Plant Syst Evol
DOI 10.1007/s00606-013-0790-1

ORIGINAL ARTICLE

Population structure and genetic analysis of different utility types of mango (*Mangifera indica* L.) germplasm of Andhra Pradesh state of India using microsatellite markers

Malathi Surapaneni · Lakshminarayan R. Vemireddy · Hameedunnisa Begum · B. Purushotham Reddy · C. Neetasri · J. Nagaraju · S. Y. Anwar · E. A. Siddiq

Received: 1 November 2012 / Accepted: 27 February 2013

Abstract Genetic analysis of 90 mango genotypes including juicy, table, dual and pickle types from different parts of Andhra Pradesh of India was carried out employing 143 mango-specific microsatellite markers. Of the 143, 34 were new mango-specific microsatellite loci isolated in the course of the present investigation by constructing an $(CA)_n$ and $(TG)_n$-enriched genomic library. Characterization of the 90 genotypes resulted in the detection of 301 alleles from 106 polymorphic loci with an average of 2.87 alleles per locus and polymorphism information content of 0.67. UPGMA cluster analysis grouped all the genotypes into two major groups with a genetic similarity range of 47–88 %. Grouping of the genotypes based on the utility type was observed only at sub-cluster level. Study of population structure by a model-based STRUCTURE analysis revealed the germplasm to exist in four gene pools. Overall F_{st} of 0.11 indicated genetic differentiation between the populations to be low. Analysis of molecular variance revealed that major proportion of the variation was within the individuals (62.25 %). The molecular marker-based study of genetic diversity suggests that the germplasm studied representing the kind of variability would be a valuable genetic resource for future breeding and association mapping in search for new and novel alleles.

Keywords Mango · Microsatellite markers · Genetic diversity · Population structure

Introduction

Mango (*Mangifera indica* L.) the pride fruit of India, is one of the choicest fruit crops of tropical and sub-tropical regions of the world, especially Asia. Its place of importance can be understood from its being referred to as 'King of fruits' in the tropical world (Singh 1996). Because of its nutritive value, delicious taste, excellent flavor, attractive fragrance and health promising qualities, mango has gained global popularity in the last two decades. It is commercially grown in over 103 countries. The major mango growing countries in the world are India, China, Mexico, Pakistan, Indonesia, Thailand, South and Central America, The Philippines, Brazil, Australia, Nigeria and Egypt (Hemanth kumar et al. 2007). The main exporting producing nations include Mexico (23 % of production), Brazil (14.3 %), Pakistan (3.2 %), Peru (10.3 %) and India (9.71 %) (Bally 2011). Mango has been under cultivation since 4,000 years in the Indian subcontinent. Endowed with rich diversity India is considered to be the center of origin of it (Ravishankar et al. 2000). As of now, more than 1,000 mango cultivars are known to exist in the country (Karihaloo et al. 2003) representing the largest mango genepool in the world. Andhra Pradesh which is considered to be 'The Mango state of India' has the richest mango

Electronic supplementary material The online version of this article (doi:10.1007/s00606-013-0790-1) contains supplementary material, which is available to authorized users.

M. Surapaneni (✉) · L. R. Vemireddy · C. Neetasri · E. A. Siddiq
Institute of Biotechnology, ANGRAU, Hyderabad, India
e-mail: malathi3@rediffmail.com

H. Begum · B. Purushotham Reddy
Vegetable Research Station, ARI, APHU, Hyderabad, India

J. Nagaraju
Laboratory of Molecular Genetics, CDFD, Hyderabad, India

S. Y. Anwar
Department of Genetics, Osmania University, Hyderabad, India

Published online: 04 April 2013

germplasm. There are 100 of mango varieties in the state and of hundreds of varieties 25–30 are commercially popular as juicy, table and pickle types.

Global production of mango has estimated to be 30 million tonnes (2010) of which India accounts for more than 50 % (15.2 million tonnes) (FAOSTAT 2010). Within India, the state of Andhra Pradesh state ranks first in terms of both area (0.48 million ha.) and production (4.05 million tonnes) (Indian Horticulture Database 2010). The other major mango growing states are Uttar Pradesh, Bihar, Karnataka, Tamil Nadu, Paschimbanga (West Bengal), Odisha and Maharashtra. Indian mango is largely exported to UAE, Bangladesh, Saudi Arabia, and the UK.

Andhra Pradesh '*The mango state of India*' is the major center of diversity in the country. The germplasm largely confined to the districts of Krishna, East Godavari and West Godavari, Khammam, Vizayanagaram, etc., may be broadly grouped into table, juicy and pickle types. They are characterized on the basis of a set of agro-botanic traits. Those fruits which can be sliced easily and are almost free of fiber come under table type. Banaganapalle (Beneshan), Himayat, Jahangir, Mulgoa, Dashehari, Alphonso, Banglora, etc., are some of the popular table varieties. Juicy varieties popularly known as '*Rasaalu*' are characterized by high fiber content and abundant juice. Peddarasam, Cherukurasam, Chinnarasam, Panchdarakalasa, Nuziveedu rasalu are some of the popular varieties of this group. The third category is pickle type, wherein fruits in raw unripened stage are used for preparation of pickles. Tellagulabi, Royal Special, etc., belong to pickle types. Some are dual types categorisable as juicy or table type. Nawab Pasand, Rumani, kumkum are some of the dual types.

Success of any crop improvement program depends essentially on nature and magnitude of genetic variability available in crop germplasm. Over the last two decades, both international (Hossaim et al. 2001; Saleh et al. 2009; Mussane et al. 2010; Rajwana et al. 2011) and national (Gosh et al. 1985; Gowda and Ramanjaneya 1994; Singh et al. 2009; Begum et al. 2012) efforts were made in understanding the extent of variability and grouping of mango germplasm into distinct genetic clusters based on morpho-physiological traits. These markers have many limitations particularly in fruit crops because of long generation time besides being influenced by environment. Molecular markers on the other hand have proved their potentiality in unraveling the limitations of morphological traits. In recent years, extensive molecular work has been carried out to estimate the genetic variation in mango germplasm using random amplified polymorphic DNA (RAPD) markers (Schnell et al. 1995; Lopez et al. 1997; Ravishankar et al. 2000; Hemant Kumar et al. 2001; Karihaloo et al. 2003; Rahman et al. 2007; Rajwana et al. 2008; Díaz-Matallana et al. 2009; Souza et al. 2011; Ramessur and Ranghoo 2011) and restriction fragment length polymorphism (RFLP) analysis (Eiadthong et al. 1999); amplified fragment length polymorphism (AFLP) markers (Eiadthong et al. 2000; Kashkush et al. 2001; Yamanaka et al. 2006; Gálvez-López et al. 2010) inter simple sequence repeat (ISSR) markers (He et al. 2005; Pandit et al. 2007; Srivastava et al. 2007; Tomar et al. 2011; Rocha et al. 2012; Samal et al. 2012) and start codon targeted (SCoT) markers (Luo et al. 2010). Among the molecular markers simple sequence repeat or (microsatellites) have been found to be the marker of choice for more precisely understanding the genetic diversity, gene mapping, and cultivar discrimination due to their abundance, co-dominant and high reproducibility nature (Gupta and Varshney 2000). Nevertheless, only few attempts have been made to develop and use mango-specific genomic SSR markers. It was Viruel et al. (2005), who reported the first set of 16 microsatellite markers and validated them in 28 cultivars. Similar attempts have been made by few others subsequently to isolate SSRs in mango (Viruel et al. 2005; Duval et al. 2005; Honsho et al. 2005; Schnell et al. 2005; Utoskit 2007; Ravishankar et al. 2011).

Understanding how genetic variation is distributed within and among populations is important to germplasm management, crop breeding and association mapping. The use of DNA-based markers offers another approach for population-level genetic analysis. Model-based clustering method has been developed to detect underlying population structure in a collection of individuals genotyped with molecular markers. The program STRUCTURE (Pritchard et al. 2000) uses a Bayesian approach and has been utilized in numerous genetic diversity and association mapping studies in plant species including tomato (Sim et al. 2011), apple (Urrestarazu et al. 2012) and strawberry (Yoon et al. 2012). In this study, an effort has been made to generate and identify more polymorphic markers to understand the genetic diversity and investigate structure of mango germplasm of India with special emphasis on genotypes collected from Andhra Pradesh state.

Materials and methods

Survey and collection of mango germplasm

A well-planned germplasm collection survey was conducted in different geographical areas of Andhra Pradesh which includes 14 major mango growing districts of state horticultural research stations and private owned mango orchards. Random sampling strategy was followed for collection of leaf samples. In all 90 genotypes comprising 55 juicy, 22 table, 7 dual (juicy/table) and 6 pickle types were used for the study (Table 1).

