Breeding Science 69: 332–344 (2019)
doi:10.1270/jsbbs.18204

Research Paper

Genetic diversity and relatedness of mango cultivars assessed by SSR markers

Shinsuke Yamanaka[1)], Fumiko Hosaka[2)], Masato Matsumura[3)], Yuko Onoue-Makishi[3)], Kenji Nashima[2)], Naoya Urasaki[4)], Tatsushi Ogata[1)], Moriyuki Shoda[4)] and Toshiya Yamamoto*[2)]

[1)] *Tropical Agriculture Research Front, Japan International Research Center for Agricultural Sciences*, 1091-1 Maezato-Kawarabaru, Ishigaki, Okinawa 907-0002, Japan
[2)] *Institute of Fruit Tree and Tea Science, NARO*, 2-1 Fujimoto, Tsukuba, Ibaraki 305-8605, Japan
[3)] *Okinawa Prefectural Agricultural Research Center Nago Branch*, 4605-3 Nago, Nago, Okinawa 905-0012, Japan
[4)] *Okinawa Prefectural Agricultural Research Center*, 820 Makabe, Itoman, Okinawa 901-0336, Japan

Assessment of genetic diversity and relatedness is an essential component of germplasm characterization and use. We analyzed 120 mango (*Mangifera indica* L.) genetic resources in Japan for their parentage, cultivar identification, genetic relatedness, and genetic diversity, using 46 polymorphic simple sequence repeat (SSR) markers. Ten sets of three SSR markers could successfully distinguish 83 genotypes with the exception of synonymous and identical accessions. We successfully assessed parentage, newly identifying or reconfirming both parents of 11 accessions, and revealing over 30 cultivars as offspring of 'Haden'. Genetic relatedness and diversity analyses revealed three distinct clusters. Two clusters correspond to the groups of USA and India, which are closely related. The other includes accessions from Southeast and East Asia. The results agree with the previous identification of genetically distinct Indian and Southeast Asian types, and suggest that the Florida accessions, which originated from hybrids between those two types, are more closely related to the Indian type.

Key Words: genetic diversity, genetic resources in Japan, *Mangifera indica*, mango, parentage.

Introduction

Mango (*Mangifera indica* L.) is a juicy stone fruit in the Anacardiaceae, which includes about 850 species of tropical fruit trees (Bompard 2009), and is an economically important cash crop produced about 40 Mt in 2012 (Mitra 2016). Mango is grown widely in the world's tropical and subtropical regions, as well as in a wide range of more marginal areas; India, China, Thailand, Mexico, Pakistan and Indonesia are the major producers (Mitra 2016). It is believed to have originated in the areas from India, where it has been grown for more than 4000 years and considered to be a primary center of diversity, to the Malay Peninsula in Southeast Asia.

More than 1000 mango cultivars exist around the world (Mukherjee 1953). They can be divided into two cultivar groups based on their embryo type: the monoembryonic (Indian) type is predominantly distributed in the subtropics, and the polyembryonic (Southeast Asian) type is most common in the tropics (Iyer and Degani 1997, Viruel *et al.* 2005). The polyembryony trait is dominant (Aron *et al.* 1998, Mukherjee and Litz 2009). The Indian type has a zygotic (sexually produced) embryo, and the fruit skin is mainly red, whereas the Southeast Asian type has several nucellar embryos (produced from the mother plant), and the skin is mainly green to yellow (Iyer and Degani 1997, Viruel *et al.* 2005).

During the 20th century, mango germplasms were introduced into Florida, USA, from the Caribbean Islands, Southeast Asia (the Philippines, Cambodia), India, and whole area extending from India to the Malay Peninsula, creating a secondary center of genetic diversity (Mukherjee and Litz 2009). In 1910, a seedling of 'Mulgoba' came into production in Florida, and the attractive selection was named 'Haden'. 'Eldon', 'Glenn', 'Lippens', 'Osteen', 'Parvin', 'Smith', 'Springfels', 'Tommy Atkins', and 'Zill' are considered to be progeny of 'Haden' (Campbell 1992). It is now estimated that most Florida cultivars are descended from only four monoembryonic Indian mango accessions 'Mulgoba', 'Sandersha', 'Amini', and 'Bombay' and the polyembryonic 'Turpentine' from the West Indies (Schnell *et al.* 2006). In the latter half of the 20th century, plantings of Florida cultivars have been established in many countries and now form the basis of the international mango trade (Mukherjee and Litz 2009).

Communicated by Hiroyuki Iketani
Received December 25, 2018. Accepted March 26, 2019.
First Published Online in J-STAGE on May 28, 2019.
*Corresponding author (e-mail: toshiya@affrc.go.jp)

Breeding Science
Vol. 69 No. 2 BS

Isozyme markers were initially used in a survey of genetic variation (Gan *et al.* 1981) and for the identification of cultivars (Degani *et al.* 1990). Schnell *et al.* (1995) used random amplified polymorphic DNA (RAPD) markers to fingerprint cultivars and estimate the genetic relationships among a group of putative 'Haden' seedlings. López-Valenzuela *et al.* (1997) used RAPD markers to estimate the genetic diversity of 15 mango cultivars and identified a specific RAPD band that was associated only with the polyembryonic type. Kashkush *et al.* (2001) used amplified-fragment-length polymorphic (AFLP) markers to estimate the genetic relationships among 16 cultivars and 7 rootstocks. These markers have been used to identify cultivars, evaluate their genetic relationships, and confirm that cross-pollination has occurred (Arias *et al.* 2012, Krishna and Singh 2007).

Simple sequence repeat (SSR), or microsatellite, markers have advantages over many other marker types: they are highly polymorphic, have multiple alleles, and are co-dominant. SSRs have been widely used for the conservation of genetic resources and in population genetics, molecular breeding, and paternity testing studies (Ellegren 2004). In mango, SSR markers are particularly important in the identification of cultivars, determination of genetic variability, conservation of germplasm, and identification of the domestication and movement of germplasm (Viruel *et al.* 2005). More than 100 SSR markers have been developed from various mango germplasms (Chiang *et al.* 2012, Dillon *et al.* 2014, Duval *et al.* 2005, Honsho *et al.* 2005, Ravishankar *et al.* 2011, Schnell *et al.* 2005, Viruel *et al.* 2005), and there are some studies on regional genetic diversity of mango using SSRs, e.g. Schnell *et al.* (2006) for Florida mango cultivars, Hirano *et al.* (2010) for Myanmar mango landraces, Tsai *et al.* (2013) for Taiwanese cultivars.

In Japan, cleaved amplified polymorphic sequence markers (Shudo *et al.* 2013) and retrotransposon-based insertion polymorphism markers (Nashima *et al.* 2017) were developed for marker-assisted selection and construction genetic linkage map in mango breeding program. Although these practical molecular tools have been developed, information of mango genetic resources in Japan is still meager.

To obtain the information for cultivar identification and diversity of Japanese mango genetic resources, in this study, we analyzed genetic diversity and relatedness of 120 accessions of mango which cover almost all mango collection in Japan, using 46 polymorphic SSR markers. Accurate parentages of many commercially grown cultivars were identified or reconfirmed. Phylogeographic relationships were discussed in comparison with previous studies.

Materials and Methods

Plant materials and DNA extraction

We analyzed 120 mango genetic resources held in Japan. They originated from the USA (Florida, Hawaii), Australia, Colombia, Egypt, Haiti, Honduras, India, Israel, Mexico, Panama, the Philippines, South Africa, Taiwan, Thailand, Trinidad and Tobago, Vietnam, and the West Indies (**Table 1**) (Campbell 1992, Hamilton 1993, Knight *et al.* 2009, Olano *et al.* 2005, Schnell *et al.* 2006). The origins of six accessions ('Barl', 'Khom-JIRCAS', 'Khom-OPARC', 'Mayer', 'Turpin', and 'Yu-Win #6-JIRCAS') are unknown. Eighty-three mango accessions were collected and maintained at the Japan International Research Center for Agricultural Sciences, Tropical Agriculture Research Front (JIRCAS, Ishigaki, Okinawa, Japan), and 37 accessions were at the Okinawa Prefectural Agricultural Research Center Nago Branch (OPARC, Nago, Okinawa, Japan).

Ninety-six F_1 individuals from the cross of 'Irwin' × 'Keitt' were used for evaluation of segregation of SSR genotypes. Plant materials were grown and maintained at the OPARC.

Genomic DNA was isolated from young leaves with a DNeasy Plant Mini Kit (Qiagen, Germany) according to the manufacturer's instructions.

SSR analysis

We preliminary tested 67 SSR markers that originated from mango. Of those, 21 were excluded because of no amplification, unstable amplification of the target band or the presence of monomorphic fragments. We used the remaining 46 SSR markers (**Table 2**), comprising 26 from Ravishankar *et al.* (2011), 6 from Schnell *et al.* (2005), and 14 from Viruel *et al.* (2005).

SSR markers were amplified in a 5-μL reaction mixture, containing 2.5 μL of Multiplex PCR Master Mix with HotStar Taq DNA Polymerase (Qiagen), 5 pmol of each primer (forward, fluorescently labeled with FAM or HEX; R, unlabeled), and 5 ng of genomic DNA. The PCR profile consisted of initial denaturation for 15 min at 95°C; 35 cycles of denaturation for 60 s at 94°C, annealing for 60 s at 55°C, and extension for 60 s at 72°C; and a final extension for 7 min at 72°C. The amplified PCR products were separated and detected in a PRISM 3130*xl* DNA sequencer (Applied Biosystems, USA). The sizes of the amplified bands were scored against internal standard DNA (400HD-ROX, Applied Biosystems) in GeneScan software (Applied Biosystems).

