdeniz.edu.tr*

Received:11.07.2011

ABSTRACT

Over a decade, researchers have developed microsatellite primer pairs for cotton (*Gossypium* L.) and most of them have been deposited in the cotton microsatellite database. However, results of the present study clearly indicated that a considerable amount of cotton microsatellite markers were redundant, in some cases, did not contain microsatellite domains and had low level of polymorphisms. In this study, a new set of 144 non-redundant microsatellite primer pairs were developed using expressed sequence tags (ESTs) and tested. Among these primer pairs, seventy were polymorphic while seventy-four were monomorphic. In the present study, suitable restriction enzymes (REs) useful in CAPS-microsatellite analyses were also determined. Results showed that these REs were suitable in the conversion of monomorphic microsatellite markers to polymorphic markers. Non-redundant microsatellite primer pairs and restriction enzymes for CAPS-microsatellite technique could be useful in detecting, manipulating and identifying genes associated with desirable agronomic and quality traits within cotton breeding programs.

Key Words: non-redundant microsatellites, restriction enzymes for ESTs, touch-down-PCR

INTRODUCTION

Four species of cotton (*Gossypium* L.) are widely grown to produce both natural textile fibers and cotton seed oil. Although traditional cotton breeding programs have produced steady improvement in a number of agronomic traits, the lack of useful economic characters in cotton still remains a major challenge. One hundred and fourty five morphological markers have been identified in cultivated cotton. However, utiliy of these markers in cotton breeding programs has remained limited due to their deleterious effect and the difficulties in accumulating several morphologic markers in a single genotype (Karaca et al., 2004). Although isozymes have been used in genetic studies as an alternative to morphological markers, their expression is often restricted to a specific developmental stage of tissues (Mellon and Triplett, 2006).

Restriction fragment length polymorphism (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP) and simple sequence repeats (SSR) or microsatellites are the most used DNA marker techniques in plant species (Zhang et al., 2002; Lascape et al., 2003; Ince et al., 2010a). Microsatellites are tandemly repeated DNA sequence motifs that usually consist of two to six nucleotide core units. They are highly abundant in eukaryotic genomes but also occur in prokaryotes at lower frequencies. The allelic variations in microsatellites are determined using polymerase chain reaction (PCR). The DNA sequences flanking the microsatellites (microsatellite domains) are generally conserved and primer pairs complementary to the flanking regions are used to amplify microsatellite markers (Ince et al., 2010b). The length of microsatellite containing fragments varies according to the number of repeated residues in microsatellite domains. Microsatellite flanking regions other than primer annealing sites may also differ among individuals and these differences can be detected using a new approach called CAPS-microsatellite marker technique (Ince et al., 2010c).

Microsatellites are considered as the marker of choice for self-pollinated crop species with little intraspecific polymorphism (Roder et al., 1998). The reproducibility of microsatellite markers is very high. Therefore, they can be efficiently used by different laboratories to produce consensus data, which makes them useful for genome mapping and genetic studies. One of the drawbacks of the conventional microsatellite technique is the availability of the primer pairs. However, development of microsatellite markers has dramatically increased with the use of expressed sequence tags (ESTs) in cotton (Saha et al., 2003).

In 2006 the Cotton Microsatellite Database (CMD) was established representing about 5,484 microsatellite loci (Blenda et al., 2006) and later on the CMD was renewed to Cotton Marker Database. CMD now contains more than 10,000 publicly available microsatellites (http://www.cottonmarker.org). However, our initial studies clearly revealed that a considerable amount of these microsatellite loci are redundant and have low level of polymorphisms (Saha et al., 2003). Therefore, the present study was conducted to identify new primer pairs and restriction enzymes suitable for CAPS-microsatellite markers using EST databases.

MATERIALS AND METHODS

A total of 11,514 microsatellite sequences, deposited in CMD database (http://www.cottonmarker.org) consisting of 379 BNL, 392 CIR, 53 CM, 200 DPL, 700 GH, 205 HAU, 309 JESPR, 84 MGHES, 2513 MON, 617 MUCS, 1316 MUSB, 554 MUSS, 3,250 NAU, 192 STV and 750 TMB, were reanalyzed to identify microsatellites using the TRA 1.5 software (Bilgen et al., 2004) and identify redundant microsatellite sequences using the Sequencher software set to minimum overlap of 100 bases and 90% identity match.

A total of 375,776 *Gossypium* expressed sequence tags (ESTs) consisting of 268,779 *G. hirsutum* L., 63,577 *G. raimondii* Ulbr., 41,768 *G. arboreum* L., and 1,023 *G. barbadense* L. were analyzed using the TRA 1.5 software to identify microsatellites; consisting of motif lengths 2 to 10 nucleotides using the selection criteria described in Karaca et al. (2005a). Primer pairs flanking the microsatellite domains were designed using PRIMER3 software according to Ince et al. (2010b).

Leaves of Texas Marker-1 (TM-1), Aydin 110, Nazilli 87 (*Gossypium hirsutum* L.), Pima 3-79 (*G. barbadense* L.) and a total of 20 F_2 individuals obtained from a cross between Aydin 110 and Nazilli 87 were used for DNA extractions. Genomic DNAs were extracted using a DNA extraction method described in Karaca et al. (2005b).

Touch-down polymerase chain reactions (Td-PCRs) were carried out in 25 µl reaction volume containing 120 nanograms genomic DNA as template, 0.5 µM of each microsatellite primer pair listed in Table 1, 80 mM Tris–HCl (pH 8.8), 19 mM $(NH_4)_2SO_4$, 0.009% Tween-20 (w/v), 0.28 mM of each dNTP, 2 or 3 mM $MgCl_2$, and 2 units of *Taq* DNA polymerase (Bioron or Fermentas).