Table 1 Mango cultivars used in the study

S. no.	Genotypes	Type	Area of collection/ district	S. no.	Genotypes	Type	Area of collection/district
1	Peddarasam	Juicy	HRS, Antharajpet, Kadapa	46	Amrigola	Juicy	HRS, Antharajpet, Kadapa
2	Cherukurasam	Juicy	HRS, Antharajpet, Kadapa	47	Janardhanpasand	Table	HRS, Antharajpet Kadapa
3	Chinnarasam	Juicy	Nuzividu, Krishna	48	Zardalu	Table	Veeraballi, Kadapa
4	Panchadarakalasam	Juicy	HRS, Antharajpet, Kadapa	49	Jalalu	Pickle	Raghavapuram, Krishna
5	Tellarasam	Juicy	Reddygudem, Krishna	50	Tellagulabi	Pickle	Reddygudem, Krishna
6	Sinduri	Juicy	Rangapuram, Krishna	51	Chinnaachar	Pickle	Panyam, Nandyala
7	Amrutham	Juicy	Panyam, Nandyala	52	Peddaachar	Pickle	Panyam, Nandyala
8	Delhipasand	Juicy	Panyam, Nandyala	53	Punasabaramasi	Pickle	HRS, Antharajpet Kadapa
9	Reddipasand	Juicy	Panyam, Nandyala	54	Royal special	Pickle	HRS, Antharajpet Kadapa
10	Doodpeda	Juicy	Banganapalle, Kurnool	55	Himayat	Table	Mylavaram, Krishna
11	Shakargola	Juicy	Yagantipalle, Kurnool	56	SuvernaRekha	Juicy/ Table	ARI, Rajngareddy
12	Natupalli	Juicy	Yagantipalle, Kurnool	57	Chinna Suvernarekha	Juicy/ Table	HRS, Antharajpet, Kadapa
13	Nadusalai	Juicy	Anantharajpet, Kadapa	58	Mylapur punasa	Juicy	FRS, Sangareddy
14	Moolky	Juicy	Anantharajpet, Kadapa	59	Aryavartharasalu	Juicy	FRS, Sangareddy
15	Panakalu	Juicy	Anantharajpet, Kadapa	60	Panchavarnam	Juicy	FRS, Sangareddy
16	NagulapalliIrasalu	Juicy	Anantharajpet, Kadapa	61	LajjatBaksha	Juicy	FRS, Sangareddy
17	Peter	Juicy	Anantharajpet, Kadapa	62	Kothapalli kobbari	Juicy	FRS, Sangareddy
18	Pulihora	Juicy	Anantharajpet, Kadapa	63	Kanthavarapadu	Juicy	FRS, Sangareddy
19	Dashehari	Juicy	Anantharajpet, Kadapa	64	Nuzeevedurasalu	Juicy	Nuziveedu, Krishna
20	Vikarabad	Juicy	Anantharajpet, Kadapa	65	Desavali	Juicy	Pithapuram, East Godavari
21	Laddupasand	Juicy	VRPalle, Kadapa	66	Paparaoguava	Juicy	Kathipudi, East Godavari
22	Lalbahar	Juicy	VRPalle, Kadapa	67	Kolanguava	Juicy	Kathipudi, East Godavari
23	Kesar	Juicy	Tpalle, Rangareddy	68	Kobbariantu	Juicy	Palakinda sankili, Srikakulam
24	Dondakayalamamidi	Juicy	Tpalle, Rangareddy	69	Kalamamidi	Juicy	Kathipudi, East Godavari
25	Panukulamamidi	Juicy	Tpalle, Rangareddy	70	Patikarasalu	Juicy	Akkulapet, Srikakulam
26	Goa	Juicy	Tpalle, Rangareddy	71	Nallayendrasulu	Juicy	Akkulapet, Srikakulam
27	Panduluivarimamidi	Juicy	Tpalle, Rangareddy	72	Mukkurasalu	Juicy	Akkulapet, Srikakulam
28	Nawabpasand	Juicy/ Table	Panyam, Nandyala	73	Mettavalayapeechumanu	Juicy	Bobbili, Vijayanagaram
29	Rumani	Juicy/ Table	Panyam, Nandyala	74	Chuttamrutham	Juicy	Bobbili, Vijayanagaram
30	Khader	Juicy/ Table	HRS, Antharajpet, Kadapa	75	Bobbilipeechumanu	Juicy	Bobbili, Vijayanagaram
31	Kumkum	Juicy/ Table	Yagantipalli, Kurnool	76	Kinthalooripeta	Juicy	Bobbili, Vijayanagaram
32	Phirangiladwa	Juicy/ Table	HRS, Antharajpet, Kadapa	77	Sannarasalu	Juicy	Tuni, East Godavari
33	Beneshan	Table	Banaganapalle, Kurnool	78	Erraarati	Juicy	Bobbili, Vijayanagaram
34	Banglora	Table	FRS, Sangareddy	79	Kobbarimamidi	Juicy	Kathipudi, East Godavari
35	Tiyyamamidi	Table	HRS, Antharajpet, Kadapa	80	Punasa	Juicy	Kathipudi, East Godavari
36	Imampasand	Table	Sanampudi, Prakasham	81	Navaneetham	Juicy	A.R.I, Rangareddy
37	Sora	Table	FRS, Sangareddy	82	ChitoorRasalu	Juicy	Narasingapuram, Chitoor

Table 1 continued

S. no.	Genotypes	Type	Area of collection/ district	S. no.	Genotypes	Type	Area of collection/district
38	Alphonsa	Table	Mylavaram, Krishna	83	MallepalliBN	Table	H.R.S, Mallepalli, Nalgonda
39	Neelum	Table	Hyderbagh, Kurnool	84	AlampurBN	Table	FRS, Sangareddy
40	Abbasi	Table	Yagantipalli, Kurnool	85	VeeraballiBN	Table	Veeraballi, Kadapa
41	Safed Damini	Table	Yagantipalli, Kurnool	86	RatiBpalli	Table	FRS, Sangareddy
42	Govander	Table	Yagantipalli, Kurnool	87	Hydersahib	Table	Kathipudi, East Godavari
43	Chalka	Table	Yagantipalli, Kurnool	88	Malgoa	Table	HRS, Antharajpet, Kadapa
44	Hublee	Table	Yagantipalli, Kurnool	89	Natlu	Juicy	Mylavaram, Krishna
45	Jahangir	Table	HRS, Antharajpet, Kadapa	90	Palli	Juicy	Yagantipalli, Kurnool

Construction and screening of microsatellite enriched library

Total genomic DNA was extracted from fresh leaf samples of the mango cultivar 'Peddarasam', the most popular juicy variety of Andhra Pradesh by using CTAB method as per the Porebski et al. (1997) protocol. Two microsatellite enriched libraries $(CA)_n$ and $(TG)_n$ were constructed following the protocol of Glenn and Schable (2005). Genomic DNA was digested with *Rsa*I and *Xmn*I restriction enzymes (New England Biolabs, USA). After digestion, DNA fragments were ligated to linkers and pre-amplification was carried out using linker-specific primers. Biotinylated probes of two dinucleotide repeats $(CA)_n$ and $(TG)_n$ were used to capture fragments containing microsatellites and linker-specific primer was used for enrichment by PCR. The enriched fragments were cloned into pGEM-T Easy vector (Promega, Madison, Wisconsin, USA). A total of 100 clones were randomly chosen and screened by colony PCR. About 50 positive clones were sequenced and clones containing microsatellites along with flanking sequences were selected for primer design. Primers were designed using Primer3 v. 0.4.0.

SSR marker genotyping

The genotypes were analyzed using 143 SSR primer pairs which include 34 SSR primers developed in the present study and 109 primers from the previous studies. PCR amplification was carried out in 10 µl reaction volume consisting of 10 mM Tris–HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM each dNTPs, 5 pmol of each primer 0.5 U of Taq DNA polymerase (New England Biolabs, USA) and 10 ng of genomic DNA. Reactions were carried out in Gen Amp PCR system 9700 (Applied Biosystems, USA) thermocycler using the following temperature profile: an initial denaturation step of 5 min at 94 °C followed by 35 cycles of 45 s at 94 °C, 45 s at 55 °C and 1 min at 72 °C, then a final extension of 5 min at 72 °C. Amplification products were resolved on 3 % metaphore agarose gels. The size of each band was determined by comparing with a size standard 50 bp DNA ladder (New England Biolabs, USA).

SSR analysis

Variability parameters like number of alleles (*Na*) per locus, observed heterozygosity (*Ho*), expected heterozygosity (*He*) were evaluated using POPGENE v 1.31 (Yeh et al. 1999). Polymorphism information content (PIC) was calculated by $PIC = (1 - \Sigma pi^2)$ as described by Anderson et al. (1993). Genetic relatedness among the genotypes was calculated using unweighted pair group method with arithmetic averages algorithm (UPGMA) cluster analysis by using the program NTSYSpc 2.02i (Rohlf 1989). A neighbor-joining tree with bootstrap values was constructed using DARwin ver 5.0.145 software (Perrier et al. 2003). Also FCA analysis was carried out by DARwin ver 5.0.145 software. The *F* statistics (F_{st}) were used to analyze the genetic differentiation of the populations and Analysis of Molecular Variance (AMOVA) was performed to assess the genetic variability among the populations using ARLEQUIN ver 3.0 (Excoffier et al. 2005). A model-based clustering method was applied to infer the genetic structure using the software SRUCTURE version 2.3.1 (Pritchard et al. 2000). Five independent runs were performed by setting the number of clusters (*K*) from 2 to 4. Each run consisted of burn-in-period of 1,00,000 steps followed by 1,00,000 MCMC replications. No prior information was used to define the clusters. The run with highest and consistent log likelihood was selected ($K = 4$) to determine inferred ancestries.