Data analysis

Using CERVUS v. 2.0 (Marshall *et al.* 1998) and MarkerToolKit v. 1.0 software (Fujii *et al.* 2008), we estimated the expected (H_E) and observed heterozygosity (H_O) at SSR marker loci in the cultivars. H_E was calculated from allele frequencies using an unbiased formula as $1 - \Sigma p_i^2 (1 \leq i \leq m)$, where m is the number of alleles at the target locus and p_i is the allele frequency of the ith allele at the target locus. H_O was calculated as the number of heterozygous individuals divided by the total number of individuals.

Parent–offspring relationships were tested by comparing the SSR alleles in each accession with those of its reported parents; the data were analyzed in MARCO software (Fujii

Table 1. Mango accessions used and their assessed parentage in this study

No.	Accession name	Origin (abbreviation)	Embryo-ny*	Source**	Accession nos.***	Parentage assessed by SSR markers in this study	Parantage from literatures****
1	Ah Ping	Hawaii, USA (HI)	M	JIRCAS	JTMG-001	offspring of Haden	
2	Ai	Taiwan (TW)	M	JIRCAS	JTMG-002	Lippens × Haden	
3	Alphonso	India (IN)	M	JIRCAS	JTMG-003		
4	Anderson	Florida, USA (FL)	M	JIRCAS	JTMG-004	offspring of Haden	Sandersha × Haden (d)
5	Bailey's Marvel	Florida, USA (FL)	M	JIRCAS	JTMG-005	offspring of Haden	Haden × Bombay (d)
6	Barl	unknown (?)	U	OPARC	Barl (OPARC)	Keitt × Tommy Atkins	
7	Becky-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-006	offspring of Haden	Haden × Brooks (d)
8	Becky-OPARC	Florida, USA (FL)	M	OPARC	Becky (OPARC)	offspring of Haden	Haden × Brooks (d)
9	Beverly	Florida, USA (FL)	M	JIRCAS	JTMG-007	offspring of Haden	offspring of Cushman (d)
10	Carabao	Philippines (PH)	P	JIRCAS	JTMG-008		
11	Carrie	Florida, USA (FL)	M	JIRCAS	JTMG-009		offspring of Julie (d)
12	Cat For Rock	Vietnam (VI)	U	JIRCAS	JTMG-010		
13	Choke Anan	Thailand (TH)	P	JIRCAS	JTMG-011		
14	Cushman	Florida, USA (FL)	M	OPARC	Cushman (OPARC)	offspring of Haden	Haden × Amini (d)
15	Dot-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-013	Carrie × Spirit of '76 (one discrepancy of LMMA11)	offspring of Zill (d)
16	Dot-OPARC	Florida, USA (FL)	M	OPARC	Dot (OPARC)	Carrie × Spirit of '76 (one discrepancy of LMMA11)	offspring of Zill (d)
17	Duncan	Florida, USA (FL)	M	JIRCAS	JTMG-014		offspring of Nam Doc Mai (d)
18	Edward-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-015	offspring of Haden	offspring of Haden (a, d)
19	Edward-OPARC	Florida, USA (FL)	M	OPARC	Edward (OPARC)	offspring of Haden	offspring of Haden (a, d)
20	Fahlan	Thailand (TH)	U	JIRCAS	JTMG-016		
21	Fairchild	Panama (PA)	U	OPARC	Fairchild (OPARC)	offspring of Alphonso	
22	Fascell	USA	M	JIRCAS	JTMG-017	Lippens × Haden	
23	Fukuda-JIRCAS	Hawaii, USA (HI)	M	JIRCAS	JTMG-018	offspring of Haden	
24	Fukuda-OPARC	Hawaii, USA (HI)	M	OPARC	Fukuda (OPARC)	offspring of Haden	
25	Glenn-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-019	offspring of Haden	offspring of Haden (a, b, d)
26	Glenn-OPARC	Florida, USA (FL)	M	OPARC	Glenn (OPARC)	offspring of Haden	offspring of Haden (a, b, d)
27	Golden Lippens-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-020	offspring of Lippens	offspring of Lippens (a, d)
28	Golden Lippens-OPARC	Florida, USA (FL)	M	OPARC	Golden Lippens (OPARC)	offspring of Lippens	offspring of Lippens (a, d)
29	Golden Nugget-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-021	offspring of Haden	offspring of Kent (d)
30	Golden Nugget-OPARC	Florida, USA (FL)	M	OPARC	Golden Nugget (OPARC)	offspring of Haden	offspring of Kent (d)
31	Gouviea	Hawaii, USA (HI)	U	JIRCAS	JTMG-023	offspring of Haden	
32	Graham	Trinidad Tobago (TT)	M	JIRCAS	JTMG-024		offspring of Julie (a)
33	Haden-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-027	offspring of Turpentine-JIRCAS	Mulgoba × Turpentine (a, b, d)
34	Haden-OPARC	Florida, USA (FL)	M	OPARC	Haden (OPARC)	offspring of Turpentine-JIRCAS	Mulgoba × Turpentine (a, b, d)
35	Hatcher	Florida, USA (FL)	M	JIRCAS	JTMG-028	offspring of Haden	Haden × Brooks (d)
36	Hodson	Florida, USA (FL)	M	JIRCAS	JTMG-029	offspring of Haden	offspring of Haden (d)
37	Honglong-JIRCAS	Taiwan (TW)	U	JIRCAS	JTMG-041	offspring of Irwin	
38	Honglong-OPARC	Taiwan (TW)	U	OPARC	Honglong (OPARC)	offspring of Irwin	
39	Irwin	Florida, USA (FL)	M	JIRCAS	JTMG-030	Lippens × Haden	Lippens × Haden (b, d)
40	Jacquelin-OPARC	Florida, USA (FL)	M	OPARC	Jacquelin (OPARC)	offspring of Haden or Pruter	Haden × Bombay (d)
41	Jacquelin-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-031	offspring of Haden	Haden × Bombay (d)
42	Jakarta	Florida, USA (FL)	M	JIRCAS	JTMG-032	offspring of Haden	Kent × Zill (d)
43	Jewel	Florida, USA (FL)	M	JIRCAS	JTMG-033		Lippens × Palmer (d)
44	Jinhuang-JIRCAS	Taiwan (TW)	U	JIRCAS	JTMG-040	White × Kent (one discrepancy of LMMA9)	
45	Jinhuang-OPARC	Taiwan (TW)	U	OPARC	Jinhuang (OPARC)	White × Kent (one discrepancy of LMMA9)	
46	Jinlong	Taiwan (TW)	U	OPARC	Jinlong (OPARC)	offspring of Irwin	
47	Jubilee	Florida, USA (FL)	M	JIRCAS	JTMG-034	Sensation × Irwin	Sensation × Irwin (d)
48	Keitt	Florida, USA (FL)	M	OPARC	Keitt (OPARC)	offspring of Haden	offspring of Brooks (b, d)
49	Keitt Red-JIRCAS	Taiwan (TW)	U	JIRCAS	JTMG-036	Irwin × Keitt	
50	Keitt Red-OPARC	Taiwan (TW)	U	OPARC	Keitt Red (OPARC)	Irwin × Keitt	
51	Kensington	Australia (AU)	P	JIRCAS	JTMG-037		
52	Kensington Pride	Australia (AU)	P	OPARC	Kensington Pride (OPARC)		
53	Kent	Florida, USA (FL)	M	JIRCAS	JTMG-038	offspring of Haden	Brooks × Haden (b, d)
54	Khom-JIRCAS	unknown (?)	U	JIRCAS	JTMG-039		
55	Khom-OPARC	unknown (?)	U	OPARC	Khom (OPARC)		
56	Lancetilla	Honduras (HN)	M	JIRCAS	JTMG-043		
57	Lily-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-044	Springfels × Sensation	Springfels × Sensation (d)
58	Lily-OPARC	Florida, USA (FL)	M	OPARC	Lily (OPARC)	Springfels × Sensation	Springfels × Sensation (d)
59	Lippens-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-045	offspring of Haden	offspring of Haden (a, d)
60	Lippens-OPARC	Florida, USA (FL)	M	OPARC	Lippens (OPARC)	offspring of Haden	offspring of Haden (a, d)
61	Madame Francis	Haiti (HT)	P	JIRCAS	JTMG-046		
62	Magshamim	Israel (IL)	M	JIRCAS	JTMG-047		

Table 1. (continued)