Td-PCR amplification profile used in the present study was as follows: initial denaturation at 94°C for 3 min, 10 cycles with denaturation at 94°C for 30 s, annealing at 60°C, or 66°C (Table 1, Profile A and B, respectively) for 30 s in the first cycle, diminishing by 0.5°C each cycle, and extension at 72°C for 1 min in a 96-well P92 thermal cycler (Thermo Hybaid). An additional 30 PCR cycles were run using the same cycling parameters mentioned above with constant annealing at 55°C or 61°C. Denaturation and extension conditions were the same as the ones indicated above. The amplifications were finished with a final extension at 72°C for 7 min. Amplicons were separated using high resolution agarose gel (Serva) electrophoresis according the procedures described in Ince et al. (2010d).

In order to identify restriction enzymes (RE) suitable for CAPS-microsatellite technique a total of 280 RE recognition sites were mined using 7,396,371 base pair of cotton microsatellite sequences deposited in CMD. Number of restriction enzymes and frequency of their cutting sites were identified using Sequencher software. CAPS-microsatellite technique was applied according to protocol described in Ince et al. (2010c).

RESULTS AND DISCUSSION

Analyses of 11,514 microsatellite sequences in the CMD database indicated that 5,705 sequences contained microsatellites. The most common microsatellites in the CMD database are di-nucleotide repeats (2,934) followed with tri-nucleotide repeats (2,022). Among the tetra- (512), penta- (96) and hexa-nucleotide (141) repeats, hexa-nucleotide repeats were more abundant in cotton microsatellites. Analyses of the present study clearly indicated that a considerable amount of public microsatellite primer pairs and microsatellite sequences deposited in the CMD were redundant (49.55%). The use of redundant microsatellite markers will definitely decrease the efficiency of microsatellite markers in cotton genomic studies. Researchers should be aware of the redundancy in CMD prior to selection of public microsatellite markers in order to increase the success of the use of cotton microsatellites.

The physical size of the cotton genome is relatively larger than many other important crop species. Cotton genome size varies from 2.246 to 2.702 megabases depending on the species (Arumuganathan and Earle, 1991). If we assume that the average physical size of a centiMorgan in cotton is about 400 kilobases (Lascape et al., 2003), then 5,615-6,755 polymorphic markers are required to completely cover the cotton genome. Therefore, a new set of microsatellites are still required in cotton genetic studies.

In the present study using 268,779 ESTs, a total of 173 EST-based microsatellite primer pairs were developed and named MK primer pairs. Twenty-four MK primer pairs failed to amplify cotton genomic DNAs although they could amplify cDNA samples of TM-1 (Ince and Karaca, 2009). A list of 144 MK primer pairs, which could successfully amplified genomic DNAs, and selected restriction enzymes suitable for CAPS-microsatellite analyses are listed in Table 1.

Table 1. MK microsatellite primer pairs, marker sizes and restriction enzymes

Primer ID	Forward Primer 5'➔3'	Reverse Primer 5'➔3'	Motif	G[1]	Size bp[2]	I[3]	RE[4]
MK001	AGAGGGGGCACAGAGAAAAC	CCGAGTGAACAGGCAAATCT	$[TC]_{15}$	A	337	M	
MK004	TCACCGCAGAAAACCACTC	TGTGTCAAGTCATAGCCGTTG	$[ATG]_{12}$	B	500; 436	M	*Rsa* I
MK006	CAGAGGCAGAAAAGCAAACC	AGGGAGTAGGCAAAGAAATCG	$[TGA]_{10}$	B	346; 331	P	
MK007	TTCTCCCTTTCAGCGTTTAGG	AAGCAACACCAACACACCAA	$[ATT]_{13}$	B	344	P	
MK008	AAGTGGTGGAGGAGGAGGAG	CCTTGGCTGGTGGTTTTG	$[GGT]_{6}$	B	1,174	M	*Aci* I
MK009	GTGGAGAAGTGCGATGAGG	CTCGGTTTGGGAGTGGTAAC	$[CACC]_{4}$	A	383	M	
MK011	CCTCCTCGTTTCTTCACTGC	CTTGTTCCATTTCACCCAAAG	$[TCT]_{12}$	B	1,000	M	*Taq* I
MK013	ACGCTGAAATCCAAAACGAC	AAGGCAGAACGAAGGTTGC	$[TATG]_{15}$	B	304	M	
MK015	TTTGTTTCCCTTTGCTTGGA	GGTGCCTGGTTCTTGTGTCT	$[AT]_{15}$	B	340	M	
MK016	CCGCACACACTTCTCTCTCTC	CTACGACTTGCTCCGTGGTT	$[AC]_{15}$	B	850; 719	M	*Bs* II
MK017	GCCCTTCAATCCATCATTCA	CCTTTCTCCGCCTCAACAG	$[ATC]_{15}$	B	367; 340	P	
MK019	CACCTCTCCACCCATCTCC	CCTCCCCATTCGTTTTCTTT	$[AAAG]_{9}$	B	1,142; 348	P	
MK020	ACCAGTTTCCCAGTTGTTGTTT	TGTCTCCCCTGCTCTGTTTT	$[AT]_{20}$	B	145	P	
MK021	CCATTTTTCTGCCACTACCC	TGCCAATCCCCTATTTCTTG	$[AAG]_{12}$	B	311	M	