Table 2 Salient features of the 34 mango-specific microsatellite markers developed in the present study

Locus	Genebank accession no.	Sequence	Tm	Allele size	Repeat motif	Na	He	Ho	PIC	Allele frequency
MGDSSR1F	JF487796	CGAAATGAGACACCTGCAAA	55	206–220	(GA)2(CA)2(AG)7 (GA)5	2	0.2485	0.0856	0.62	
MGDSSR1R		TTTCCTCCATTGCTTTTTCG								
MGDSSR2F	JF487797	GGGAATGGTAGAGACGGACA	55	200–270	(AG)8	2	0.1629	0.0842	0.57	0.9345
MGDSSR2R		ATCCAAGCAGTCACCATCAA								
MGDSSR3F	JF487798	TGAAAAAGGTTTAGGCGAAAAA	51	196	(TG)15	–		–	–	
MGDSSR3R		CTCAAGCTATGCATCCAACG								
MGDSSR4F	JF487800	AGCAGAATCCATTCTTGATTGA	51	262	(AC)16	–		–	–	
MGDSSR4R		CCTAGCTAGCAGAATCACAAAAA								
MGDSSR5F	JF487801	CGATAGTGCCAATCTGGTGA	57	210–225	(GT)10	2	0.4178	0.1548	0.6	0.6964
MGDSSR5R		TCATCTCACACACTCTCTCTCTCTC								
MGDSSR6F	JF487825	ACTTGGAAATGTTTTATCTTTTGAA	51	200–250	(CA)5(GA)5 (GA)8	2	0.3228	0.3000	0.48	0.7976
MGDSSR6R		TGGTTTTTCATAGCCAAATGC								
MGDSSR7F	JF487820	GCTAGCAGAATCCTAGCAGAATC	51	162	(CT)9	–		–	–	
MGDSSR7R		GCAGAATCACAAATATCATCCATC								
MGDSSR8F	JF487821	AAGGCCATACACTGGAAACG	53	330	(AT)7(TG)13	–		–	–	
MGDSSR8R		CCCTTTCGGTGGTCTCTCTC								
MGDSSR9F	JF487818	TCAAGAAAAGCAAAGAAAAGCA	53	123	(TG)8(AG)6	–		–	–	
MGDSSR9R		AGCAGAATCGCTTCAGATGT								
MGDSSR11F	JF487797	GGGAATGGTAGAGACGGACA	58	150–210	(TA)3(AG)8	3	0.5396	0.1748	0.79	0.4881
MGDSSR11R		TTCATCATAGGTCCCACACG								
MGDSSR12F	JF487799	TCGGTAAACATTAGACAGGATTGA	58	117	(GT)6					
MGDSSR12R		CCAATTACCAAACCCTCATTTT								
MGDSSR13F	JF487802	GGTTAACTCCAAAATGAAGACGA	57	180–200	(CT)6	2	0.4132	0.1667	0.62	0.6905
MGDSSR13R		GCTAGCAGAATCTCTCTGGAATG								
MGDSSR14F	JF487803	AATGCTGAGCCTGGTAAGGA	58	160–310	(AG)6	3	0.5184	0.5778	0.5	0.4643
MGDSSR14R		CAACATCCTCTTTCTTCCCTGT								
MGDSSR15F	JF487804	GCATATGAATTGAGCCCTTG	58	250–320	(TGCA)3(TGTA)2 (TG)6	2	0.4541	0.6310	0.33	0.6369
MGDSSR15R		AGCAGAATCCAACCATGCTA								
MGDSSR16F	JF487809	AATGCCCAGCTAGGGAGAAT	53	220	(GT)6					
MGDSSR16R		CAACATCTTCTTTCTTCCCTGT								
MGDSSR17F	JF487806	AGCAGAATCGCTTCACAACA	57	160–180	(AC)7	2	0.4670	0.4762	0.41	0.6429
MGDSSR17R		GAATCACGTGCTGCGTTAAA								
MGDSSR18F	JF487805	GCATATGAATTGAGCCCTTG	53	259	(TG)6	–		–	–	
MGDSSR18R		GGCCTAGCTAGCAGAATCCA								
MGDSSR19F	JF487826	CTTTTCGCGCGTATACATGA	51	200–210	(CT)5	2	0.0854	0.0595	0.49	0.9702
MGDSSR19R		ACACGCGTAGAGAACACACG								
MGDSSR20F	JF487819	TCAGCTAGCTTCCCACCAAC	55	200–220	(GT)7	2	0.0648	0.0500	0.53	0.963
MGDSSR20R		GCTAGCAGAATCCTAGCAGAATC								
MGDSSR21F	JF487822	GCAGAATCACATGAAAACGAGA	55	198	(CA)5	–		–	–	
MGDSSR21R		TCTCCATGCATTAGCTTCCA								
MGDSSR22F	JF487823	GGGATTCCTATTGGTCCACA	55	155–210	(AC)7	3	0.4847	0.5833	0.56	0.6250
MGDSSR22R		GCAATCCCTTTGGGTAAACA								
MGDSSR23F	JF487806	CAAGCAAAATCGCTTCACAA	56	170–200	(CA)7	2	0.4608	0.6905	0.24	0.6548
MGDSSR23R		GAATCACGTGCTGCGTTAAA								
MGDSSR24F	JF487807	TCACAGAGAAAATGCTTCTGAAT	55	150–180	(CT)11	2	0.1986	0.2143	0.48	0.8810
MGDSSR24R		TGTTTCCCAGCGAACAAACT								
MGDSSR25F	JF487808	GCAAAATCACTTTGGGTTTCA	56	220–240	(GA)8(AG)2 (AAG)3	2	0.2158	0.0800	0.6	0.8810
MGDSSR25R		CCCACCTTTGACATTTGATG								
MGDSSR26F	JF487809	CAAAATCCCTGGAGGTGAGA	58	251	(GA)9AGG (GA)4	–		–	–	
MGDSSR26R		AGGGCAACAATTTGAAGCTG								

Table 2 continued

Locus	Genebank accession no.	Sequence	Tm	Allele size	Repeat motif	Na	He	Ho	PIC	Allele frequency
MGDSSR27F	JF487810	CAAAATCCCTGGAGGTGAGA	57	103	(GA)9	–		–	–	
MGDSSR27R		CCGTTGCTTTTCTAAACATCTCT								
MGDSSR28F	JF487811	CCTTTTCCATGCAGTTTTAC	57	188	(GA)9	–		–	–	
MGDSSR28R		TGAATAGTATTATTTGTGTGCAT								
MGDSSR29F	JF487812	GTTTAAGGCCTAGCAAGCA	55	140–160	(TG)9	2	0.0221	0.0550	0.51	0.9881
MGDSSR29R		TTCGAGCAGTCTTCATCAC								
MGDSSR30F	JF487813	TCACTTTGGGTTTCAACTTTCA	55	117	(AG)7 (AAG)2	–		–	–	
MGDSSR30R		ATCCGAGCTTGAAACAGCAT								
MGDSSR31F	JF487814	AAGCAGAATCACAGCCTCTTG	55	150–170	(AG)18	2	0.1899	0.1667	0.47	0.9167
MGDSSR31R		AAAGACAGCCATGACCATCC								
MGDSSR32F	JF487815	CACAGAGAAAATGCTTCTGAATTA	58	150–170	(TC)12	2	0.1720	0.0833	0.54	0.8988
MGDSSR32R		TGTTTCCCAGCGAACAAACT								
MGDSSR33F	JF487816	CAAACCCACACTTGCACAAA	59	270–310	(CA)5	3	0.0863	0.0560	0.68	0.9464
MGDSSR33R		CATTTGCTCCAGCAACTTGA								
MGDSSR34F	JF487817	GAAAGTGAGACCTTCGGTTCC	58	160–190	(GA)17 AGG (GA)5	3	0.5340	0.9643	0.4	0.5060
MGDSSR34R		AAGGCCCCTTCTTCACATTT								
MGDSSR35F	JF487799	CATTAGACAGGATTGATGCTCAC	53	108	(GT)6	–		–	–	
MGDSSR35R		CCAATTACCAAACCCTCATTTT								

Na number of alleles, *Ho* observed heterozygosity, *He* expected heterozygosity, *PIC* polymorphic information content

Results

Microsatellite development

SSR-enriched library was constructed from the promising cultivar 'Peddarasam' following the protocol of Glenn and Schable (2005). A total of 100 clones from the enriched genomic library was randomly chosen and screened by colony PCR. Of them 50 showed positive signal, an enrichment of 52 %. Of the 50 clones sequenced, 38 readable sequences were obtained containing $(CA)_n$ and $(TG)_n$ repeats and the rest were deleted due to non-availability of flanking regions for primer. In all, 34 genomic SSR primers were designed with melting temperature ranging from 52 to 57 °C. Thus developed microsatellite sequences have been deposited with the Genbank and the details of them are given in Table 2.

NCBI blast searches showed a highly significant homology of the one microsatellite flanking regions of one of the clones sequenced (MGDSSR28, accession number JF487811) with *Vitis vinifera* transcription factor TT2-like (LOC100250940), mRNA (accession no.—XM_002278193.2).

Table 4 Nei's (1978) unbiased measures of genetic distance (below diagonal) and pairwise differences values (upper diagonal) between four utility type mango populations of Andhra Pradesh

Population	Juicy	Dual	Table	Pickle
Juicy	****	0.0522**	0.029*	0.084***
Dual	0.0569	****	0.047**	0.091***
Table	0.0300	0.0500	****	0.058**
Pickle	0.0901	0.0885	0.0611	****

*** Significance at the 0.1 % level, ** significance at the 1 % level, * significance at 5 % level

Table 3 Genetic diversity indices of mango accessions in four utility type populations

Population	*N*	N_{at}	N_{pl}	N_r	N_{al}	N_{pr}	Ho	He	Average gene diversity
Juicy	55	301	106	8	2.86	9	0.2897	0.4509	0.450936 ± 0.217703
Dual	8	252	84	0	2.35	0	0.2677	0.3914	0.391352 ± 0.199786
Table	21	287	99	16	2.66	0	0.2686	0.4132	0.413232 ± 0.202605
Pickle	6	238	77	0	2.16	0	0.2720	0.3752	0.375214 ± 0.196087

N number of individuals per population, N_{at} total number of alleles per population, *Npl* number of alleles at polymorphic loci, *Nr* number of rare alleles, N_{al} mean number of alleles per locus, *Npr* number of private alleles per population, *He* Expected homozygosity, *Ho* observed heterozygosity