No.	Accession name	Origin (abbreviation)	Embryo-ny*	Source**	Accession nos.***	Parentage assessed by SSR markers in this study	Parantage from literatures****
63	Maha Chanok	Thailand (TH)	U	JIRCAS	JTMG-048		
64	Mallika	India (IN)	M	JIRCAS	JTMG-049	offspring of Neelumlate (one discrepancy of LMMA9)	Neelum × Dashehari (a, b)
65	Manilita	Mexico (MX)	P	JIRCAS	JTMG-050		
66	Manzanillo	Mexico (MX)	M	JIRCAS	JTMG-051	Haden × Kent	
67	Mapulehu	Florida, USA (FL)	M	JIRCAS	JTMG-052		offspring of Step (d)
68	Mayer	unknown (?)	M	JIRCAS	JTMG-053	offspring of Turpentine-JIRCAS	
69	Momi-K	Hawaii, USA (HI)	U	JIRCAS	JTMG-054	offspring of Haden	
70	N-13	Israel (IL)	U	OPARC	N-13 (OPARC)		
71	Nam Doc Mai #2-JIRCAS	Thailand (TH)	M	JIRCAS	JTMG-056		
72	Nam Doc Mai #2-OPARC	Thailand (TH)	M	OPARC	Nam Doc Mai #2 (OPARC)		
73	Nam Doc Mai #4-JIRCAS	Thailand (TH)	P	JIRCAS	JTMG-057		
74	Nam Doc Mai #4-OPARC	Thailand (TH)	P	OPARC	Nam Doc Mai #4 (OPARC)		
75	Naomi	Israel (IL)	M	JIRCAS	JTMG-058		offspring of Palmer (e)
76	Neelumlate	India (IN)	M	JIRCAS	JTMG-059		
77	Niku	Taiwan (TW)	U	JIRCAS	JTMG-060		
78	Oro	Mexico (MX)	M	JIRCAS	JTMG-061		
79	Osteen	Florida, USA (FL)	U	JIRCAS	JTMG-062	offspring of Haden	offspring of Haden (a, b, d)
80	Palmer	Florida, USA (FL)	M	JIRCAS	JTMG-063	offspring of Haden	offspring of Haden (b, d)
81	Paris	Hawaii, USA (HI)	P	JIRCAS	JTMG-064	offspring of Turpentine	
82	Parvin	Florida, USA (FL)	U	OPARC	Parvin (OPARC)	offspring of Haden	offspring of Haden (a)
83	Piva-JIRCAS	South Africa (ZA)	M	JIRCAS	JTMG-065		
84	Piva-OPARC	South Africa (ZA)	M	OPARC	Piva (OPARC)		
85	Pruter	Florida, USA (FL)	U	JIRCAS	JTMG-066	offspring of Haden	
86	R2E2	Australia (AU)	P	JIRCAS	JTMG-067	Kensington Pride × Kent	
87	Rad	Thailand (TH)	P	JIRCAS	JTMG-068		
88	Rapoza	Hawaii, USA (HI)	M	JIRCAS	JTMG-069	Irwin × Kent or offspring of Haden	
89	Ruby	Florida, USA (FL)	M	JIRCAS	JTMG-070	offspring of Haden	offspring of Haden (d)
90	S-01	Florida, USA (FL)	U	OPARC	S-01 (OPARC)	offspring of Haden	offspring of Haden (d)
91	Sensation	Florida, USA (FL)	M	JIRCAS	JTMG-071	offspring of Haden	Brooks × Haden (b, d)
92	Shiba	Taiwan (TW)	U	JIRCAS	JTMG-072		
93	Sonsien-JIRCAS	Taiwan (TW)	U	JIRCAS	JTMG-073		
94	Sonsien-OPARC	Taiwan (TW)	U	OPARC	Sonsien (OPARC)		
95	Spirit of '76-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-074	offspring of Haden	Zill × Haden (a, d)
96	Spirit of '76-OPARC	Florida, USA (FL)	M	OPARC	Spirit of '76 (OPARC)	offspring of Haden	Zill × Haden (a, d)
97	Springfels-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-075	offspring of Haden	offspring of Haden (a, d)
98	Springfels-OPARC	Florida, USA (FL)	U	OPARC	Springfels (OPARC)	offspring of Haden	offspring of Haden (a, d)
99	Tahar	Israel (IL)	M	JIRCAS	JTMG-076	offspring of Irwin	
100	Tainoung No. 1-JIRCAS	Taiwan (TW)	M	JIRCAS	JTMG-077		
101	Tainoung No. 1-OPARC	Taiwan (TW)	M	OPARC	Tainoung No. 1 (OPARC)		
102	Taiwan	Taiwan (TW)	U	JIRCAS	JTMG-078		
103	Tommy Atkins	Florida, USA (FL)	M	JIRCAS	JTMG-079	offspring of Haden	offspring of Haden (a, b, d)
104	Turpentine-JIRCAS	West Indies (WI)	P	JIRCAS	JTMG-081		
105	Turpentine-OPARC	West Indies (WI)	P	OPARC	Turpentine (OPARC)		
106	Turpin	unknown (?)	P	JIRCAS	not applicable		
107	Valencia Pride-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-082	offspring of Haden	offspring of Haden (a, d)
108	Valencia Pride-OPARC	Florida, USA (FL)	M	OPARC	Valencia Pride (OPARC)	offspring of Haden	offspring of Haden (a, d)
109	Vallenato	Colombia (CO)	P	JIRCAS	JTMG-083	offspring of Haden	
110	Van Dyke-JIRCAS	Florida, USA (FL)	M	JIRCAS	JTMG-084	offspring of Haden	offspring of Haden (b, d)
111	Van Dyke-OPARC	Florida, USA (FL)	M	OPARC	Van Dyke (OPARC)	offspring of Haden	offspring of Haden (b, d)
112	White-JIRCAS	Taiwan (TW)	P	JIRCAS	JTMG-085		
113	White-OPARC	Taiwan (TW)	P	OPARC	White (OPARC)		
114	White Pirie	Jamaica (JA)	P	JIRCAS	JTMG-086		
115	Yu-Win	Taiwan (TW)	U	JIRCAS	JTMG-025	offspring of Irwin	
116	Yu-Win #2	Taiwan (TW)	U	OPARC	Yu-Win #2 (OPARC)	Jinhuang × Irwin	
117	Yu-Win #6-JIRCAS	unknown (?)	U	JIRCAS	JTMG-026	Jinhuang × Irwin	Jinhuang × Irwin (c)
118	Yu-Win #6-OPARC	Taiwan (TW)	U	OPARC	Yu-Win #6 (OPARC)	Jinhuang × Irwin	Jinhuang × Irwin (c)
119	Zebda	Egypt (EG)	M	JIRCAS	JTMG-087		
120	Zillate	Florida, USA (FL)	M	JIRCAS	JTMG-088	offspring of Keitt	offspring of Keitt (d)

* M: monoembryony; P: polyembryony; U: unknown.

** JIRCAS: Japan International Research Center for Agricultural Sciences, Tropical Agriculture Research Front; OPARC: Okinawa Prefectural Agricultural Research Center Nago Branch.

*** Accessions of OPARC are maintained using cultivar name.

**** Parentage was described in literatures of a: Campbell (1992), b: Knight *et al.* (2009), c: Lee *et al.* (2009), d: Schnell *et al.* (2006), and e: Tomer *et al.* (1993).

Table 2. Characteristics of SSR markers applied for mango accessions

SSR loci	No. of alleles	H_E	H_O	References (Genbank accession nos.)
MiIIHR01	4	0.372	0.349	Ravishankar *et al.* (2011), EF592181
MiIIHR02	8	0.734	0.590	Ravishankar *et al.* (2011), EF592182
MiIIHR03	3	0.547	0.675	Ravishankar *et al.* (2011), EF592183
MiIIHR05	6	0.756	0.843	Ravishankar *et al.* (2011), EF592185
MiIIHR07	4	0.521	0.482	Ravishankar *et al.* (2011), EF592187
MiIIHR10	2	0.024	0.000	Ravishankar *et al.* (2011), EF592190
MiIIHR11	3	0.330	0.386	Ravishankar *et al.* (2011), EF592191
MiIIHR12	6	0.530	0.530	Ravishankar *et al.* (2011), EF592192
MiIIHR13	2	0.493	0.494	Ravishankar *et al.* (2011), EF592193
MiIIHR14	4	0.428	0.422	Ravishankar *et al.* (2011), EF592194
MiIIHR16	7	0.544	0.554	Ravishankar *et al.* (2011), EF592196
MiIIHR17	11	0.826	0.867	Ravishankar *et al.* (2011), EF592197
MiIIHR20	5	0.473	0.386	Ravishankar *et al.* (2011), EF592200
MiIIHR21	5	0.116	0.072	Ravishankar *et al.* (2011), EF592201
MiIIHR22	5	0.637	0.482	Ravishankar *et al.* (2011), EF592202
MiIIHR24	8	0.758	0.747	Ravishankar *et al.* (2011), EF592204
MiIIHR25	3	0.231	0.241	Ravishankar *et al.* (2011), EF592205
MiIIHR26	8	0.748	0.747	Ravishankar *et al.* (2011), EF592206
MiIIHR27	3	0.070	0.072	Ravishankar *et al.* (2011), EF592207
MiIIHR28	7	0.775	0.711	Ravishankar *et al.* (2011), EF592208
MiIIHR29	8	0.727	0.735	Ravishankar *et al.* (2011), EF592209
MiIIHR30	9	0.834	0.880	Ravishankar *et al.* (2011), EF592210
MiIIHR32	8	0.641	0.663	Ravishankar *et al.* (2011), EF592212
MiIIHR33	4	0.590	0.554	Ravishankar *et al.* (2011), EF592213
MiIIHR34	6	0.754	0.783	Ravishankar *et al.* (2011), EF592214
MiIIHR35	8	0.783	0.687	Ravishankar *et al.* (2011), EF592215
MiSHRS-4	4	0.661	0.711	Schnell *et al.* (2005), AY942818
MiSHRS-26	2	0.193	0.217	Schnell *et al.* (2005), AY942821
MiSHRS-29	5	0.560	0.590	Schnell *et al.* (2005), AY942822
MiSHRS-32	7	0.535	0.482	Schnell *et al.* (2005), AY942824
MiSHRS-33	5	0.355	0.434	Schnell *et al.* (2005), AY942825
MiSHRS-39	7	0.616	0.639	Schnell *et al.* (2005), AY942829
LMMA1	9	0.834	0.880	Viruel *et al.* (2005), AY628373
LMMA2	7	0.650	0.458	Viruel *et al.* (2005), AY628374
LMMA4	5	0.663	0.554	Viruel *et al.* (2005), AY628376
LMMA5	3	0.307	0.289	Viruel *et al.* (2005), AY628377
LMMA6	11	0.694	0.735	Viruel *et al.* (2005), AY628378
LMMA7	6	0.716	0.687	Viruel *et al.* (2005), AY628379
LMMA8	9	0.747	0.747	Viruel *et al.* (2005), AY628380
LMMA9	7	0.806	0.711	Viruel *et al.* (2005), AY628381
LMMA10	11	0.799	0.880	Viruel *et al.* (2005), AY628382
LMMA11	6	0.764	0.735	Viruel *et al.* (2005), AY628383
LMMA12	7	0.713	0.747	Viruel *et al.* (2005), AY628384
LMMA14	4	0.400	0.301	Viruel *et al.* (2005), AY628386
LMMA15	6	0.561	0.566	Viruel *et al.* (2005), AY628387
LMMA16	6	0.748	0.843	Viruel *et al.* (2005), AY628388
Average	6.0	0.577	0.569	

et al. 2010). Minimal Marker software (Fujii *et al.* 2013) was used to identify minimal marker subsets needed to distinguish all cultivars and to find identical genotypes generated from the 46 SSR markers among the 120 accessions.