TABLE 1 continued

MK022	GCTGGTTGAAGGAAATGCTG	GACTTGCTCTGGTCTGCTTTG	$[GAGCGG]_5$	A	218	M	
MK023	CCTTCCCATTTCTCTTTCTCC	GCCTTTTGTCCCTGGTTTTA	$[AGAGA]_6$	A	201	P	
MK024	TTCCACAGGCATCAAATCAA	TTTTCCAGCAATCCGAGAAC	$[ATGTAT]_{11}$	A	197	M	
MK025	TCCGCATTTCTTTCTCTTTCTC	GGAGGCAATCTAAGCACACC	$[TTCTTG]_5$	A	251	M	
MK026	TTACAAACACATCCATCACACG	TGCGTCTCCCTTCCATTT	$[TA]_{16}$	A	144	M	
MK027	TCCATCTCATCCTGCTCTCC	GTTCACGCTCCCACTTTCAG	$[CTCATT]_6$	A	254	P	
MK028	GAACATACGAACGAACGAAACA	GCAGGTGGAGAACTGGGTTA	$[AT]_{16}$	A	218	M	
MK029	GATGGAAAGGGCAGTGTCAT	GGGTTTTACGGTGGCATTAG	$[TA]_{19}$	A	188; 171	P	
MK030	GCAGGAGAACTGATGGAAAAA	GATGGTGAAAGGATGGAAGC	$[TCA]_{11}$	A	178	M	
MK031	AAAACCCCCTTGTCAGTGG	GGCGATGATGGAAGAAGAAG	$[TCT]_{11}$	A	592; 394; 296	M	
MK033	CCACCCTGATTACTGAACAATG	GCAAGAGATGAAAATGCCAAC	$[AT]_{15}$	B	204	P	
MK035	GGCGACCTACCTTCCACTC	GCTTGATTTATTGGGGGATG	$[CAC]_{11}$	A	183; 157	M	
MK037	CTTGGAAAAAGGAAGAGCAGAA	TTGGCTGGAAGTGATGAAGA	$[ATAC]_{12}$	A	259	P	
MK038	AGAAGAAAGAAAGGAAACCCTACG	GGCTTGGAGCAAACACAGAC	$[CATA]_{18}$	A	220	M	
MK039	TTGGGGTGTTGACTTTGGTT	GAGGGCAAGCGTTCTCATC	$[ATGCCC]_6$	A	156; 146; 129	M	
MK040	TACGCAAAACCACTCCAACA	CCAGAAAAGTAGCGGGATGA	$[TAG]_{10}$	B	198	M	
MK041	TTCGCATTTCTATTCTCCTCCT	CCCCTCTCTCTCTCCCTCTC	$[CTT]_{12}$	A	229	P	
MK042	TAAGACGATGCGAGGGTCA	GCTTGGGAGAGTCAGAAGGAG	$[ATGTGA]_6$	A	358; 324	P	
MK043	AACAGGGTTTGGAGCAGTGA	GTTGACCGCAAGATGGAGAT	$[TGGTGA]_7$	A	739	P	
MK044	CACTTGGGGATTGGTTCAA	CTGCTGTTGTTGGGTCTGAG	$[GAA]_{10}$	A	240	P	
MK045	ATCAGCGAAGGTAGCCAATG	CAGCCACACCGTATTCTGG	$[TATAT]_6$	A	290; 265	P	
MK046	GTGTTCCATTTCTCCCAAGG	AGGGCAGTCATAAGTTGAGGTC	$[TCA]_{11}$	A	229	P	
MK047	GGTGTCATCGTCTCCATTTTG	AGCAGCGGTTCCTTCTTTCT	$[GAA]_7$	A	285; 265	M	*Dde* I
MK048	TTTGGGCTTTCTTTTCTCTCTC	AGACTTTGTGTCCCGCTCA	$[CT]_{17}$	B	1,500	M	*Nla* III
MK048	TTTGGGCTTTCTTTTCTCTCTC	AGACTTTGTGTCCCGCTCA	$[CT]_{17}$	A	1,500	P	
MK049	GTGTGTGTGCCCTGTGAGAT	CACTGCCCCTAAGAAGTTGC	$[AG]_{21}$	A	340	M	
MK050	CGTCCGCTTCAATCTGTTTT	GGTTCCTGGCTGGTCTTCTT	$[CCAAC]_6$	A	215	P	
MK051	CCTGATACACCCCGCAATAG	CGTCAAAAAGGAAGGAGCAA	$[TGCTCC]_5$	A	252	P	
MK052	ACTAAAAGGTTCTCCCACGGTA	TGATGTTGGCTTCAGAGGTG	$[ATA]_{14}$	A	274	P	
MK053	CTCACCGTTACTCGCCTTG	CGGATTTTCCAACCCTGTT	$[CAT]_{13}$	A	164; 138	P	
MK054	GCTGACACGAAAGCACTCC	CGCCTTGGAAACTCTACCC	$[GAT]_{13}$	A	252	P	
MK055	CGACCACACCCCTAAAGTAACA	TCTTCCTCAGCACTCCAAGG	$[AG]_{19}$	A	243	M	*Dde* I
MK056	GGGAAGAAGAGGAAGCAAGG	TAAGTTCGGGGTTCGGTTG	$[ATG]_{10}$	A	200	M	
MK057	TAAGTTCGGGTTCGGTTGG	CACTGGGAGAAGAGGAAGCA	$[TCA]_{10}$	A	203	M	
MK058	CCCCCTCAACCAAATGTAAA	TCCTACTGTTTTTCCCCATAGC	$[ACATAG]_6$	A	289; 229; 93	M	
MK059	TGATGATGTTGGGTCTTTTGG	GAAGGAGATGGGGAAGTGGT	$[ACCCCA]_8$	B	248; 204	M	
MK060	GGGTGAGGGGGTAAGACAAT	AGAAGAAAGAAGCACAGATGACG	$[CTG]_{10}$	B	242	P	
MK061	CTCCATCTCTCTGCCCTCAA	CGAACTCAGGCTCAAACCAG	$[CCTCTT]_5$	B	188; 159	P	
MK062	GGCTTCTTTCTTGCTGCTGT	GTTTCATCCCAGACCCAATC	$[TCT]_{16}$	B	282; 257; 226	P	
MK064	TCAGACCAAACCCTTCCTTTC	GCCAGTTGATTCCCTCTTGA	$[CAGCAC]_5$	B	335	P	
MK065	CCCCACTCACTCCCTCTTC	GCAGGTCCGTAGCAGTTGA	$[CCACCG]_6$	B	299; 263	P	
MK066	GCAGTAGCCGATGGTGATG	CGATGCTTTCTTCCTTTGTCTC	$[CAGCAA]_6$	B	186; 170	M	*Aci* I
MK067	AGAAAAACAGCGGCAACAAC	CTTCGTCGTCCTCGTCATCT	$[GAG]_{11}$	A	2,000	M	*Hinf* I
MK068	TAGATTTGTTGCGGGTGCT	TCTTCCTCTTCGCCTCCAT	$[GA]_{24}$	B	154	M	*Taq* I