Table 5 Analysis of molecular variance (AMOVA) among the populations

Source of variation	*df*	Variance of components	Percentage of variation
Among populations	3	0.87993	3.67
Among individuals within population	86	8.17205	34.08
Within Individuals	90	14.92778	62.25

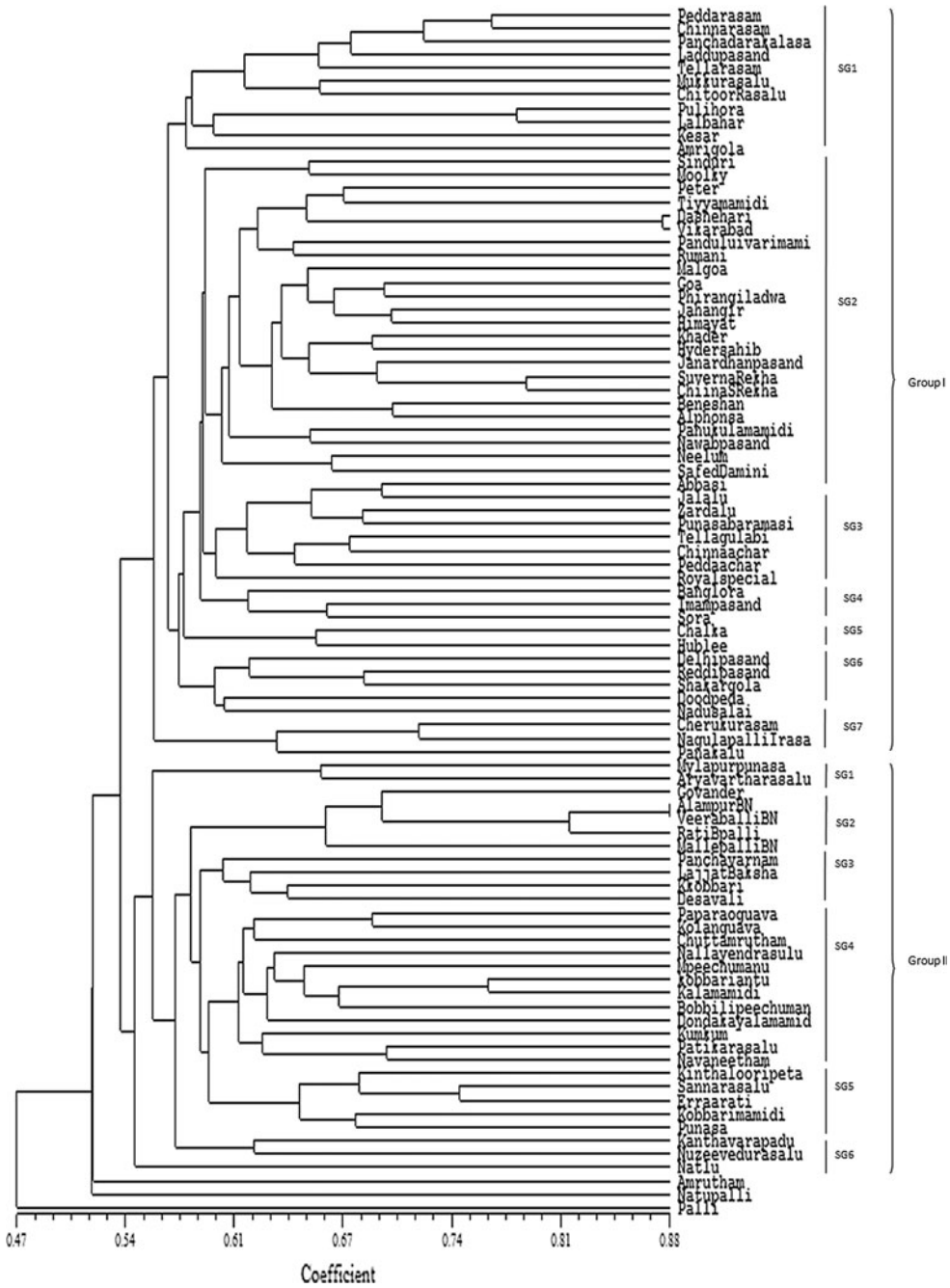

Fig. 1 Dendrogram using UPGMA cluster analysis based on the genetic diversity of 90 mango genotypes

Overall SSR diversity

In all, 143 microsatellite primer pairs (34 newly developed in the current study and 109 from previous studies) were used to assess the genetic diversity in the 90 genotypes. Of them, 106 (74.1 %) were found to have consistent, clear and polymorphic amplification profiles. A total of 301 alleles were detected across the 106 polymorphic loci, with

the product size ranging from 110 to 340 bp. The number of alleles per locus ranged from 2 to 6 with an average of 2.87 per locus. All the microsatellite loci showed moderate to high PIC values, the range being between 0.23 (MGDSSR-23) and 0.97 (SSR-44) with an average of 0.67. Average percentage of high frequency alleles was 66 with the range being from 27 % (SSR85) to 97 % (MGDSSR29). Overall observed heterozygosity (*Ho*) per marker ranged from 0 to 0.87 with a mean of 0.29 and the expected heterozygosity (*He*) ranging from 0 to 0.81 with the mean of 0.39 (Sup Table 1 and 2).

Population differentiation

All the accessions were divided into four populations according to the utility type viz., juicy (54), table (22), dual (8) and pickle (6). Total number of alleles per population varied from 238 in pickle mango to 301 in juicy mango. Allelic frequency was moderate to high and as a consequence, low number of rare alleles (alleles present in fewer than 5 % of the individuals in a population) were observed. The number of rare alleles was in the range of 8 in juicy and 16 in table. The mean number of alleles per locus and per population was found to range from 2.16 in pickle to 2.86 in juicy varieties (Table 3). Nine cultivar-specific alleles were observed in the juicy population, while no such alleles were observed in other populations. The observed heterozygosity per population ranged from 0.26 (dual) to 0.28 (juicy) while the expected heterozygosity per population ranged from 0.37 (pickle) to 0.45 (juicy).

The tests for pairwise genetic differentiation among populations were significant but the F_{st} values were relatively low suggesting that differentiation was not strong enough. The highest F statistics value among the utility

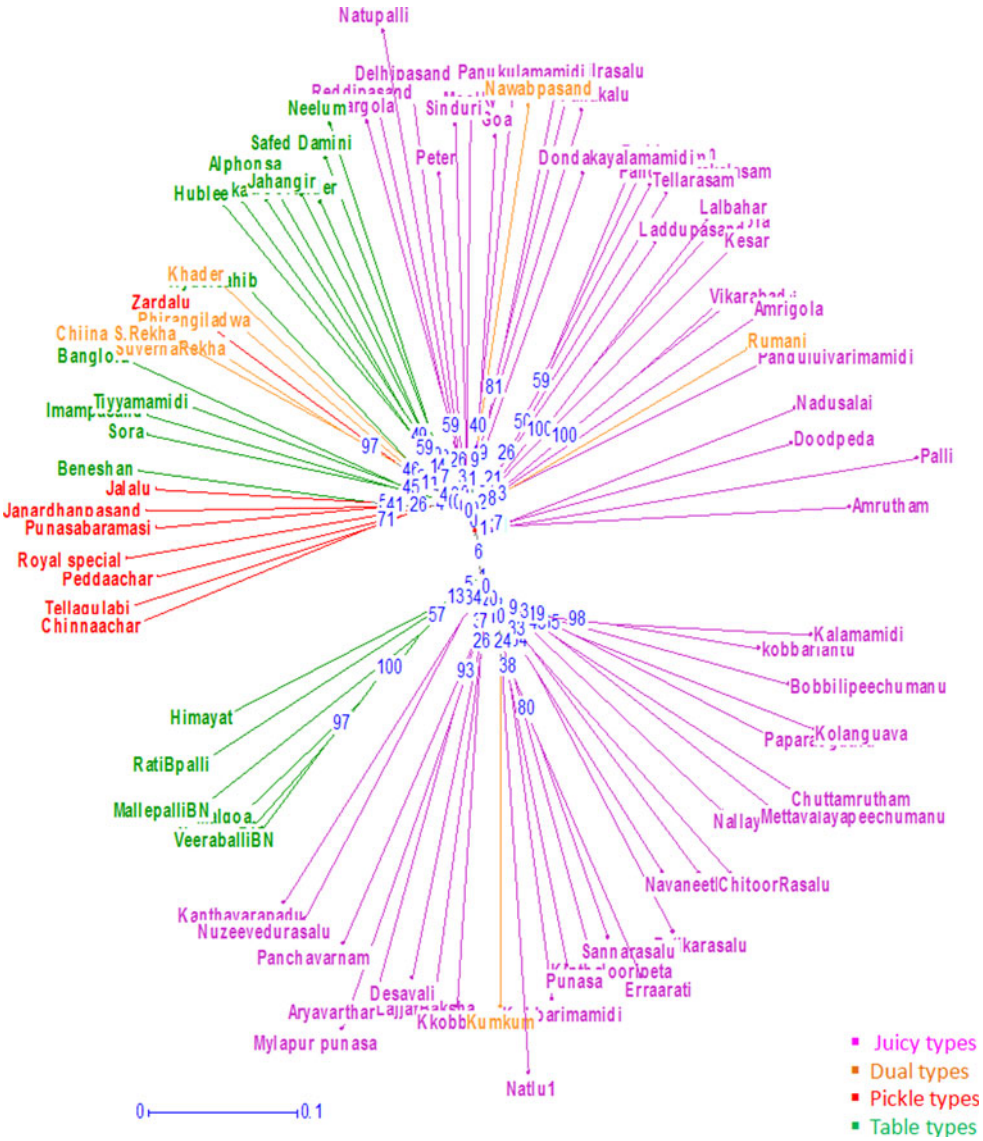

Fig. 2 An unrooted neighbor-joining tree showing the genetic relationships among the genotypes. *Pink color* indicates juicy types, *green* for table, *red* for pickle and *orange* for dual types