A phenogram of the 120 accessions was constructed by using the unweighted pair-group method with arithmetic mean (UPGMA) based on the similarities between genotypes estimated by Dice's coefficient: $D_c = 2n_{xy}/(n_x + n_y)$, where n_x and n_y represent the number of putative SSR alleles for materials X and Y, and n_{xy} represents the number of putative SSR alleles shared between X and Y. The phenogram was drawn in NTSYS-pc v. 2.1 software (Rohlf 1998).

To survey genetic diversity, we calculated the genetic distance between accessions from the allele size of each SSR locus in GenAlEx v. 6.5 software (Peakall and Smouse 2012). Principal coordinates analysis (PCoA) based on genetic distance was conducted in GenAlEx 6.5.

To analyze population structure, we applied a Bayesian model clustering algorithm to microsatellite data to infer genetic structure and to define the number of clusters in STRUCTURE v. 2.3.4 software (Pritchard *et al.* 2000), using an admixture model for ancestry and an independent model for allele frequency, without any prior information about the origin of samples. For each value of K (number of inferred ancestral populations) from 2 to 10, analyses were performed 10 times with 100 000 iterations after a burn-in period of 100 000 iterations. ΔK was used to estimate the appropriate K value according to the criterion of Evanno *et al.* (2005).

Segregation of SSR alleles were evaluated for 46 SSR loci used in this study to validate if each SSR is derived from single locus or multiple ones, by using 96 F_1 individuals obtained from the cross of 'Irwin' × 'Keitt'. JoinMap ver. 4.1 software (Kyazma B.V., the Netherlands; Van Ooijen 2011) was used. We also picked up significant linkages between two SSR loci for alleles of 'Irwin' as well as 'Keitt', calculated by JoinMap ver. 4.1 software.

Results

Genetic identification of mango accessions using SSR markers

We identified 274 putative alleles in the 120 accessions (**Table 2**). The number of alleles per locus ranged from 2 at 3 of the loci (MiIIHR10, MiIIHR13, MiSHRS-26) to 11 at 3 of the loci (MiIIHR17, LMMA6, LMMA10), with an average value of 6.0 (**Table 2**). H_E ranged from 0.024 at MiIIHR10 to 0.834 at MiIIHR30 and LMMA1, with an average value of 0.577. H_O ranged from 0 at MiIIHR10 to 0.880 at MiIIHR30, LMMA1, and LMMA10, with an average value of 0.569. The 120 accessions could be differentiated and classified into 83 genotypes excluding identical accessions by the 46 SSR markers (**Fig. 1**).

Thirty groups showing identical SSR genotypes were found in this study (**Table 3**). Twenty-three out of 30 groups included accessions with the same names maintained at different organizations, JIRCAS and OPARC. On the other hand, 13 groups included synonymous accessions. For example, three accessions ('Ai', 'Fascell', and 'Irwin') were identified as the same genotype 1. Similarly, 'Bailey's Marvel' vs. 'Beverly' (Genotype 2), 'Duncan' vs. 'Nam Doc Mai #2-JIRCAS' (Genotype 5), 'Gouviea' vs. 'Momi-K' (Genotype 11), 'Haden-JIRCAS' vs. 'Mayer' (Genotype 12), 'Honglong-JIRCAS' vs. 'Jinlong' (Genotype 13), 'Jakarta' vs. 'Valencia Pride-JIRCAS' (Genotype 14), 'Kensington' vs. 'Kensington Pride' (Genotype 17), 'Nam Doc Mai #4-JIRCAS' vs. 'Turpin' (Genotype 21), 'Nam Doc Mai #4-OPARC' vs. 'Paris' (Genotype 22), 'Osteen' vs. 'Springfels-OPARC' (Genotype 23), 'White-JIRCAS' vs. 'White Pirie' (Genotype 29), and 'Yu-Win #2' vs. 'Yu-Win

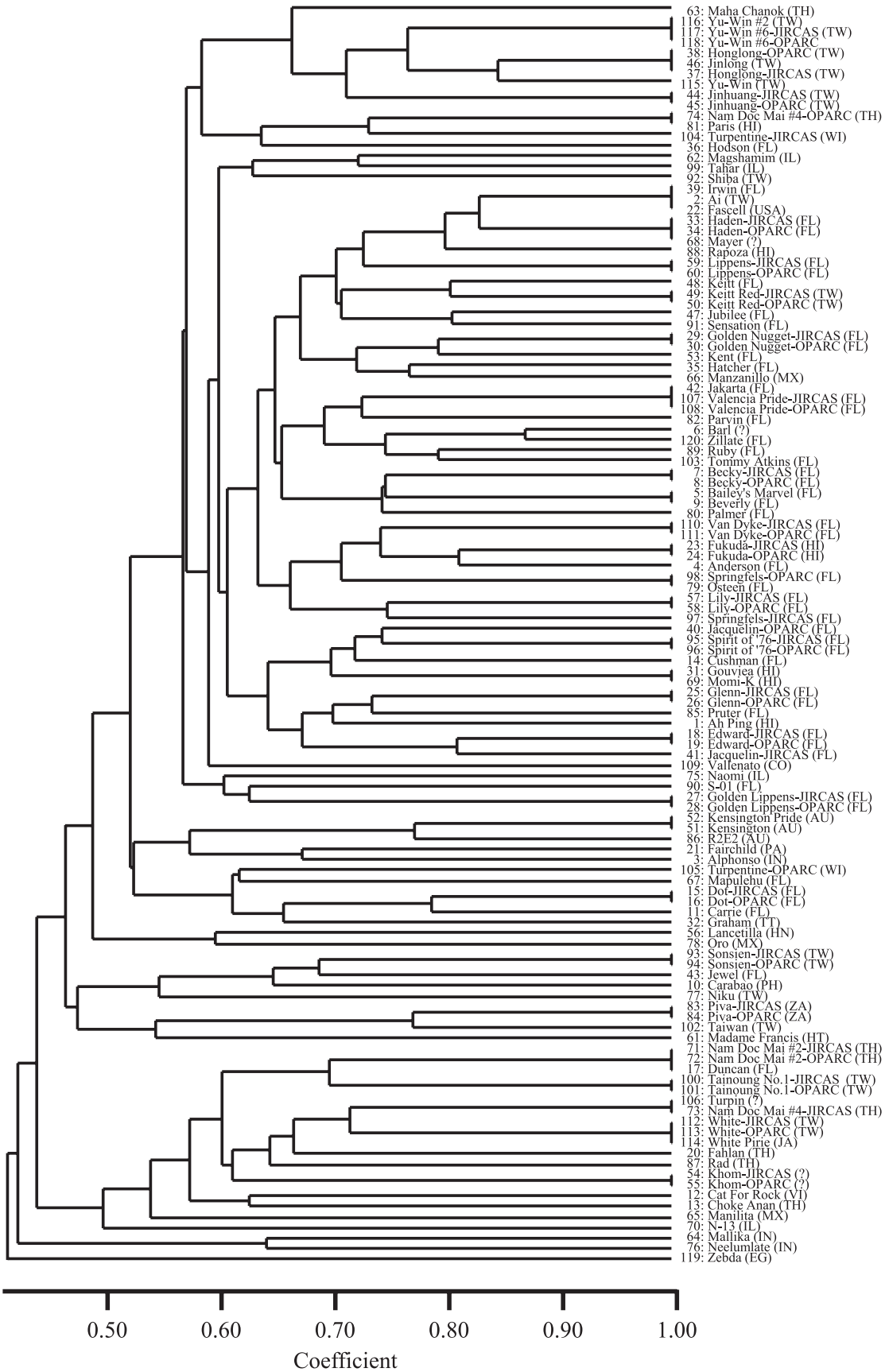

Fig. 1. Phenogram of the 120 mango genetic resources evaluated. The phenogram was produced using the UPGMA method based on Dice's coefficient. Origins of accessions are indicated as two-letter ISO 3166 codes or US state abbreviations; "?" = unknown.

Table 3. Mango accessions showing identical genotypes

Genotype	Accession name (Code No.)*
1	Ai (2), Fascell (22), Irwin (39)
2	Bailey's Marvel (5), Beverly (9)
3	Becky-JIRCAS (7), Becky-OPARC (8)
4	Dot-JIRCAS (15), Dot-OPARC (16)
5	Duncan (17), Nam Doc Mai #2-JIRCAS (71), Nam Doc Mai #2-OPARC (72)
6	Edward-JIRCAS (18), Edward-OPARC (19)
7	Fukuda-JIRCAS (23), Fukuda-OPARC (24)
8	Glenn-JIRCAS (25), Glenn-OPARC (26)
9	Golden Lippens-JIRCAS (27), Golden Lippens-OPARC (28)
10	Golden Nugget-JIRCAS (29), Golden Nugget-OPARC (30)
11	Gouviea (31), Momi-K (69)
12	Haden-JIRCAS (33), Haden-OPARC (34), Mayer (68)
13	Honglong-JIRCAS (37), Honglong-OPARC (38), Jinlong (46)
14	Jakarta (42), Valencia Pride-JIRCAS (107), Valencia Pride-OPARC (108)
15	Jinhuang-JIRCAS (44), Jinhuang-OPARC (45)
16	Keitt Red-JIRCAS (49), Keitt Red-OPARC (50)
17	Kensington (51), Kensington Pride (52)
18	Khom-JIRCAS (54), Khom-OPARC (55)
19	Lily-JIRCAS (57), Lily-OPARC (58)
20	Lippens-JIRCAS (59), Lippens-OPARC (60)
21	Nam Doc Mai #4-JIRCAS (73), Turpin (106)
22	Nam Doc Mai #4-OPARC (74), Paris (81)
23	Osteen (79), Springfels-OPARC (98)
24	Piva-JIRCAS (83), Piva-OPARC (84)
25	Sonsien-JIRCAS (93), Sonsien-OPARC (94)
26	Spirit of '76-JIRCAS (95), Spirit of '76-OPARC (96)
27	Tainoung No. 1-JIRCAS (100), Tainoung No. 1-OPARC (101)
28	Van Dyke-JIRCAS (110), Van Dyke-OPARC (111)
29	White-JIRCAS (112), White-OPARC (113), White Pirie (114)
30	Yu-Win #2 (116), Yu-Win #6-JIRCAS (117), Yu-Win #6-OPARC (118)

* Representative accessions of identical genotypes group were indicated underlined.