MK069	GATTGGGGCATTCAAGGAG	CGCTGGAGGAGGGTGTAAC	$[ACCCCA]_5$	B	248; 231	M	*Mae* III
MK070	GAGACGGTGGTGATGATGG	CCTTGTCAGTGTCCGAGTTG	$[AAG]_{16}$	A	2,000	M	*Mbo* II
MK071	CTGGAAGGAGCAGACACAGAG	TGAATCCCAAAGCACGGTAT	$[GA]_{21}$	A	159; 122	M	
MK072	CCTTCAAATCCCTTCGTCCT	GCCGACATTGCCTTCTTTT	$[CCGCCA]_5$	B	188; 162	P	
MK073	GAGAGCAAAAGCACGAGACC	CAGGGGTGACATAGGGTTCA	$[TCAGGC]_5$	B	182	P	
MK074	CCTTCCACTACTCCTCCCCTTA	GATTGAACCTGCGGCTGA	$[CAACGC]_5$	B	220	M	
MK075	CCCCTTTTGCCTTTATTGCT	CACACACCACCTCCTTCCTT	$[TC]_{15}$	A	205	P	
MK077	TCAACACCCCATCCACTATG	TCACACACGAAGTAAATCTCAGC	$[TTA]_{11}$	B	229	P	
MK078	GGAGAGTGGGAAAAGAGCAG	AGAGAAAGGAGCGGACCAG	$[GAT]_{11}$	B	230; 220; 210	P	
MK081	CAGGTCTAAGATGTACCACAAGC	GAAAACAACTCTCTTCCCCATA	$[CAG]_8$	B	318	M	
MK082	CTCACCGTTACTCGCCTTG	CGGATTTTCCAACCCTGTT	$[CAT]_{13}$	A	164; 155	P	
MK083	GTCCACTAAGCAGCAAAACTTG	TCCATTCTCCAAGTACCTAGTATCG	$[TCAGGC]_7$	B	155; 136	M	
MK084	TTGCAGCCCTTCTAGTCCTC	GGGTCAAAGGAAGTGAAACC	$[CCA]_5$	B	224; 212	P	
MK085	CAAACTCCCATCATCAGCAA	TCTTCTGCTGGTGCAATGAC	$[CAG]_{12}$	A	92; 82	P	
MK086	CCACCAGTTTGGTAGGTATGAAC	TCAACAGTGCAAGGACTTCATC	$[CAT]_8$	B	124; 100	M	
MK087	TGAAATCCCAGAGCTTTTTCTC	TGTTCTCATCGACCAATCCA	$[CT]_{15}$	B	258; 249	P	
MK088	TTTCTTTGGTGGTGAAACTGG	AGACCTTCGTTCCCCAGAC	$[CT]_{14}$	B	139; 111	P	
MK089	CAATGTCCATCGCCACCT	GGGCGGAGAAGGTTAAAAAC	$[CCA]_7$	B	189; 176	P	
MK090	CCTCCGAGGATTATGTTGGA	CGAATTCACAAATCCTTCACC	$[GTTT]_6$	B	232	M	
MK091	AGTCAGGCTCAGGCTCAGG	GCCTTGCTTCTCGCCATT	$[TCAGGC]_5$	B	298; 280	P	
MK092	ATTGATGCCAACTGAACGAG	ATTGACGGGAAAGTGAAACC	$[ACAT]_{10}$	B	244; 228	M	*Rsa* I
MK094	ACACACAGCATCCCATTCC	AAGACCAAAGGCAAGGACAC	$[ATAC]_8$	B	432	M	*Nla* III
MK095	AAACTGCAAACCCCACACTC	GGAGAGGCTATTCAGGGAGA	$[CTT]_5$	B	317	P	
MK096	CAGCTGGCCTTCTCCGTA	AGGAGAAAGGGTCTCCGATG	$[CATA]_5$	B	271; 257	P	
MK097	TCGTGGGGAAGTACAGTGTG	GGAGGCGTTTCAAATGAGAG	$[GAAGGA]_4$	B	612	M	*Alw* I
MK098	TCACAAGAGGCTTTCAATGCT	TTACACCTCCAGGCATCAAA	$[ATTT]_5$	B	212	M	
MK099	AGATTTGGGAATTAGCCAATTAGA	GCTTCCAAAATTGAGAAAGCTC	$[AC]_{11}$	B	254; 215	P	
MK100	GACTCTTGGACTGAATATTGACAT	GAGTCACAAAAGCCCAAATT	$[TAAA]_9$	B	1,479	P	
MK101	TCATCATCATCCTCGTCTTGA	TTATGGCCCAATCCTCTCAC	$[CAT]_{15}$	B	246	P	
MK102	CATCCCCACCCACATCTC	TCCACTTGATGGGGTAGGAC	$[ACCCCA]_8$	B	229; 196	M	
MK103	TGGGGAAGGACAAGCATAAG	ACCAGGATAGCCAGATGGTG	$[GT]_{13}$	B	236	P	
MK104	GTTTTGGATTGGGAAGGTCA	TCACACTCTGTCTCTGGCTCTT	$[AGCG]_5$	B	634	M	*Apo* I
MK105	CAAAGATGCCGAAAGAGAGG	GTAAGATCGGCGGGTCATC	$[CCG]_{12}$	B	177; 158; 139	P	
MK106	CACAAACGCTGTCACTACGAA	CGGGCAAGTTTAGGTCAAAG	$[CAC]_{14}$	B	195; 178	M	
MK107	AAGAAAGGCAAGCGTTCAAA	CGATGGTCATCGTTTCCAC	$[CTTCCG]_{12}$	B	244;212	P	
MK108	GGATTCTGGTTATGTATTGC	GTGTTTGATGATGGTGATG	$[CAG]_{19}$	B	186; 165	P	
MK109	TCCGATTCCCCCTTTCTAAC	CTGGAACTCGCCACCATAAT	$[CTT]_{10}$	B	194; 158	M	
MK110	ATGGGGGAGGAAGAGAAAAA	CGACGAACTTTACGAGCACA	$[CAA]_{12}$	B	300; 279	P	
MK111	ACGTGGACGAAGAACAGCA	GTGATTCTCGACCGCTTTTT	$[AGG]_{10}$	B	279; 262	M	
MK112	TGGAAGCTTCTCTTCTCTTTTGA	GCTTACTCAATCAAACAAATCCAA	$[ATTTT]_4$	B	374	P	
MK113	CAAATGGCTGCTACCTTCC	CTTCTTCCAAACCCTCTCGT	$[ATC]_{12}$	B	764	P	
MK114	ATGGTCATTCCGATGCTGTT	CCAATGGTCCCTACATGACC	$[CTG]_5$	B	228	P	
MK115	CCCGAGTTTCTCTATTTCTCAGG	TCCGCGAAGCTTCAACTAAC	$[CAT]_{11}$	A	828	M	*Apo* I