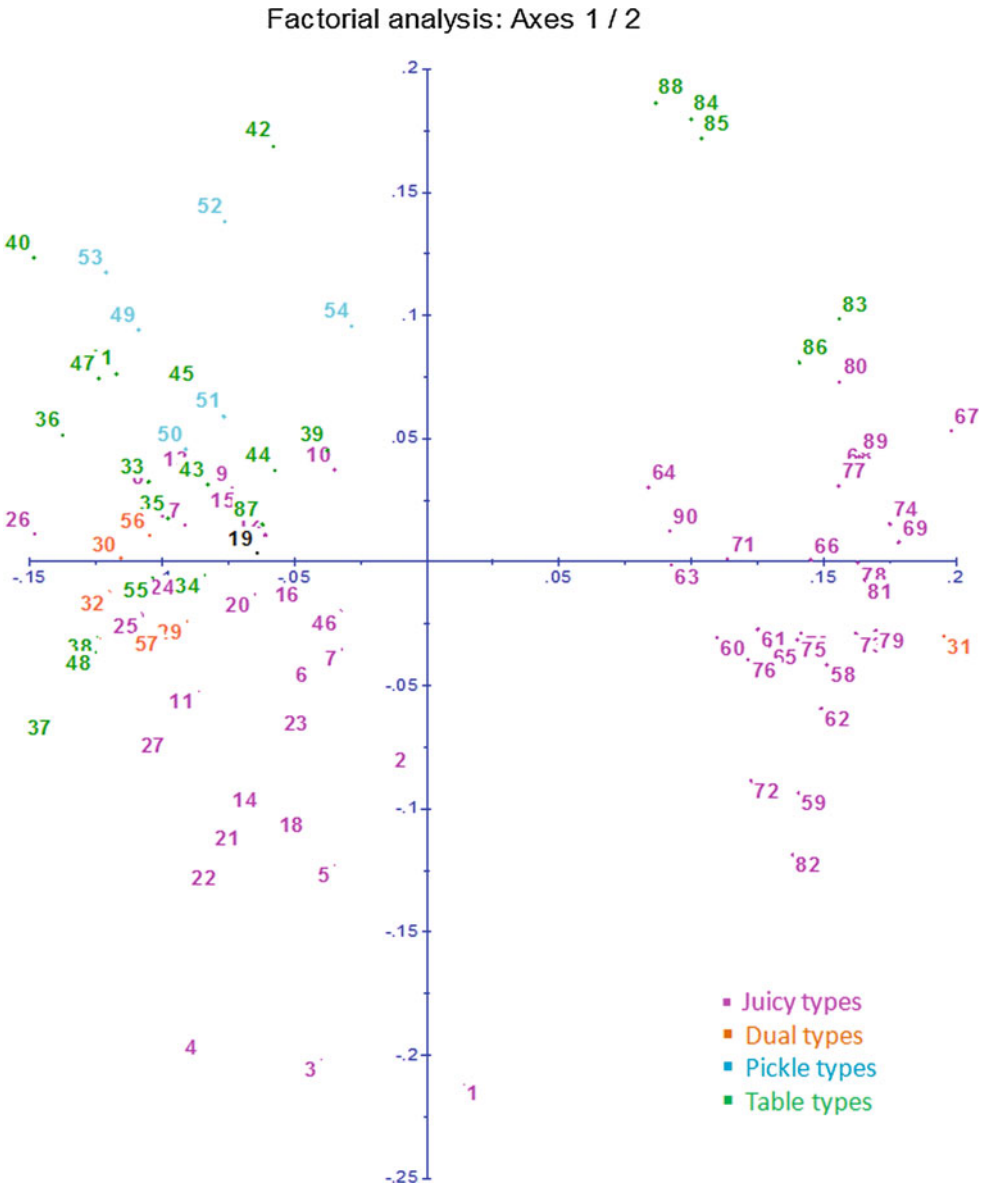

Fig. 3 *Plot* of the factorial correspondence analysis (FCA) obtained with all SSR data for the 90 genotypes. *Pink color* indicates juicy types, *green* for table, *blue* for pickle and *orange* for dual types

Table 6 Proportion of ancestry of each population defined with the model-based clustering method

Population	$K = 2$		$K = 3$			$K = 4$			
	A	B	A	B	C	A	B	C	D
Juicy	0.4868	0.5132	0.4872	0.0802	0.4326	0.30	0.1646	0.349	0.1871
Dual	0.9855	0.0144	0.0111	0.4151	0.5734	0.371	0.016	0.0165	0.5965
Table	0.7015	0.2984	0.2985	0.5582	0.1431	0.102	0.2011	0.1180	0.5780
Pickle	0.9098	0.0901	0.0431	0.946	0.011	0.011	0.0611	0.0275	0.9005

A–D clusters

type populations was observed between dual type and pickle followed by juicy and pickle (Table 4). A similar pattern of differentiation among the populations was observed in Nei's genetic distance. AMOVA analysis showed that difference among populations accounted for 3.67 %, while among individuals within population was 34.08 % and highest genetic variation was observed within individuals accounting for 62.25 % (Table 5).

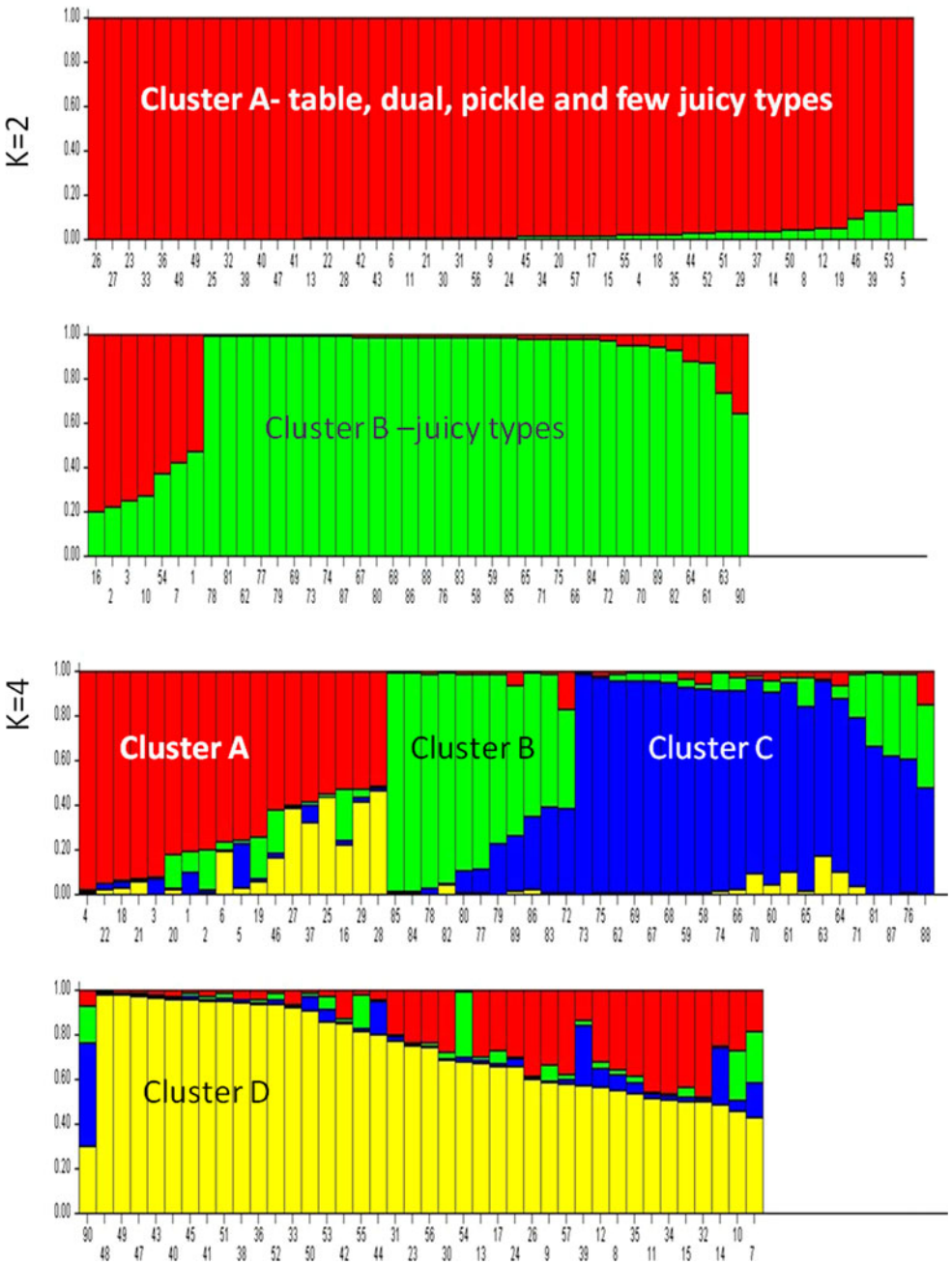

Fig. 4 Analysis on a model-based clustering method at $K = 2$ (*first panel*) and $K = 4$ (*second panel*). Each individual is represented by a *column* and the *different colors* refer to the different gene pools

Genetic diversity analysis

UPGMA cluster analysis method was employed to construct a dendrogram based on the allelic data obtained from 143 SSR markers. The cluster analysis showed a significant genetic variation among the mango genotypes studied with a similarity coefficient ranging from 0.47 to 0.88. The dendrogram revealed the genotypes studied to form two distinct groups and each with a few subgroups at the similarity coefficient of 0.47 (Fig. 1). Group I comprised 64 % of the genotypes covering all the four utility type mangoes with similarity coefficient ranging from 0.55 to 0.87. The group was further divided into seven subgroups (SG) and among the subgroups juicy cultivars such as Peddarasam, Chinnarasam and Panchdarakalasa clustered together to form SG1 with a similarity of 71 %, while all the pickle type varieties clustering together formed the subgroup 3 (SG3) with a genetic similarity of 61 %. Most of the varieties clustered according to their utility types. For instance juicy varieties clustered in three closely placed subgroups viz SG1, SG6 and SG7, while pickle varieties in SG3, table types in SG4 and SG5. The subgroup SG2 represented a mixture of juicy, table, dual types. Interestingly 'Beneshan' (Banaganapalli) the popular table type

cultivar of Andhra Pradesh clustered with 'Alphonso', another popular table type from Konkan region of Maharastra with 70 % genetic similarity in SG2. Group II comprising 32 % of the total genotypes studied, consisted largely of juicy types with a similarity coefficient ranging from 0.55 to 0.80. Locally popular juicy varieties of Andhra Pradesh like 'Amrutham', 'Natupalli' and 'Palli' were separated from the rest of the genotypes and found to be more divergent. The dendrogram analysis revealed that clustering of varieties according to their utility type was only at subgroup level, and not along the major group level indicating low genetic differentiation among all the four utility types.