#6-JIRCAS' (Genotype 30), showed identical SSR genotypes (**Table 3**). These synonymous accessions should be carefully identified by using genetic resources maintained at the different organizations. One representative accession was chosen from each genotype group by taking into account the record of introduction background of each genetic resources such as passport data, and used for further analysis.

Out of 27 homonymous cultivars maintained in both JIRCAS and OPARC with same cultivar name, four cultivar sets ('Jacquelin', 'Nam Doc Mai #4', 'Springfels', 'Turpentine') showed different SSR genotypes between the two organizations. These accessions should be treated and inventoried according to the introduction record, passport data, phenotypic traits data and so on.

Ten sets of three markers (e.g., MiIIHR02, MiSHRS-4, and LMMA1, **Supplemental Data 1a**) were enough to distinguish all 83 representative accessions (83 genotypes) on the basis of at least one difference in SSR genotype identified by Minimal Marker software (Fujii *et al.* 2013). Furthermore, 124 marker subsets consisting of five SSR markers each (e.g., MiIIHR02, MiIIHR17, MiIIHR24, MiIIHR28, and MiIIHR30, **Supplemental Data 1b**) could differentiate all 83 representative accessions on the basis of two or more differences.

Parentage analysis

We analyzed the parentages of the 120 accessions by using 274 putative alleles at 46 polymorphic SSR loci. Many accessions were identified as offspring of 'Haden-JIRCAS' crossed with unidentified cultivars not tested in this study ('Ah Ping', 'Anderson', 'Bailey's Marvel', 'Becky-JIRCAS', 'Cushman', 'Edward-JIRCAS', 'Fukuda-JIRCAS', 'Glenn-JIRCAS', 'Golden Nugget-JIRCAS', 'Gouviea', 'Hatcher', 'Hodson', 'Jacquelin-OPARC', 'Jacquelin-JIRCAS', 'Keitt', 'Kent', 'Lippens-JIRCAS', 'Osteen', 'Palmer', 'Parvin', 'Pruter', 'Ruby', 'S-01', 'Sensation', 'Spirit of '76-JIRCAS', 'Springfels-JIRCAS', 'Tommy Atkins', 'Valencia Pride-JIRCAS', 'Vallenato', 'Van Dyke-JIRCAS'; **Table 1**). The results revealed both parents of 11 accessions: 'Barl' ('Keitt' × 'Tommy Atkins'), 'Dot-JIRCAS' ('Carrie' × 'Spirit of '76-JIRCAS', except for one discrepancy at LMMA11), 'Irwin' ('Lippens-JIRCAS' × 'Haden-JIRCAS'), 'Jinhuang-JIRCAS' ('White-JIRCAS' × 'Kent', except for one discrepancy at LMMA9), 'Jubilee' ('Sensation' × 'Irwin'), 'Keitt Red-JIRCAS' ('Irwin' × 'Keitt'), 'Lily-JIRCAS' ('Springfels-JIRCAS' × 'Sensation'), 'Manzanillo' ('Haden-JIRCAS' × 'Kent'), 'R2E2' ('Kensington' × 'Kent'), 'Rapoza' ('Irwin' × 'Kent' or offspring of 'Haden-JIRCAS'), and 'Yu-Win #6-JIRCAS' ('Jinhuang-JIRCAS' × 'Irwin') (**Table 1**). The single discrepancies in 'Dot-JIRCAS' and 'Jinhuang-JIRCAS' may be due to allele mutations. Since there were no discrepancies at the other 45 SSR loci, we assumed that the parentages of 'Dot-JIRCAS' and 'Jinhuang-JIRCAS' were correct.

Genetic relatedness

We constructed a phenogram of the 120 accessions based on SSR analysis (**Fig. 1**). Many accessions from Florida were grouped in the upper part of the phenogram, while accessions from India ('Alphonso', 'Mallika', 'Neelumlate'), Thailand ('Choke Anan', 'Fahlan', 'Nam Doc Mai #2-JIRCAS', 'Nam Doc Mai #4-JIRCAS', 'Rad'), Vietnam ('Cat For Rock'), and Egypt ('Zebda') were grouped in the lower part. Nevertheless, the accessions were mingled.

Genetic diversity of mango genetic resources

For further genetic diversity analyses to characterize mango genetic resources in Japan, we also employed 83 independent accessions selected by SSR genotyping in this study as a representative collection in Japan. As for the PCoA, the first and second principal components explained 14.25% and 7.17% of the variation, respectively. Overall, all 83 accessions distributed sparsely on the scatter plot, suggesting that genetic resources in Japan possess a certain level of genetic diversity in terms of SSR variation. Based on their origin, it was revealed that they tended to form three groups: "USA", "India", and "Thailand, Taiwan, the Philippines and Vietnam" (**Fig. 2**), in contrast to the UPGMA

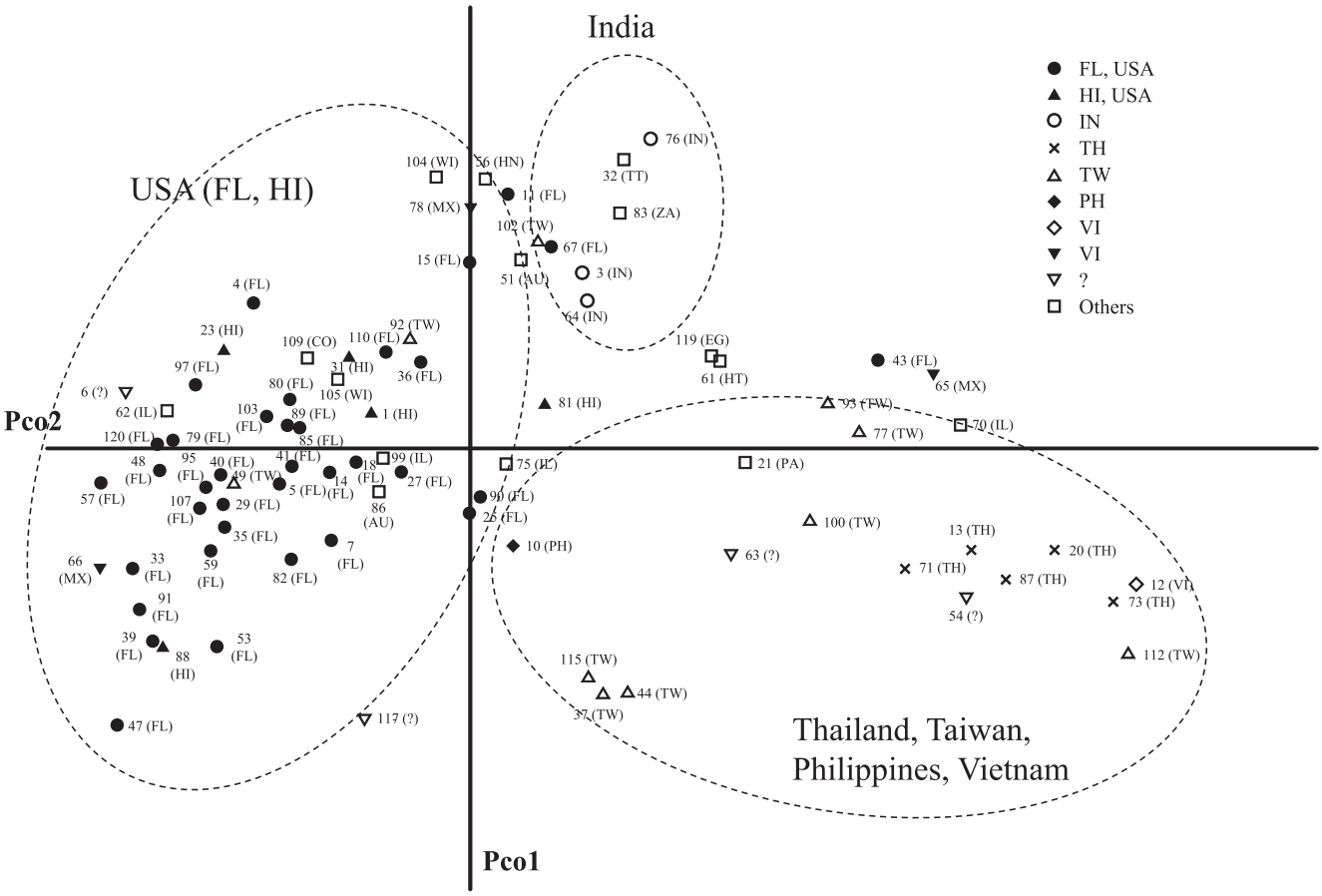

Fig. 2. Scatter plot of 83 mango genetic resources based on principal coordinates analysis. For accession numbers, see **Table 1**. Origins of accessions are indicated as two-letter ISO 3166 codes or US state abbreviations; "?" = unknown.