TABLE 1 continued

MK116	CTCGGGAAATCAATTCGTTG	ACAAATTCCCATTAAGCAAACC	$[CAA]_{11}$	B	208; 201; 191	P	
MK117	ACTCCAAAAACCCCCTCAAC	GGCGACTGCAAAGGAAATAC	$[CAA]_{7}$	B	268; 252	M	
MK118	TGGTAGGTAGAGCTTGTGGTTG	GGTTCAAGGTTTTGGATTCTTG	$[CT]_{14}[GT]_{11}$	B	211	M	
MK121	TATCCACCACCCAAGTCACC	GGCTCCTTTGTGGTTTAGGA	$[CCGCAT]_{6}$	B	235; 224; 214	M	
MK122	GCTTTGCCTCCTCATTACCA	TTGGGTGGTGGATAAACTGG	$[CCGCAT]_{5}$	B	196; 184	M	
MK123	TCTTCCTTCCCAACACAACA	TTGGATGAACAAAACAGAGAGAA	$[GGC]_{10}$	B	272	M	
MK124	AAAAACCCTTAAATCTCGTAAACA	CGATGTCGAGCTACCTTTCT	$[TA]_{8}$	B	213	M	
MK125	TGAGGGAAAGCAATACGACA	GCGATGGTTGTGATGAAGAA	$[GAA]_{10}$	B	300	M	
MK127	GTTTAGATTATGGCATTACATAT	AAGGCTATATCTGACTTTGG	$[TA]_{18}$	B	172	M	
MK128	TCAAGGACTACCAGCAGCAG	CATTTGCACGCTGTTGATTC	$[CAG]_{9}$	B	238; 221	M	
MK129	GCTGATGCTGATTCCTCCAT	TGCCCTTCATCTCGTTTCTT	$[CAA]_{8}$	B	241; 234	M	*Hinf* I
MK130	CTCACGGCTATCCACCATCT	ACCATGAGCACCGTGAGAC	$[TG]_{14}$	B	243; 234; 221	P	
MK131	CCATGCAAATCCATGCTAGA	TTTCTTTGGTGGTGAAACTGG	$[TCAGGC]_{6}$	A	245; 240; 230	P	
MK132	AGCAAGGCATGAGCGATACT	GGTGGTACCTTCCCATGTTG	$[TCAGCC]_{6}$	B	192	M	
MK133	GGCCGTTCAAGCTCAGTATC	TATAAACCCCTCCCCCTTGT	$[TCT]_{5}$	B	242	P	
MK134	AAGCTAAGCCAAGAGCACCA	TTCCGAGAAGGAATCCTCAA	$[GAA]_{6}$	B	199	M	
MK135	CAACCAGACTTGCAGAGCAG	GATTCTTGCTTCCGTCCAAA	$[GAT]_{8}$	B	179	M	
MK136	AACATGTTTTCTCCGATCT	CCGGGATACTCCTATATCTT	$[TC]_{12}$	B	1,786	M	*Mbo* II
MK137	AAAATTGCCAAACAAAAGCTATG	CAATCGAACCACACCAAAAA	$[CTT]_{6}$	B	262; 233	M	
MK138	GGGTTAGCAGAAAAGGAGGAA	TGAAACTCCCACAAGGAAGC	$[TCAGCC]_{3}$	A	268	P	
MK139	GACCAAACCCTTCCTTTCAA	AGATTGTGGTAGCCCCAGTG	$[CAGCAC]_{7}$	B	286; 256	P	
MK140	GGAAGAGGAAGGGGGAAA	CTGAGAGCAAAAACAAACCATC	$[GAGTTG]_{7}$	B	185; 163	M	
MK141	CCATTTCTACTTTCCATCAGATCC	TCCTCTCCATTTTCGTTTGG	$[TC]_{18}$	B	181; 164	M	
MK143	CACAAAACCAATCACCACCA	CAAGGGAGAACTCGGAGAAA	$[GA]_{16}$	B	239; 225; 211	P	
MK144	TCTTTGGCAATGGAAAACAA	CATGAAGCGAGACAAGACCA	$[TTC]_{7}$	B	257; 252	M	
MK145	TTAGCTCACCGGCCTGAG	TCAGGATCACCCTTTTCACC	$[GTAGTGAGA]_{2}$	B	240	P	
MK146	ATGGAGGCTGCAAAGACTGT	CCACTCCGACTAAAAGATCAGC	$[GTAGTGAGA]_{3}$	B	184	M	
MK147	TCGTCTTCTTCTTCGTCTTCG	TCAGCCGCACATTAGGTAAA	$[GA]_{7}$	B	180	P	
MK154	TGTGGATATGGAGGACTTTG	GCTCTTCCATCTCACCATC	$[TGA]_{10}$	B	235	P	
MK156	CACTCATTCTTTGTATCCATGTATTG	ACATGTTTCTGAGGCCAACC	$[TTA]_{11}[TAG]_{7}$	B	208; 196	P	
MK157	AACCCAAGGAATCGGAGAAG	TTCGCCACCTCTCTAGGTACA	$[CTT]_{5}$	B	242	M	
MK158	CTTCCAGTTCCACCATAGCC	ACCAAATCCAGGTTCCACAG	$[AC]_{14}$	B	1,473	M	*Nla* III
MK161	ATCTCTTCCACCCCTTCCAC	ATCCACCCCCATTGTTCTTT	$[TC]_{24}$	B	292; 261	M	
MK163	GTTGAAAATGGCGTGCTTG	GACTCGTGCGGTGAAAATG	$[AT]_{18}$	B	300	P	
MK165	GTTTCTCGTCTCTCCGATTT	ACCCACCTTTCAACAACG	$[AAG]_{10}$	B	415	M	*Hinf* I
MK166	AAAACGAAGTTGGGAACAGT	GAGGGACAGCCTAATAATGG	$[ATA]_{20}$	B	215	M	
MK167	GCTATGGAGATGCGAAGCA	TTGATGAGGGTCTGGGATTG	$[ATG]_{12}$	B	356	P	
MK168	CAGGAGGGACAACAAGAACA	GCGAAGATTGGGAACAAGA	$[CTT]_{13}$	B	158	P	
MK169	GCCGCCAGTGTTGTATGC	CGAGGAATGAAAGCGAGAAAG	$[TCT]_{15}$	B	244	P	
MK170	ATCCGCCCACAAATAAAGC	CATCGTGAGAGAAAGTGAAGGA	$[TTC]_{14}$	B	226	P	
MK171	ATTACAGTGGATGTTCCTTG	CTTATGGGATGATGAAAGAC	$[ATGT]_{9}$	B	345; 327	P	
MK172	ATAGGGAAAAGGTAGGGATT	ACAAATGAACCAAGACGAG	$[CATA]_{9}$	B	331	P	
MK173	GGGGTCCACAGATACAGG	GTCCAAAACTTGTCCCATTAG	$[TATG]_{9}$	B	809	M	