Genetic distance-based unrooted neighbor-joining tree (UNJ) (Fig. 2) drawn using DARwin was found to be in agreement with the UPGMA clustering method, in dividing all the genotypes into two major groups. Most of the clusters showed moderate to high bootstrap support (35–100 %). Numbers in the blue color indicate the bootstrap values. Genetic variability in the genotypes studied was evaluated by factorial correspondence analysis (FCA). The scatter plot based on the FCA (Fig. 3) analysis separated the individuals into different groups defined by the first two axes. Along the first axis 50 % of the juicy varieties were found to plot separately from the other type's. The other 50 % of the juicy types along with table, dual and pickle type varieties was observed along the second axis. Significantly, the UPGMA, FCA and UNJ tree analyses as well are in conformity with the same conclusion.

Population structure analysis

To infer the population structure of all the genotypes and assign individuals to populations, a model-based clustering method STRUCTURE was employed. The probability of data (highest likelihood) was maximum when the number of subpopulations (K) were at two (estimated Ln probability = −13,794) and four (estimated Ln probability = −13,319). Two largest values of ΔK, those for $k = 2$ and $k = 4$ more closely observed, assuming that only individuals with highest estimated membership coefficient belonged to a cluster. With $k = 2$, Cluster A comprised all the subpopulations (majority of table, dual, pickle and fewer juicy accessions) (Table 6). Cluster B corresponded to the juicy varieties (Fig. 4). When k was set at 4, four clusters were observed with juicy genotypes falling in all the clusters, but majority in cluster A and C. Cluster B was the smallest comprising seven juicy and four table type varieties. Cluster C was largely of juicy along with two table types. Despite admixtures of subpopulations, as observed in case of the cluster D all the pickle type varieties were confined to cluster D only. Most of the table and dual subpopulations were also defined to cluster D. Overall F_{st} for these four subpopulations was 0.11 indicating the genetic differentiation among populations was not high enough.

Discussion

Knowing the extent and structure of genetic variation in germplasm collections is essential for the conservation, utilization of biodiversity in any crop species and for efficient germplasm organization. Also, knowledge of the genetic diversity and population structure of germplasm collections is an important foundation for crop improvement. It is essential to first define the population structure within the germplasm to avoid spurious associations for performing association mapping studies (Flint-Garci et al. 2005). Bayesian clustering analyses have proven to be powerful tools to analyze population structure (Pritchard et al. 2000) and it is widely used in conservation biology to quantify relationships and differences among populations (Breton et al. 2008). Assessment of genetic diversity and population structure have been reported in horticulture crops such as apple (Urrestarazu et al. 2012), plum (Horvath et al. 2011), strawberry (Yoon et al. 2012), olive (Belaj et al. 2007), cucumber (Lv et al. 2012).

Several research groups have worked on diversity analysis of mango germplasm based on morphological traits (Mussane et al. 2010; Fitmawati et al. 2010; Begum et al. 2012) and molecular markers (Eiadthong et al. 1999; Rajwana et al. 2008; Yamanaka et al. 2006; Pandit et al. 2007; Luo et al. 2010). However, stray reports are available on development of mango-specific SSR markers (Viruel et al. 2005; Schnell et al. 2005; Duval et al. 2005; Honsho et al. 2005; Ravishankar et al. 2011) and characterization of mango germplasm using SSR markers (Galvez-Lopez et al. 2009; Vasugi et al. 2012). On the other hand, reports on population structure analysis in mango were very scarce both at international (Hirano et al. 2010; Dillona et al. 2013) and national level (Singh and Bhat 2009). In this communication, the present study was conceived to improve our understanding about the genetic diversity and structure of the mango population of Andhra Pradesh (India).

Microsatellite development

In the course of the present study, a genomic library enriched with $(CA)_n$ and $(TG)_n$ repeats has been constructed and 34 microsatellites were newly developed and of them 20 were polymorphic. Microsatellite enrichment procedure adapted using *Rsa*I library was highly successful, as evident from the fact that over 50 % of the clones had microsatellites. These results are in agreement with

Viruel et al. (2005) who have also reported high success rate of capturing microsatellite repeats using *Rsa*I library. Predominately the isolated SSRs were of dinucleotide repeats. Of them dinucleotide-$(AG)_n$ repeat containing motifs (18) was maximum followed by $(TG)_n$, $(AC)_n$, $(CT)_n$ and $(AT)_n$ repeats. Besides dinucleotide repeats, one trinucleotide repeat $(ACC)_n$ and two tetranucleotide $(TGTA)_n$, $(TGCA)_n$ repeats were also obtained. The finding that dinucleotide repeats were predominant and of them it was $(AG)_n$ repeat that was maximum, is in agreement with the earlier reports (Viruel et al. 2005; Duval et al. 2005; Honsho et al. 2005; Schnell et al. 2005). High percentage of dinucleotide-$(AG)_n$ repeat containing sequences have also been reported in other fruit crops like banana (Miller et al. 2010). Generally, it is assumed that SSRs with more the number of repeats, higher the probability of their being polymorphic. In contrast to this view, in the present study the SSR marker MGDSSR22, though having less repeat number, i.e., $(AC)_7$ showed more polymorphism than MGDSSR21 with more repeats of $(AG)_{18}$.

SSR diversity

The SSRs bring out polymorphism better than any other molecular marker because of the allelic diversity caused by replication slippage (Tautz and Renz 1984). In the present study also high level of SSR polymorphism (94.6 %) has been observed, wherein of the 318 bands generated 301 were polymorphic. This is as against 88 alleles generated in 28 mango varieties by Viruel et al. (2005) and 103 alleles detected by analyzing 241 mango genotypes by Singh et al. (2009). Relatively higher number of alleles (Na) detected in the present study in comparison to the earlier reports may be attributed to relatively large and diverse accessions used. Different levels of polymorphism ranging from 73 to 100 % using markers like RAPD (Ravishankar et al. (2000); Karihaloo et al. (2003); Rahman et al. (2007); Rajwana et al. (2008); Souza et al. (2011), ISSR polymorphism ranging from 85 to 97 % (Pandit et al. (2007); Samal et al. (2012); Tomar et al. (2011) AFLP ranging from 84 to 96 % (Yamanaka et al. (2006); Gálvez-López et al. (2010) and 73 % of polymorphism by SCoT markers (Luo et al. 2010) have been reported in earlier studies involving mango.

Markers with higher PIC values are said to possess greater potential to reveal allelic variation. The average PIC value of SSR markers tested by different researchers varies with number of SSR markers used and number of genotypes tested. In the course of the present investigation, average PIC (0.67) value obtained was higher than the values reported by Schnell et al. (2005) and Singh and Bhat (2009), and slightly lower than the average value reported by Hirano et al. (2010) (0.7), Ravishankar et al. (2011) (0.68), Vasugi et al. (2012) (0.76) and Dillona et al. (2013) (0.72). The average values of expected heterozygosity (0.39) detected in the present study were higher than those reported by Singh et al. (2009) (0.26), but lower than Viruel et al. (2005) (0.65), Hirano et al. (2010) (0.7) and Dillona et al. (2013) (0.75) which can be explained by the fact that the cultivars chosen in their studies were from different geographical regions like Florida (USA), Mexico, Australia, India and Southeast Asia. As against the maximum *Na* and *He* observed at locus MichR022 reported by Galvez-Lopez et al. (2009) in the Mexico germplasm and at the locus MiSHRS1 by Singh and Bhat (2009) in the Indian germplasm, in the present study it was the locus MiIIHR36 (i.e., SSR84) that exhibited the maximum *Na* and *He* indicating that this could be a valuable marker in genetic investigation of Indian mango germplasm.

SSR markers provide a reliable and reproducible approach for genotype-specific fingerprinting for cultivar identification. Unique alleles which were observed for nine juicy varieties can be used as molecular IDs in fingerprinting studies and such information could be used in assessment of genetic purity of the varieties. The presence of 'rare alleles' in juicy and table types represent a unique source of genetic diversity among the mango germplasm of Andhra Pradesh (India).