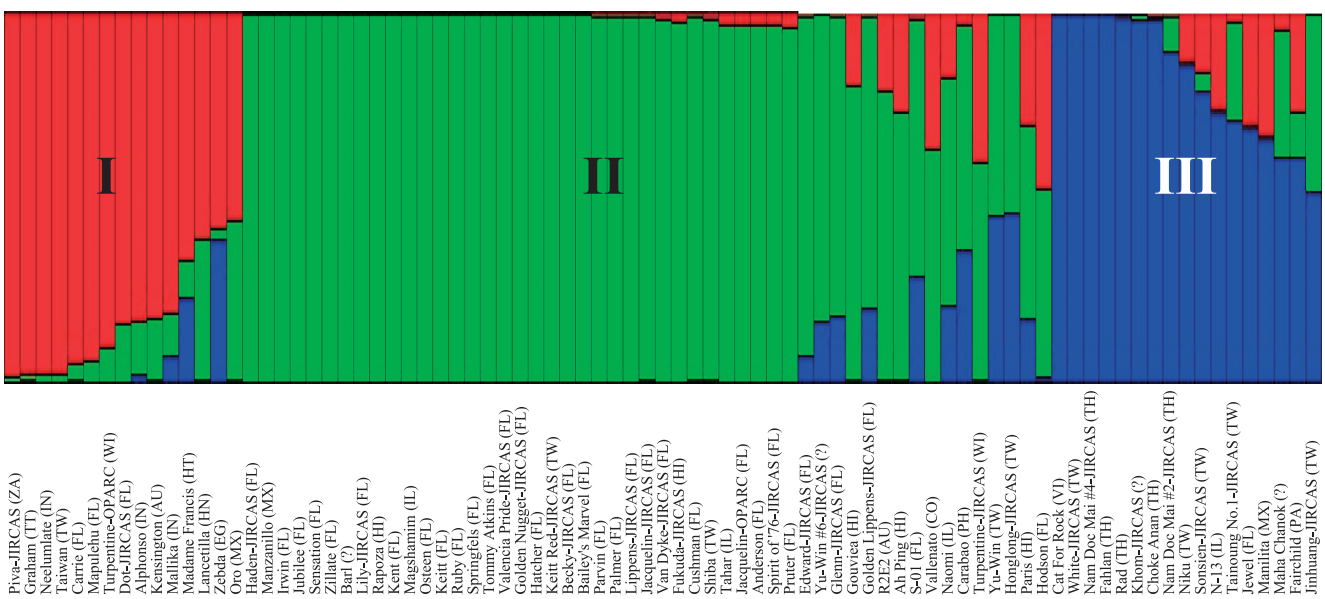

Fig. 3. Bar plot of 83 mango genetic resources by structure analysis ($K = 3$) with 46 SSR loci. Origins of accessions are indicated as two-letter ISO 3166 codes or US state abbreviations; "?" = unknown.

phenogram (**Fig. 1**).

In the analysis of population structure, ΔK showed a maximum at $K = 3$, suggesting three genetically distinct clusters (I, II, and III in **Fig. 3**). Cluster I included accessions from India ('Alphonso', 'Mallika', and 'Neelumlate'), suggesting that typical Indian type accessions were included.

Cluster II included predominantly US accessions from Florida and Hawaii. Cluster III included mostly Asian accessions from Thailand, Vietnam, and Taiwan, in which accessions of Southeast Asian type were predominant. These clusters were generally consistent with the groups obtained from PCoA as mentioned above. As for the relationship between population structure and embryo types of the seed, monoembryonic accessions were predominant in clusters I and II, showing a relationship between embryony and cultivar clusters identified by population structure analysis (**Supplemental Fig. 1**). Polyembryonic accessions were predominant in cluster III and also featured in cluster II.

Segregation of SSR loci

In order to characterize whether SSR alleles were derived from single locus or multiple loci used in this study, segregations of SSR genotypes were evaluated by using 96 F_1 individuals obtained from the cross of 'Irwin' × 'Keitt' (**Table 4**). Thirty-five SSR loci showed segregations of SSR genotypes in the 96 F_1 individuals of 'Irwin' × 'Keitt', whereas no segregation was observed for 11 SSR loci. Eighteen SSR loci showed binary segregations (a/a: a/b, a/c: b/c, a/b: a/c), and 17 of them fitted to the expected segregation ratio of 1:1, whereas only one SSR locus MiIIHR13 showed skewed segregation at 5% level. Out of the 13 SSR

Table 4. Segregation of SSR genotypes for 96 F_1 plants from Irwin × Keitt

SSR loci	SSR genotypes of Irwin (bp)	SSR genotypes of Keitt (bp)	Segregation for F_1 hybrids of Irwin × Keitt	Expected ratio	chi-square value	Signif.
MiIIHR01	252/252	246/252	246/252:252/252 = 48:48	1:1	0.00	ns
MiIIHR02	171/175	175/189	171/175:171/189:175/175:175/189 = 26:24:23:23	1:1:1:1	0.25	ns
MiIIHR03	235/235	235/236	235/235:235/236 = 47:49	1:1	0.04	ns
MiIIHR05	209/216	209/215	209/209:209/215:209/216:215/216 = 26:18:29:23	1:1:1:1	2.75	ns
MiIIHR07	170/170	170/174	170/170:170/174 = 51:45	1:1	0.38	ns
MiIIHR10	190/190	190/190	no segregation			
MiIIHR11	221/221	212/221	212/221:221/221 = 55:41	1:1	2.04	ns
MiIIHR12	177/177	177/177	no segregation			
MiIIHR13	190/197	197/197	190/197:197/197 = 39:57	1:1	3.38	*
MiIIHR14	354/354	342/354	342/354:354/354 = 54:42	1:1	1.50	ns
MiIIHR16	208/208	208/208	no segregation			
MiIIHR17	244/274	244/276	244/244:244/274:244/276:274/276 = 17:27:18:34	1:1:1:1	8.08	**
MiIIHR20	190/190	190/190	no segregation			
MiIIHR21	239/239	239/239	no segregation			
MiIIHR22	227/241	234/241	227/234:227/241:234/241:241/241 = 19:20:26:31	1:1:1:1	3.92	ns
MiIIHR24	247/247	247/252	247/247:247/252 = 42:54	1:1	1.50	ns
MiIIHR25	151/151	151/151	no segregation			
MiIIHR26	145/164	149/151	145/149:145/151:149/164:151/164 = 26:25:24:21	1:1:1:1	0.58	ns
MiIIHR27	197/197	197/197	no segregation			
MiIIHR28	112/120	114/120	112/114:112/120:114/120:120/120 = 22:27:27:20	1:1:1:1	1.58	ns
MiIIHR29	157/157	153/161	153/157:157/161 = 56:40	1:1	2.67	ns
MiIIHR30	202/204	198/202	198/202:198/204:202/202:202/204 = 27:24:16:29	1:1:1:1	4.08	ns
MiIIHR32	188/190	190/190	188/190:190/190 = 40:56	1:1	2.67	ns
MiIIHR33	180/180	168/180	168/180:180/180 = 52:44	1:1	0.67	ns
MiIIHR34	236/246	243/246	not tested			
MiIIHR35	193/201	201/201	193/201:201/201 = 43:53	1:1	1.04	ns
MiSHRS-4	135/139	133/139	133/135:133/139:135/139:139/139 = 24:20:24:28	1:1:1:1	1.33	ns
MiSHRS-26	281/281	281/284	281/281:281/284 = 51:45	1:1	0.38	ns
MiSHRS-29	186/188	186/188	186/186:186/188:188/188 = 26:49:21	1:2:1	0.56	ns
MiSHRS-32	211/211	207/211	207/211:211/211 = 40:56	1:1	2.67	ns
MiSHRS-33	254/257	254/257	254/254:254/257:257/257 = 24:48:24	1:2:1	0.00	ns
MiSHRS-39	374/374	359/374	374/374:374/359 = 47:49	1:1	0.04	ns
LMMA1	208/210	206/208	206/208:206/210:208/208:208/210 = 24:25:20:27	1:1:1:1	1.08	ns
LMMA2	285/297	285/295	285/285:285/295:285/297:295/297 = 30:19:26:21	1:1:1:1	3.08	ns
LMMA4	237/237	231/247	231/237:237/247 = 56:40	1:1	2.67	ns
LMMA5	288/288	288/288	no segregation			
LMMA6	112/131	112/131	112/112:112/131:131/131 = 23:48:25	1:2:1	0.08	ns
LMMA7	206/206	206/212	206/206:206/212 = 53:43	1:1	1.04	ns
LMMA8	263/263	263/263	no segregation			
LMMA9	178/188	178/178	178/178:178/188 = 44:52	1:1	0.67	ns
LMMA10	162/181	177/181	162/177:162/181:177/181:181/181 = 16:20:29:31	1:1:1:1	6.42	*
LMMA11	238/246	238/255	238/238:238/246:238/255:246/255 = 20:26:28:22	1:1:1:1	1.67	ns
LMMA12	211/211	207/211	207/211:211/211 = 48:48	1:1	0.00	ns
LMMA14	177/177	177/177	no segregation			
LMMA15	217/225	217/225	217/217:217/225:225/225 = 32:46:18	1:2:1	4.25	ns
LMMA16	240/245	245/250	240/245:240/250:245/245:245/250 = 20:20:21:35	1:1:1:1	6.75	*

* and ** showed distortion at 5% and 1% level.

Table 5. Significant linkages between SSR loci for Irwin

SSR locus 1	SSR locus 2	Recombination frequency	LOD score
MiIIHR05	MiIIHR26	0.031	23.06
MiIIHR17	MiIIHR32	0.094	15.55
MiSHRS-4	LMMA2	0.115	14.07
MiIIHR22	LMMA10	0.156	10.10

Table 6. Significant linkages between SSR loci for Keitt

SSR locus 1	SSR locus 2	Recombination frequency	LOD score
MiIIHR14	MiIIHR24	0.000	28.57
MiIIHR14	LMMA16	0.052	20.05
MiIIHR24	LMMA16	0.052	20.05
MiIIHR01	MiSHRS-39	0.073	18.04
MiIIHR05	MiIIHR26	0.094	16.29
MiIIHR07	LMMA12	0.094	16.09
MiIIHR02	MiSHRS-32	0.115	15.45
MiSHRS-4	LMMA2	0.125	13.13
MiIIHR22	LMMA10	0.146	11.51
MiIIHR29	LMMA11	0.208	7.63
MiIIHR29	MiIIHR33	0.229	6.31

loci showing 1:1:1:1 segregations (a/c: a/d: b/c: b/d, a/a: a/c: a/b: b/c), ten SSR loci fitted to the expected segregation ratio of 1:1:1:1. Two SSRs (LMMA10 and LMMA16) showed distorted segregation at 5% level, and one SSR (MiIIHR17) showed distorted segregation at 1% level. All four SSR loci showing a/a: a/b: b/b segregation fitted to the expected segregation ratio of 1:2:1. These results indicated that almost all SSR loci used in this study were derived from single locus, which can lead to evaluate considerably exact genetic diversity and relatedness of mango cultivars.