[1:] Touch-down PCR annealing temperature (°C); A: amplification starts at 60°C annealing, B: amplification starts at 66°C annealing, [2:] Size of PCR amplified products of Texas Marker-1 TM-1, *Gossypium hirsutum* L. in base pairs; [3:] M: monomorphic, P: polymorphic among TM 1, Pima 3-79, Aydin 110 and Nazilli 8; [4:] RE: restriction enzymes used to digest PCR amplified products. List of RE usable in cotton CAPS-microsatellite analyses are also shown in Table 2.

Among MK primer pairs (Table 1) developed in this study, seventy produced polymorphic markers based on the genomic DNA analyses of TM-1 and Pima 3-79. Figure 1 shows amplification profiles of 36 MK primer pairs between TM-1 and Pima 3-79.

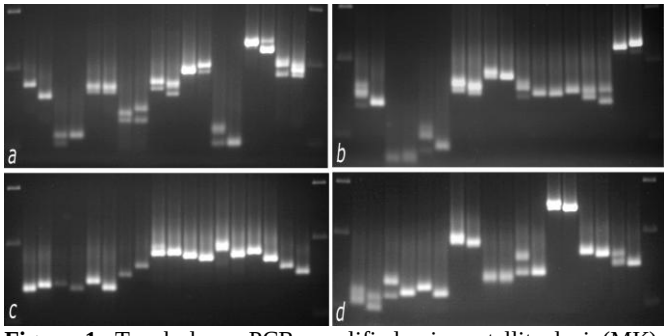

Figure 1. Touch-down PCR amplified microsatellite loci (MK). Panels a, b, c and d contain DNA size markers ranging from 150-500 bp at both sides. Panel a shows amplified products of MK139 MK143, MK145, MK156, MK027, MK037,MK053, MK064 and MK065. Panel b shows amplified products of MK007, MK019, MK033, MK041, MK045, MK051, MK52, MK054 and MK060. Panel c shows amplified products of MK082, MK085, MK088, MK092, MK112, MK131, MK154, MK170 and MK172. Panel d shows amplified products of MK061, MK072, MK073, MK091, MK100, MK108, MK113, MK130 and MK138. All the MK loci in figure above are amplified products of TM-1 and Pima 3-79 in the respective order.

Among 144 MK primer pairs 74 produced monomorphic amplicons whose size ranged from 200 bp to 2,000 bp. Several larger amplified products of EST-based microsatellites were detected in the present study as they were also reported in other plant species including the cotton (Ince et al., 2010c,d). Analayes of Ince et al. (2010b) showed that these larger amplified products obtained using the EST databes have intron(s) between the primer flanking regions, resulting in very big products. Previous studies of different research groups found that the occurrence of larger amplified products were between 10% and 50% of EST-microsatellite primer pairs depending on the species (Cristancho and Escobar 2008; Minamiyama et al., 2006; Wang et al., 2008). In many cases these larger products obtained using EST-based microsatellite primer pairs are not polymorphic and not sutable in genetic studies (Ince et al., 2010c).

In the present study research was conducted to determine restriction enzymes (REs) suitable in CAPS-microsatellite studies.DNA sequences of the CMD was mined to identify REs and a total of 27 REs were determined. Table 2 shows REs determined and the number of recognition sites in microsatellite containing sequences.