The overall F_{st} among the four populations (juicy, table, dual and pickle types) was 0.11 demonstrating weak population differentiation. F_{st} values greater than 0.25 indicate significant population differentiation (Hartl and Clark 1997). Similarly, Hirano et al. (2010) also reported weak population differentiation ($F_{st} = 0.12$) among the four mango populations from different geographical regions. AMOVA results indicate that the major proportion (62.25 %) of variation was exhibited within the individual which is not unexpected because mango is a highly cross-pollinated crop. Galvez-Lopez et al. (2009) detected that highest ratio of molecular variance and high genetic differentiation ($F_{st} = 0.19$) corresponded to mango populations within the states of Mexico. Singh and Bhat (2009) reported high variability (57.8 %) among the cultivars within populations of India using 18 SSR markers. RAPD analysis by Souza et al. (2011) revealed 72.8 % of variation within the populations of Brazil mangoes.

Cluster analysis

Based on UPGMA, UNJ and FCA analyses the test genotypes could be differentiated into two major groups and subgroups within each of them. In group I clustering of genotypes is according to their utility type at subgroup level only and not at major group. The group II predominantly constituted of the juicy germplasm (50 %) reveals a sort of geographical pattern, given the majority of them being

collections from North coastal districts of the state viz., Srikakulam, Vijayanagaram and East Godavari. Interestingly, among the hundreds of mango germplasm being maintained at the regional Fruit Research Station of the state Horticultural University at Sangareddy juicy genotypes included in the present study, also clustered in the group II. It is quite probable that most of the juicy germplasm found in the north coastal region of the state might have been from a common juicy genepool either by clonal or seed propagation over as long period of time. Group I seems to represent clusters of utility type rather than geographical pattern. Previous studies from India also show that mango varieties from different geographical zones though differ genetically, do not fall into distinct geographical region specific groups (Karihaloo et al. 2003; Pandit et al. 2007). Rocha et al. (2012) also have reported that there was no accession grouping according to sample locations in their genetic diversity study of 'Uba'' mango tree using ISSR markers. Observation of non differentiation of genotypes according to geographical pattern and admixtures in the populations in the current study can be attributed to more than one reason. Primarily, mango being a cross-pollinated crop, gene flow among the varieties, especially orchards situated in close proximity could have tinkered with the genetic purity of the given variety. Secondly, the earlier practice of propagation by seed as well as the desire of the orchardists to bring in types not belonging to a given region are equally important causal factors for exceptions observed. All the more, it is significant that popular varieties irrespective of the region of their existence remain true to type as evident from their molecular profile. If studies are undertaken by generating more mango-specific SSRs and applying them in diversity studies could reveal grouping of the genotypes according to their geographic regions and utility types. Also intervarietal variability could be assessed precisely leading to establishment of type/variety-specific molecular signature which could be valuable in certifying to the identity/quality of popular variety.

Very recently, Ramessur and Ranghoo (2011) have reported that pickle varieties of Mauritius cluster together in their study using RAPD analysis. All the pickle varieties analyzed in the current study also clustered together as revealed by FCA, UPGMA, UNJ and STRUCTURE analyses. However, number of pickle varieties used in the present study were not enough to understand their distinctiveness in the mango germplasm. Locally popular juicy varieties like 'Amrutham', and 'Palli', clustered distinctively from all other genotypes including the popular juicy varieties could be to some extent genetically tinkered ones either by mutations or cross-pollination with atypical juicy varieties. Though they do not cluster with typical juicy varieties 'Amrutham' and 'Palli' can be commercialized as juicy types and used in breeding programs.

Population structure analysis

To our knowledge, the present study is the first report using different utility types of mango of Andhra Pradesh (India) to analyze the population genetic structure in the germplasm using microsatellite generated data. The Bayesian model-based STRUCTURE analysis has enabled detection of four distinct clusters at $K = 4$. The results obtained by this analysis left room for alternate explanations for the kind of clustering pattern of mango germplasm. It can be inferred from Table 6, mango germplasm of the state exist in four distinct gene pools. Popular juicy varieties characterized by high fiber content and extensive commercial propagation of these promising varieties extensively in the state form one genepool (Cluster A). The juicy types collected from northern parts of the state form the second genepool (Cluster C) which is in congruent with the UPGMA analysis. The third (Cluster B) is the smallest cluster of very few juicy and table types. The fourth genepool (cluster D) representing all the pickle types, majority of the dual and table types and a few juicy types is the most admixtured cluster. Singh and Bhat (2009) reported structure analysis of 15 mango populations from different geographical regions of India resulted in admixturing and no distinct geographical differentiation among the populations. Similarly, Dillona et al. (2013) revealed admixtured clusters of five mango populations analyzed from different geographical regions indicating presence of ancestors in all the five clusters. Also, Hirano et al. (2010) reported no distinct differentiation of Myanmar accessions from other populations of India, Florida and southeast accessions.

In the current study, a close examination of the juicy component in all the gene pools leads to an interesting assumption that the table and pickle types might have derived from juicy types. This differentiation towards table and pickle types could have been due to gradual reduction in fiber content in juicy types largely through natural crossing of juicy types with table types and their propagation by seeds leading to human selection for desired quality.

In the present investigation, robustness and usefulness of microsatellite markers have proved their effectiveness in revealing the extent of genetic diversity of mango germplasm in general and of the state of Andhra Pradesh in particular. Also, SSRs developed in the study exhibited high polymorphism. To our knowledge, this study is the first report involving different utility types of mango of Andhra Pradesh (India) using high number of mango-specific SSRs available. SSR diversity analysis reported was highly informative and will be very useful in future applications like cultivar identification, DNA fingerprinting of a commercial varieties, selection of appropriate parents

for breeding programs, maintaining the genetic purity of germplasm by discarding the duplicates. Understanding the population structure of mango is a footstep which would benefit to make use of these germplasms in breeding programs as well as applying them in association mapping.

Acknowledgments Authors gratefully acknowledge the Department of Biotechnology (DBT), New Delhi, for the financial support. Thanks are due to the Institute of Biotechnology, Acharya N.G. Ranga Agricultural University, Hyderabad (India), for providing necessary laboratory facilities to carry out the research work. We are very thankful to Dr. Trushar Shah, ICRISAT, India, for the STRUCTURE analysis.

References

Anderson JA, Churchill GA, Autrique JE, Tanksley SD et al (1993) Optimizing parental selection for genetic linkage maps. Genome 36:181–186

Bally ISE (2011) Advances in research development in mango industry. Rev Bras Frutic 057–063

Begum H, Reddy TM, Malthi S, Reddy BP, Archack S, Nagaraju J, Siddiq EA (2012) Molecular analysis for genetic distinctiveness and relationships aof indigeneous landraces with popular cultivars of mango in Andhra Pradesh, India. The Asian and Australian Journal of Plant biotechnology 6(1):24–37

Belaj A, Munoz-Diez C, Baldoni L, Porceddu A, Barranco D, Satovic Z (2007) Genetic Diversity and Population Structure of wild olives from the North-western Mediterranean assessed by SSR markers. Ann Bot 100(3):449–458

Breton C, Pinatel C, Médail F, Bonhomme F, Bervillé A (2008) Comparison between classical and Bayesian methods to investigate the history of olive cultivars using SSR-polymorphisms. Plant Sci 175:524–532

Díaz-Matallana M, Schuler-García I, Ruiz-García M, Hodson de Jaramillo E (2009) Analysis of diversity among six populations of Colombian mango (*Mangifera indica* L. cvar. Hilacha) using RAPDs markers. Electron J Biotechnol 12(3), Available from: http://www.ejbiotechnology.info/content/vol12/issue3/full/10/index.html

Dillona NL, Ballya Ian SE, Wrighta CL, Hucksa L, Innesb DJ, Dietzgen RG (2013) Genetic diversity of the Australian National Mango Genebank. Sci Hortic 150:213–226

Duval MF, Bunel J, Sitbon C, Risterucci AM (2005) Devleopment of microsatellite markers for mango (*Mangifera indica* L.). Mol Ecol Notes 5(4):824–826

Eiadthong W, Yonemori K, Sugiura A, Utsunomiya N, Subhadrabandhu S (1999) Analysis of phylogenetic relationship in *Mangifera* by restriction site analysis of an amplified region of cpDNA. Science Horticulture 80:145–155

Eiadthong W, Yonemori K, Kanzaki S, Sugiura A, Utsunomiya N, Subhadrabandhu S (2000) Amplified fragment length polymorphism analysis for studying genetic relationships among Mangifera species in Thailand. J Am Soc Hortic Sci 125(2):160–164

Excoffier L, Guillaume L, Stefan S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol. Bioinform 1:47–50

FAOSTAT (2010) Available at: http://faostat.fao.org/site/339/default.aspx

Flint-Garci SA, Thuillet AC, Yu J, Pressoir G, Romer SM, Mitchell SE, Doebley J, Kresovich S, Goodman M, Buckler E (2005) Maize association population: a high-resolution platform for quantitative trait locus dissection. Plant J 44(6):1054–1064

Fitmawati AH, Bambang S, Purwoko (2010) Diversity of Indonesian mango (*Mangifera indica*) cultivars based on morphological and RAPD markers. SABRAO J Breed Genet 42(2):83–94

Galvez-Lopez D, Hernandez-Delgado S, González-Paz M, Becerra-Leor EN, Salvador-Figueroa M, Mayek-Pérez N (2009) Genetic analysis of mango landraces from Mexico based on molecular markers. Plant Genet Res: Character Util 7(3):244–251

Gálvez-López D, Salvador-Figueroa M, Becerra-Leor EN, González-Paz M, Hernández-Delgado S, Mayek-Pérez N (2010) Molecular diversity and genetic relationships of mango germplasm from chiapas, méxico. Agrociencia 44:907–915

Glenn TC, Schable NA (2005) Isolating microsatellite DNA loci. Methods Enzymol 395:202–222

Gosh SK, Dhau RS, MItra RS (1985) Studies on physiochemical characteristics of some mango cultivars grown at West Bengal. Indian food Packer 39(1):46–50