Significant linkages between SSR loci were also evaluated for alleles of 'Irwin' (**Table 5**). Four SSR combinations (MiIIHR05 vs. MiIIHR26, MiIIHR17 vs. MiIIHR32, MiSHRS-4 vs. LMMA2, and MiIIHR22 vs. LMMA10) showed significant linkages of 0.031 to 0.156 with the recombination frequency, suggesting that these SSR loci are located at close positions. Nevertheless, since No. of alleles, H_E and H_O were rather different for 83 representative mango accessions for these linked two SSR loci, it could be no problem to obtain exact genetic diversity and relatedness of mango cultivars.

Significant linkages between SSR loci were also evaluated for alleles of 'Keitt' (**Table 6**). Eleven SSR combinations showed significant linkages of 0.000 to 0.229 with the recombination frequency. No. of alleles, H_E and H_O for 11 SSR combinations for 83 representative mango accessions were rather different from each other. SSR loci MiIIHR14 and MiIIHR24 showed a complete linkage of 0.000 with the recombination frequency, however, they revealed different No. of alleles, H_E and H_O for 83 representative mango accessions.

Discussion

Mango shows the third biggest production of tropical fruits in the world, next to the bananas and the pineapples (FAOSTAT), and has been cultivated world-widely in the tropical and subtropical areas. In contrast to bananas and pineapples, however, mango has not been comprehensively studied as industrial plantations led by major commercial companies. Therefore, there have been conserved hundreds number of mango cultivars which may possess a certain genetic diversity with regionally uniqueness in the production areas. In Japan, mango commercial production started in 1980s. Because of the limited cultivation history and production areas in Japan, mango has not yet become major fruit crop in Japan (Ogata *et al.* 2016).

In this study, 120 mango accessions in Japan were clearly distinguished into 83 genotypes excluding synonymous and identical accessions by the SSR markers. There has been considerable confusion in the nomenclature of mango cultivars because of the use of synonyms for many cultivars, which increases the difficulty of identifying them (Krishna and Singh 2007). The use of SSR markers can differentiate mango cultivars and identify genetic diversity (Chiang *et al.* 2012, Duval *et al.* 2005, Honsho *et al.* 2005, Ravishankar *et al.* 2011, Schnell *et al.* 2005, Viruel *et al.* 2005). Some synonymous (identical SSR genotypes with different cultivar names) and homonymous (different SSR genotypes with the same cultivar name) accessions were pointed out in this study. Therefore, introduction background of mango accessions such as passport data should be carefully examined and considered again for validation as genetic resources, which will be utilized for breeding programs.

Using 11 SSR markers, Dillon *et al.* (2013) determined genetic diversity of 254 *M. indica* accessions maintained in the Australian National Mango Genebank, but found it difficult to identify parentage. Olano *et al.* (2005) analyzed 63 Florida cultivars to identify their pedigrees by using SSR markers, and Schnell *et al.* (2006) performed DNA analysis of 203 cultivars using SSR markers. The pedigree data that we obtained are in good accordance with those of Olano *et al.* (2005) and Schnell *et al.* (2006), including the many offspring of 'Haden' and the parentages of 'Irwin', 'Jubilee', and 'Lily'. The parentage of 'Dot-JIRCAS' ('Carrie' × 'Spirit of '76-JIRCAS') was newly identified in this study, confirmed by all loci except LMMA11. Similarly, the parentage of 'Jinhuang-JIRCAS' ('White-JIRCAS' × 'Kent') was confirmed by all loci except LMMA9. These discrepancies may be due to high mutation rates of SSR loci: estimates of mutation rates among loci vary over the range of 10^{-3} to 10^{-5} (Weber and Wong 1993) in human SSRs, exceeding mutation rates for non-SSR loci by up to four orders of magnitude (Lacy 1987). Moriya *et al.* (2011) likewise concluded that allele mutation occurred at one out of 46 SSR loci in 'Ozenokurenai' apple and its parents 'Morioka #47' × 'Morioka #46'.

PCoA indicated that accessions from India had a close

relationship with accessions from the USA, while accessions from Thailand, Taiwan, the Philippines, and Vietnam seemed to be genetically separate (**Fig. 2**). These groupings appear to correspond to the previously defined Indian and Southeast Asian types (Iyer and Degani 1997, Viruel *et al.* 2005). Structure analysis also identified three clusters: cluster I included accessions from India and some of Florida, cluster II contained most accessions from the Florida and Hawaii of USA, and cluster III included many accessions from Southeast Asia. Moreover, monoembryonic accessions predominated in clusters I and II, and polyembryonic accessions predominated in cluster III (**Supplemental Fig. 1**). These results were in good accordance with previous studies (Iyer and Degani 1997, Viruel *et al.* 2005).

Unstable flowering is one of the most important issues in mango cultivation and production to be solved, not only in Japan but also in Southeast Asia. It may be due in part to unstable climatic conditions such as obscurity seasonal change from rainy to dry period, and in part to higher temperatures during the flower initiation period as influenced by global warming (Normand *et al.* 2015). The mechanism of flower initiation tends to differ between the Indian and Southeast Asian types, reflecting the climate features of each region (Davenport 2009): flower initiation in the Indian type is induced mainly by low temperature, whereas that in the Southeast Asian type is induced mainly by drought stress in the dry season. It is important to understand cultivar characteristics and genetic diversity for choosing the appropriate genetic resources in order to maintain stable flowering in the practical field. Our results reveal the genetic structural distribution of the Indian and Southeast Asian types of mango genetic resources in Japan.

There has been no practical information about genetic diversity of mango in Japan. It is partly because the commercial production in Japan is quite recently (started from 1980s) and substantially monoculture of 'Irwin' (occupies >90% production in Japan), so there had been no strong interest about characteristics among genetic resources and also no intensive introduction of other new cultivar. However, recently, mango has been focused as one of the potential cash crops for premium fruit with high price in the commercial markets in Japan. The accessions that we examined cover almost all mango cultivars in Japan, therefore, their genetic information will pave the way to the use of the genetic resources for breeding and/or direct use of domestic production in Japan. Since the mango accessions used in this study have been mainly selected and established in Florida, and disseminated to the major production countries/areas (Mukherjee and Litz 2009), it is considered that mango accessions evaluated here could reflect the representative genetic diversity among major cultivars in the world.

Molecular markers have been used to create genetic linkage maps of mango (Arias *et al.* 2012, Kashkush *et al.* 2001, Kuhn *et al.* 2017, Luo *et al.* 2016). Although a lot of SSR markers have been developed (Chiang *et al.* 2012, Dillon *et al.* 2014, Duval *et al.* 2005, Honsho *et al.* 2005, Ravishankar *et al.* 2011, Schnell *et al.* 2005, Viruel *et al.* 2005), SSR-based genetic linkage maps were not constructed and reported. In this study, we evaluated 46 SSR markers with 96 F_1 individuals from 'Irwin' × 'Keitt', and identified that 35 SSR markers might be mapped in the genetic linkage maps of 'Irwin' and/or 'Keitt'. Four SSR combinations showing significant linkages for alleles of 'Irwin', i.e., MiIIHR05 vs. MiIIHR26, MiIIHR17 vs. MiIIHR32, MiSHRS-4 vs. LMMA2, and MiIIHR22 vs. LMMA10, could be positioned in the same linkage groups of 'Irwin'. Eleven SSR combinations showing significant linkages for alleles of 'Keitt' could be used for genome mapping of 'Keitt'. SSR markers provide a reliable method for evaluation of genetic diversity and construction of genetic maps because of their co-dominant inheritance and the allelic abundance (Weber and May 1989). Reference genetic linkage maps constructed with genome-wide molecular markers such as SSR markers are important for many genetic and breeding applications in fruit trees including marker-assisted selection (MAS), mapping of quantitative trait loci, and map-based gene cloning (Yamamoto and Terakami 2016). MAS can accelerate the selection process and reduce the number of progeny needed and thus the cost of raising individuals to maturity in the field (Luby and Shaw 2001).

Recently, high-density, almost saturated linkage maps in mango were developed through the use of next-generation sequencing-based and transcriptome-based single nucleotide polymorphism markers (Kuhn *et al.* 2017, Luo *et al.* 2016). Genetic maps are valuable tools for quantitative trait locus mapping and MAS of plants with desirable traits. Significant associations between traits and single nucleotide polymorphism markers for branch habit and for fruit bloom, ground skin color, blush intensity, beak shape, and pulp color (Kuhn *et al.* 2017) will be valuable for MAS in mango breeding programs.

With these advantages of recent molecular tools, mango genetic resources characterized in this study will be utilized to accelerate for promotion of mango cultivation in Japan and will contribute to provide information for breeding and/or adoption appropriate cultivar for stable production in the world.

Acknowledgments

This study was partially supported by an Okinawa special promotion grant. This work was also partially conducted as "Evaluation and utilization of diverse genetic materials in tropical field crops (EDITS)" in JIRCAS.

Literature Cited

Arias, R.S., J.W. Borrone, C.L. Tondo, D.N. Kuhn, B.M. Irish and R.J. Schnell (2012) Genomics of Tropical Fruit Tree Crops. *In*: Schnell, R.J. and P.M. Priyadarshan (eds.) Genomics of Tree Crops. Springer, New York, pp. 209–239.