Table 2. List of restriction enzymes useable in cotton CAPS-microsatellite analyses

Restriction Enzymes	Recognition Sequence	Number of Recognition Sites
Aci I	C↓CGC/ GGC↑G	21,094
Alw I	GGATC(N)$_4$↓ CCTAG(N)$_5$↑	11,128
Apo I	R↓AATTY YTTAA↑R	19,761
Bsl I	CC(N)$_5$↓(N)$_2$GG GG(N)$_2$↑(N)$_5$CC	14,236
Bst7 II	GCAGC(N)$_8$↓ CGTCG(N)$_{12}$↑	15,731
*Cac*8 I	GC(N)↓(N)GC CG(N)↑(N)CG	13,237
Dde I	C↓T(N)AG GA(N)T↑C	19,477
*Fnu*4*H* I	GC↓(N)GC CG(N)↑CG	22,702
Hae III	GG↓CC CC↑GG	12,730
Hinf I	G↓A(N)TC CT(N)A↑G	22,574
Hph I	GGTGA(N)$_8$↓ CCACT(N)$_7$↑	14,394
Mae I	C↓TAG GAT↑C	14,865
Mae III	↓GT(N)AC CA(N)TG↑	16,243
Mbo I	↓GATC CTAG↑	22,717
Mbo II	GAAGA(N)$_8$↓ CTTCT(N)$_7$↑	17,262
Msp I	C↓CGG GGC↑C	11,086
Mwo I	GC(N)$_5$↑(N)$_2$GC CG(N)$_2$↓(N)$_5$CG	18,876
Nla III	CATG↓ ↑GTAC	30,621
Nla IV	GG(N)↓(N)CC CC(N)↑(N)GG	14,508
Rsa I	GT↓AC CA↑TG	11,814
*Sau*96 I	G↓G(N)CC CC(N)G↑G	11,665
ScrF I	CC↓(N)GG GG(N)↑CC	12,562
Sec I	C↓C(N)$_2$GG GG(N)$_2$C↑C	14,994
SfaN I	GCATC(N)$_5$↓ CGTAG(N)$_9$↑	11,507
Taq I	T↓CGA AGC↑T	19,114
Tfi I	G↓AWTC CTWA↑G	14,698
*Tru*9 I	T↓TAA AAT↑T	16,979

R: A or G; W: A or T; Y: C or T; N: any of the four bases

A total of 30 MK primer pairs selected from 74 MK primer pairs producing monomorphic markers were further analyzed using a strategy called CAPS-microsatellite analysis (Ince et al., 2010b) utilizing *Mbo* I, *Rsa* I, *Apo* I, *Mae* I, *Mbo* II, *Sau*96 I, *Mae* III, *NIa* III, *Dde* I, *Mwo* I, *NIa* IV, *Taq* I, *Tru*9 I and *AIw* I restriction enzymes selected from Table 2. Results clearly indicated that these restriction enzymes were useful in CAPS-microsatellite technique, converting 21 monomorphic microsatellite markers into polymorphic markers (Figure 2).

In order to investigate the application of CAPS-microsatellite technique, genomic DNAs of Aydin 110, Nazilli 87 and several F_2 individual derived from a cross between Aydin 110 and Nazilli 87 were amplified using several MK primer pairs which produced monomorphic markers ranging in size 200-2,000 bp. Larger monomorphic markers were converted into polymorphic markers which segregated in co-dominant Mendelian fashion and in some cases single enzyme digestion of larger amplicons produced