Gowda TND, Ramanjaneya KH (1994) Studies on physiochemical characteristics of some commercial cultivars of mango. Indian Food Packer 48:45

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

Hartl DL, Clark AG (1997) Principles of Population Genetics, 3rd edn. Sinauer Associates, Sunderland

He XH, Li YR, Guo YZ, Tang ZP, Li RB (2005) Genetic analysis of 23 mango cultivar collection in Guangxi province revealed by ISSR. Mol. Plant Breed. 3(6):829–834

Hemant Kumar NV, Narayanswamy P, Theertha Prasad D, Mukund GK, Sondur S (2001) Estimation of genetic diversity of commercial mango (*Mangifera indica* L.) cultivars using RAPD markers. J. Hort. Sci. and Biotech 76(5):529–533

Hirano R, Than H, Watanabe KN (2010) Myanmar mango landraces reveal genetic uniqueness over common cultivars from Florida, India, and Southeast Asia. Genome 53(4):321–330

Honsho C, Nishiyama K, Eiadthong W, Yonemori K (2005) Isolation and characterization of new microsatellites markers in mango (*Mangifera indica*). Mol Ecol Notes 5:152–154

Horvath A, Balsemin E, Barbot JC, Christmann H, Manzano G, Reynet P, Laigret F, Mariette S (2011) Phenotypic variability and genetic structure in plum (Prunus domestica L.), cherry plum (P. Cerasifera Ehrh.) and sloe (P. Spinosa L.). Sci Hortic 129:283–293

Hossaim MM, HaqueM RM, Rahman MH (2001) Physio-Morphological and compositional variation in Ripe fruit of three mango varieties. Online journal of Biological Sciences 1(11): 1101–1102

Indian Horticultural Database (2010) National Horticulture Board. Gurgoan, India

Karihaloo JL, Dwivedi YK, Archak S, Galkwab AB (2003) Analysis of genetic diversity of Indian Mango cultivars using RAPD markers. J. Hort. Sci. Biotechnol 78:285–289

Kashkush K, JingguiF TE, Hillel J, Lavi U (2001) Cultivar identification and genetic map of mango (*Mangifera indica*). Euphytica 122(1):129–136

Lopez VJA, Martinez O, Lopez OP (1997) Geographic differentiation and embryo type identification in *Mangifera indica* L cultivars using RAPD markers. Hort Sci 32(6):1105–1108

Luo C, He XH, Chen H, Ou SJ, Gao MP (2010) Analysis of diversity and relationships among mango cultivars using Start Codon Targeted (SCoT) markers. Biochem Syst Ecol 38(6):1176–1184

Lv J, Qi J, Shi Q, Shen D, Zhang S et al (2012) Genetic Diversity and Population Structure of Cucumber (Cucumis sativus L.). PLoS ONE 7(10):e46919. doi:10.1371/journal.pone.0046919

Miller NGR, Passos MAN, Natalia MPN, Souza MT, Marcos MC, Azevedo VR, Edson PA, Pappas GJ, Ciampi AY (2010) Short Report Characterization of novel microsatellite markers in Musa acuminata subsp.burmannicodes, var. calcutta 4. BMC Research Notes. 3(1):48

Mussane CRB, Biljon AV and Herselman L (2010) Morphological and genetic characterization of mango varieties in Mozambique. Second RUFORUM Biennial Meeting, Uganda, pp 991–995

Pandit SS, Mitra S, Giri AP, Pujari KH, Patil BP, Jambhale ND (2007) Genetic diversity analysis of mango cultivars using inter simple sequence repeat markers. Current Sci 93:1135–1141

Perrier X, Flori A, Bonnot F (2003) Data analysis methods. In: Hamon P, Seguin M, Perrier X, Glaszmann JC (eds) Genetic diversity of cultivated tropical plants. Science Publishers, Montpellier, pp 43–76

Porebski S, Grant Bailey L, Benard RB (1997) Modification of CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol Biol Report : 8–15

Pritchard JK, Stephens M, Donnelly P (2000) Inference of populationstructure using multilocus genotype data. Genetics 155: 945–959

Rahman ML, Rabbani MG, Siddique MNA, Rahman MA, Garvey EJ, Rahaman EHMS (2007) Molecular characterization of 28 mango germplasm using RAPD. Plant Tissue Cult. Biotechnol. 17(1): 71–77

Rajwana IA, Tabbasam N, Malik AU, Malik SA, Rahman M, Zafar Y (2008) Assessment of genetic diversity among mango (Mangifera indica L) genotypes using RAPD markers. Sci Hortic 117(3):297–301

Rajwana IA, Khan IA, Malik AU, Saleem BA, Khan AS, Khurram Z, Raheel A, Amin M (2011) Morphological and biochemical markers for varietal characterization and quality assessment of potential indigenous mango (*Mangifera indica*) germplasm. Int J Agri Biol 13(2):151–158

Ramessur AD, Ranghoo VMS (2011) RAPD Marker-assisted identification of genetic diversity among mango (*Mangifera indica*) varieties in Mauritius. Int. J. Agric. Biol. 13:167–173

Ravishankar KV, Lalitha A, Dinesh MR, Anand L (2000) Assessment of genetic relatedness among mango cultivars of India using RAPD markers. J. Hort. Sci. Biotech 75(2):198–201

Ravishankar KV, Hanumantha BM, Lalitha A, Dinesh MR (2011) Development of new microsatellite markers from Mango (*Mangifera indica*) and cross-species amplification. Am J Bot 98:96–99

Rocha A, Luiz CS, Tania FS, Cosme DC, de Dalmo LS (2012) Genetic diversity of 'Uba′ mango tree using ISSR markers. Mol Biotechnol 50(2):108–113

Rohlf FJ (1989) Numerical taxonomy and multivariate analysis system. Version 2.02i. Exeter Publishing, New York

Saleh MA, Zaied NS, Fouad AA, Nagiub MM, Fekria HK (2009) Assessment of some mango species by fruit characters and fingerprint. Aust J Basic Appl Sci 3:1920–1924

Samal KC, Jena RC, Swain SS, Das BK, Chand PK (2012) (2011) Evaluation of genetic diversity among commercial cultivars, hybrids and local mango (Mangifera indica L.) genotypes of India using cumulative RAPD and ISSR markers. Euphytica 185:195–213

Schnell RJ, Ronning CM, Knight RJ (1995) Identification of cultivars and validation of genetic relationship in (*Mangifera indica* L) using RAPD markers. Theor Appl Gent 90(2):269–274

Schnell RJ, Olano CT, Quintanilla WE, Meerow AW (2005) Isolation and characterization of 15 microsatellite loci from mango (*Mangifera indica* L.) and cross-species amplification in closely related taxa. Mol Ecol Notes 5:625–627

Sim SC, Robbins MD, VanDeynze A, Michel AP, Francis DM (2011) Population structure and genetic differentiation associated with breeding history and selection in tomato (*Solanum lycopersicum* L.). Heredity 106(6):927–935

Singh RN (1996) Mango. ICAR, NewDelhi

Singh S, Bhat KV (2009) Molecular characterization and analysis of geographical differentiation of Indian mango (*Mangifera indica* L.) germplasm. Acta Horticult 839:599–606

Singh S, Gaikwad AB, Karihaloo JL (2009) Morphological and molecular analysis of intra cultivar variation in Indian mango (Mangifera indica L.) cultivars. Acta Horti 829:205–212

Souza IGB, Valente SES, Britto FB, de Souza VAB, Lima PSC (2011) RAPD analysis of the genetic diversity of mango (Mangifera indica) germplasm in Brazil. Genet Mol Res 10(4):3080–3089

Srivastava AP, Chandra R, Rajan S, Ranade SA, Prasad V, Sangeeta S (2007) A PCR-based assessment of genetic diversity, and parentage analysis among commercial mango cultivars and hybrids. J Hortic Sci Biotechnol 82:951–959

Tautz D, Renz M (1984) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res 12:4127–4138

Tomar RS, Gajera HP, Viradiya RR, Patel SV, Golakiya BA (2011) Phylogenetic relationship among Mango (*Mangifera indica* L.) Landraces of Saurashtra based on DNA fingerprinting. J Hortic For 3(13):379–385

Urrestarazu J, Miranda C, Santesteban LG, Royo JB (2012) Genetic diversity and structure of local apple cultivars from Northeastern Spain assessed by microsatellite markers. Tree Genet Genom. doi:10.1007/s11295-012-0502-y

Utoskit K (2007) Development of microsatellite markers in mango (Mangifera indica L.) using 5′ anchored PCR. Thammasat Int J Sc Tech 12(3)

Vasugi C, Dinesh MR, Sekar K, Shivashankara KS, Padmakar B, Ravishankar KV (2012) Genetic diversity in unique indigenous mango accessions (Appemidi) of the Western Ghats for certain fruit characteristics. Curr Sci 103(2):199–207

Viruel M, Escribano P, Barbieri M, Ferri M, Hormaza J (2005) Fingerprinting, embryo type and geographic differentiation in mango (*Mangifera indica* L., Anacardiaceae) with microsatellites. Mol Breed 15(4):383–393

Yamanaka N, Hasran M, He-Xu D, Tsunematsu H, Idris S, Ban T (2006) Genetic relationships and diversity of four *Mangifera* species revealed through AFLP analysis. Genet Res Crop Evol 53:949–954

Yeh FC, Yan R, Boyle T (1999) A Microsoft Window based freeware for population genetic analysis. Version 1.31, University of Alberta, Canada

Yoon M-Y, Moe KT, Kim D-Y, Il-Rae Rho S, Kim K-TK, Won M-K, Chung J-W, Park Y-J (2012) Genetic diversity and population structure analysis of strawberry (*Fragaria* × *ananassa* Duch) using SSR markers. Electron J Biotechnol 15:2