Aron, Y., S. Gazit, H. Czosnek and C. Degani (1998) Polyembryony in

mango (*Mangifera indica* L.) is controlled by a single dominant gene. HortScience 33: 1241–1242.

Bompard, J.M. (2009) Taxonomy and Systematics. *In*: Litz, R.E. (ed.) The Mango, 2nd Edition: Botany, Production and Uses. CAB International, New York, pp. 19–41.

Campbell, R.J. (1992) A Guide to Mango in Florida. Fairchild Tropical Garden, Miami, Florida, USA.

Chiang, Y.C., C.M. Tsai, Y.K.H. Chen, S.R. Lee, C.H. Chen, Y.S. Lin and C.C. Tsai (2012) Development and characterization of 20 new polymorphic microsatellite markers from *Mangifera indica* (Anacardiaceae). Am. J. Bot. 99: e117–e119.

Davenport, T.L. (2009) Reproductive Physiology. *In*: Litz, R.E. (ed.) The Mango, 2nd Edition: Botany, Production and Uses. CAB International, New York, pp. 97–169.

Degani, C., R.E. Batsri and S. Gazit (1990) Enzyme polymorphism in mango. J. Am. Soc. Hortic. Sci. 115: 844–847.

Dillon, N.L., I.S.E. Bally, C.L. Wright, L. Hucks, D.J. Innes and R.G. Dietzgen (2013) Genetic diversity of the Australian National Mango Genebank. Sci. Hortic. 150: 213–226.

Dillon, N.L., D.J. Innes, I.S.E. Bally, C.L. Wright, L.C. Devitt and R.G. Dietzgen (2014) Expressed sequence tag-simple sequence repeat (EST-SSR) marker resources for diversity analysis of mango (*Mangifera indica* L.). Diversity 6: 72–87.

Duval, M.F., J. Bunel, C. Sitbon and A.M. Risterucci (2005) Development of microsatellite markers for mango (*Mangifera indica* L.). Mol. Ecol. Notes 5: 824–826.

Ellegren, H. (2004) Microsatellites: simple sequences with complex evolution. Nat. Rev. Genet. 5: 435–445.

Evanno, G., S. Regnaut and J. Goudet (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14: 2611–2620.

FAOSTAT, http://www.fao.org/faostat/en/#home

Fujii, H., H. Yamashita, T. Shimada, T. Endo, T. Shimizu and T. Yamamoto (2008) MarkerToolKit: an analysis program for data sets consist of DNA marker types obtained from various varieties. DNA Polymorph. 16: 103–107.

Fujii, H., H. Yamashita, F. Hosaka, S. Terakami and T. Yamamoto (2010) Development of a software to presume the parent-child relationship using the result of DNA marker typing. Hort. Res. 9 (Suppl. 1): 34.

Fujii, H., T. Ogata, T. Shimada, T. Endo, H. Iketani, T. Shimizu, T. Yamamoto and M. Omura (2013) Minimal marker: an algorithm and computer program for the identification of minimal sets of discriminating DNA markers for efficient variety identification. J. Bioinform. Comput. Biol. 11: 1250022.

Gan, Y.Y., S. Zaini and A. Idris (1981) Genetic variation in the grafted vegetatively propagated mango *Mangifera indica*. Pertanika 4: 53–63.

Hamilton, R.A. (1993) Origin and classification of mango varieties in Hawaii. Proceedings, Conference on Mango in Hawaii, pp. 28–33.

Hirano, R., O.T. Htun and K.N. Watanabe (2010) Myanmar mango landraces reveal genetic uniqueness over common cultivars from Florida, India, and Southeast Asia. Genome 53: 321–330.

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

Iyer, C.P.A. and C. Degani (1997) Classical breeding and genetics. *In*: Litz, R.E. (ed.) The Mango, Botany, Production and Uses. CAB International, New York, pp. 49–68.

Kashkush, K., F. Jinggui, E. Tomer, J. Hillel and U. Lavi (2001) Cultivar identification and genetic map of mango (*Mangifera indica*). Euphytica 122: 129–136.

Knight, Jr., R.J., R.J. Campbell and I. Maguire (2009) Important Mango Cultivars and their Descriptors. *In*: Litz, R.E. (ed.) The Mango, 2nd Edition: Botany, Production and Uses. CAB International, New York, pp. 42–66.

Krishna, H. and S.K. Singh (2007) Biotechnological advances in mango (*Mangifera indica* L.) and their future implication in crop improvement: a review. Biotechnol. Adv. 25: 223–243.

Kuhn, D.N., I.S.E. Bally, N.L. Dillon, D. Innes, A.M. Groh, J. Rahaman, R. Ophir, Y. Cohen and A. Sherman (2017) Genetic map of mango: a tool for mango breeding. Front. Plant Sci. 8: 577.

Lacy, R.C. (1987) Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection, and population subdivision. Conserv. Biol. 1: 143–158.

Lee, W.L., K.D. Chiou and I.S. Weng (2009) Study on genetic variations of mango (*Mangifera indica* L.) germplasm. J. Taiwan Agric. Res. 58: 243–253.

López-Valenzuela, J.A., O. Martínez and O. Paredes-López (1997) Geographic differentiation and embryo type identification in *Mangifera indica* L. cultivars using RAPD markers. HortScience 32: 1105–1108.

Luby, J.J. and D.V. Shaw (2001) Does marker-assisted selection make dollars and sense in a fruit breeding program? HortScience 36: 872–879.

Luo, C., B. Shu, Q. Yao, H. Wu, W. Xu and S. Wang (2016) Construction of a high-density genetic map based on large-scale marker development in mango using specific-locus amplified fragment sequencing (SLAF-seq). Front. Plant Sci. 7: 1310.

Marshall, T.C., J. Slate, L.E.B. Kruuk and J.M. Pemberton (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. 7: 639–655.

Mitra, S.K. (2016) Mango production in the world—present situation and future prospect. Acta Hortic. 1111: 287–296.

Moriya, S., H. Iwanami, K. Okada, T. Yamamoto and K. Abe (2011) A practical method for apple cultivar identification and parent-offspring analysis using simple sequence repeat markers. Euphytica 177: 135–150.

Mukherjee, S.K. (1953) The mango—Its botany, cultivation, uses and future improvement, especially as observed in India. Econ. Bot. 7: 130–162.

Mukherjee, S.K. and R.E. Litz (2009) Introduction: Botany and Importance. *In*: Litz, R.E. (ed.) The Mango, 2nd Edition: Botany, Production and Uses. CAB International, New York, pp. 1–18.

Nashima, K., S. Terakami, M. Kunihisa, C. Nishitani, M. Shoda, M. Matsumura, Y. Onoue-Makishi, N. Urasaki, K. Tarora, T. Ogata *et al.* (2017) Retrotransposon-based insertion polymorphism markers in mango. Tree Genet. Genomes 13: 110.

Normand, F., P.E. Lauri and J.M. Legave (2015) Climate change and its probable effects on mango production and cultivation. Acta Hortic. 1075: 21–31.

Ogata, T., S. Yamanaka, M. Shoda, N. Urasaki and T. Yamamoto (2016) Current status of tropical fruit breeding and genetics for three tropical fruit species cultivated in Japan: pineapple, mango, and papaya. Breed. Sci. 66: 69–81.

Olano, C.T., R.J. Schnell, W.E. Quintanilla and R.J. Campbell (2005) Pedigree analysis of Florida mango cultivars. Proc. Fla. State Hort. Soc. 118: 192–197.

Peakall, R. and P.E. Smouse (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28: 2537–2539.

Pritchard, J.K., M. Stephens and P. Donnelly (2000) Inference of

population structure using multilocus genotype data. Genetics 155: 945–959.

Ravishankar, K.V., B.H.R. Mani, L. Anand and M.R. Dinesh (2011) Development of new microsatellite markers from mango (*Mangifera indica*) and cross-species amplification. Am. J. Bot. 98: e96–e99.

Rohlf, F.J. (1998) NTSYS, numerical taxonomy and multivariate analysis system, Ver. 2.01. Exeter Publishing, Ltd., Setauket, New York.

Schnell, R.J., C.M. Ronning and R.J. Knight, Jr. (1995) Identification of cultivars and validation of genetic relationships in *Mangifera indica* L. using RAPD markers. Theor. Appl. Genet. 90: 269–274.

Schnell, R.J., C.T. Olano, W.E. Quintanilla and A.W. Meerow (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.

Schnell, R.J., J.S. Brown, C.T. Olano, A.W. Meerow, R.J. Campbell and D.N. Kuhn (2006) Mango genetic diversity analysis and pedigree inferences for Florida cultivars using microsatellite markers. J. Am. Soc. Hortic. Sci. 131: 214–224.

Shudo, A., K. Tarora, Y. Makishi, R. Ichi, K. Takahashi, M. Matsumura, S. Shimabukuro, N. Matsuda, S. Nakasone and N. Urasaki (2013) Development of CAPS markers and their application in breeding for mango, *Mangifera indica* L. Euphytica 190: 345–355.

Tomer, E., U. Lavi, C. Degani and S. Gazit (1993) 'Naomi': A new mango cultivar. HortScience 28: 755–756.

Tsai, C.C., Y.K.H. Chen, C.H. Chen, I.S. Weng, C.M. Tsai, S.R. Lee, Y.S. Lin and Y.C. Chiang (2013) Cultivar identification and genetic relationship of mango (*Mangifera indica*) in Taiwan using 37 SSR markers. Sci. Hortic. 164: 196–201.

Van Ooijen, J.W. (2011) Multipoint maximum likelihood mapping in a full-sib family of an outbreeding species. Genet. Res. (Camb) 93: 343–349.

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

Weber, J.L. and P.E. May (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am. J. Hum. Genet. 44: 388–396.

Weber, J.L. and C. Wong (1993) Mutation of human short tandem repeats. Hum. Mol. Genet. 2: 1123–1128.

Yamamoto, T. and S. Terakami (2016) Genomics of pear and other Rosaceae fruit trees. Breed. Sci. 66: 148–159.