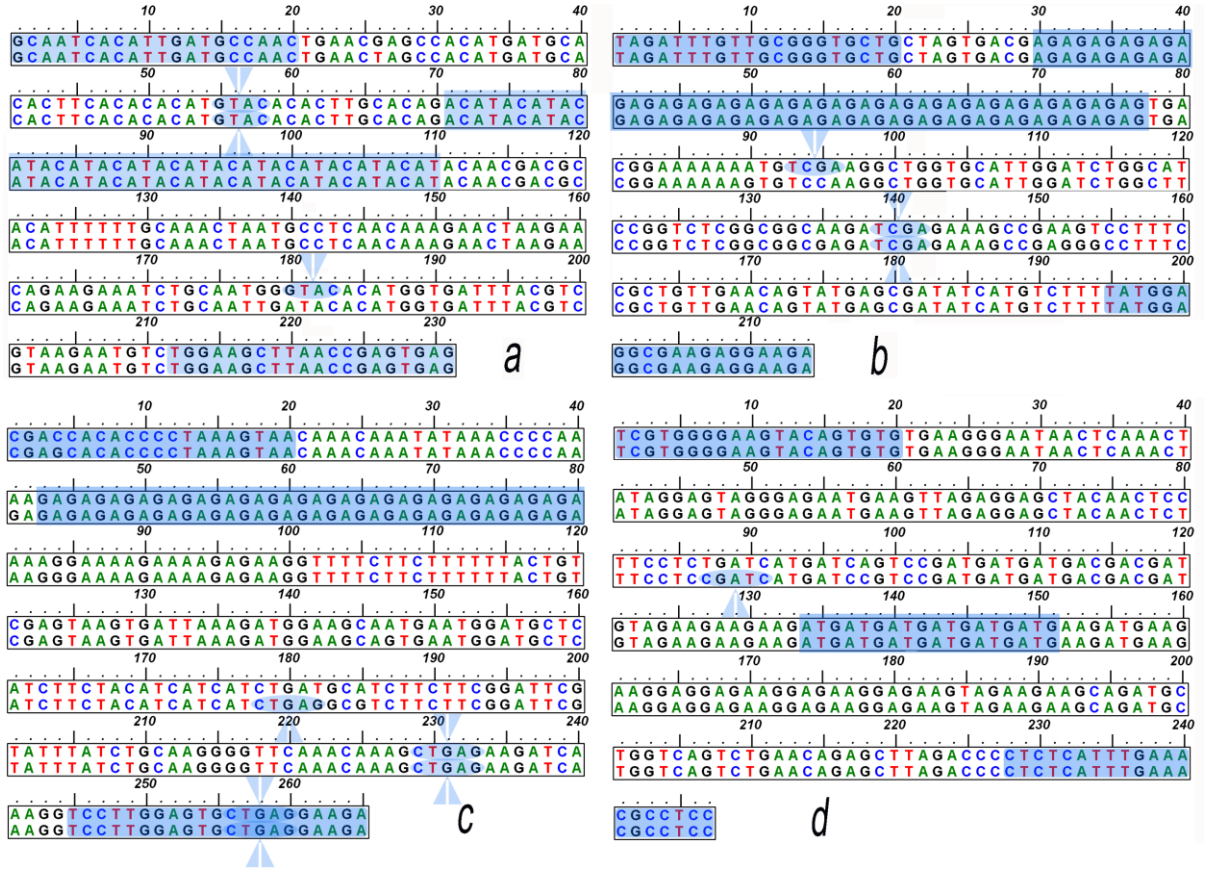

Figure 2. Monomorphic amplicon sequences of TM-1 and Aydin 110. Panel a amplified product sequence of MK092 digested with *Rsa* I restriction enzyme. Panel b: amplified product sequence of MK068 digested with *Taq* I restriction enzyme. Panel c amplified product sequence of MK055 digested with *Dde* I restriction enzyme. Panel d amplified product sequence of MK097 digested with *Alw* I restriction enzyme. In all the panels shown above shaded sequences are primer annealing (at the both ends) and restriction enzyme recognition sites shown by arrows.

dominant and co-dominant markers within the same reaction as observed in Ince et al. (2010c) in *Capsicum*.

In conclusion this study reports new polymorphic microsatellite markers (EST-SSRs and CAPS-microsatellites), which are non-redundant, highly reproducible, polymorphic, co-dominant and could be useful in mapping studies, determination of cultivar purity, efficient use and management of genetic resource collections and the establishment of property rights in cotton.

ACKNOWLEDGMENTS

This research is supported by the Scientific Research Projects Coordination Unit of Akdeniz University and The Scientific and Technological Research Council of Turkey.

LITERATURE CITED

Arumuganathan, K., E. Earle, 1991. Nuclear DNA content of some important plant species. Plant Molecular Biology Reporter, 9:208-218.

Bilgen, M., M. Karaca, A.N. Onus, A.G. Ince, 2004. A software program combining sequence motif searches with keywords for finding repeats containing DNA sequences. Bioinformatics, 20:3379-3386.

Blenda, A., J. Scheffler, B. Scheffler, M. Palmer, J.M. Lacape, J.Z. Yu et al, 2006. CMD: a Cotton Microsatellite Database resource for *Gossypium* genomics. BMC Genomics, 7:132.

Cristancho, M., C. Escobar, 2008. Transferability of SSR markers from related Uredinales species to the coffee rust *Hemileia vastatrix*. Genetics and Molecular Research, 7:1186-1192.

Ince, A.G., M. Karaca, 2009. The MAGi RNA extraction method: highly efficient and simple procedure for fresh and dry plant tissues. Journal of the Science of Food and Agriculture, 89:168-176.

Ince, A.G., M. Karaca, A.N. Onus, 2010a. Genetic Relationships within and between *Capsicum* species. Biochemical Genetics, 48:83-95.

Ince, A.G., M. Karaca, A.N. Onus, 2010b. Polymorphic microsatellite markers transferable across *Capsicum* species. Plant Molecular Biology Reporter, 28:285-291.

Ince, A.G., M. Karaca, A.N. Onus, 2010c. CAPS-microsatellites: use of CAPS method to convert non-polymorphic microsatellites into useful markers. Molecular Breeding, 25:491-499.

Ince, A.G., M. Karaca, A.N. Onus, 2010d. Differential expression patterns of genes containing microsatellites in *Capsicum annuum* L. Molecular Breeding, 25:645-658.

Karaca, M., S. Saha, F.E. Callahan, J.N. Jenkins, J.J. Read, R.G. Percy, 2004. Molecular and cytological characterization of a cytoplasmic-specific mutant in pima cotton (*Gossypium barbadense* L.). Euphytica, 139:187-197.

Karaca, M., M. Bilgen, A.N. Onus, A.G. Ince, S.Y. Elmasulu, 2005a. Exact tandem repeats analyzer (E-TRA) for DNA sequence mining. Journal of Genetics, 84:49-54.

Karaca, M., A.G. Ince, S.Y. Elmasulu, A.N. Onus, K. Turgut 2005b. Coisolation of genomic and organelle DNAs from 15 genera and 31 species of plants. Analytical Biochemistry, 343:353-355.

Lascape, J.M., T.B. Nguyen, S. Thibivilliers, B. Bojinov, B. Courtois, R.G. Cantrell, B. Burr, B. Hau, 2003. A combined RFLP-SSR –AFLP map of tetraploid cotton based on a *Gossypium hirsutum* x *Gossypium barbadense* backcross population. Genome, 46:612-626.

Mellon, J.E., B.A. Triplett (2006) De novo synthesis of peroxidase in cotton ovule culture medium. Physiologia Plantarum, 77:302-307.

Minamiyama, Y., M. Tsuro, M. Hirai, 2006. An SSR-based linkage map of *Capsicum annuum*. Molecular Breeding, 18:157-169.

Roder, M.S., V. Korzun, K. Wendehake, J. Plaschke, M.H. Tixier, P. Leroy, M.A. Ganel, 1998. A microsatellite map of wheat. Genetics, 149:2007-2023.

Saha, S., M. Karaca, J.N. Jenkins, A.E. Zipf, O.U.K. Reddy, A.E. Pepper, R. Kantety, 2003. Simple sequence repeats as useful resources to study transcribed genes of cotton. Euphytica, 130:355-364.

Wang, Y., R. Ren, Z. Yu, 2008. Bioinformatics mining of EST-SSR loci in the Pacific oyster, *Crassostrea gigas*. Animal Genetics, 39:287-289.

Zhang, J., W. Guo, T. Zhang, 2002. Molecular linkage map of allotetraploid cotton (*Gossypium hirsutum* L. x *Gossypium barbadense* L.) with a haploid population. Theoretical and Applied Genetics, 105:1166-1174.