Theor Appl Genet (1995) 90: 269–274

R. J. Schnell · C. M. Ronning · R. J. Knight, Jr.

Identification of cultivars and validation of genetic relationships in *Mangifera indica* L. using RAPD markers

Received: 6 December 1993 / Accepted: 21 June 1994

Abstract Twenty-five accessions of mango were examined for random amplified polymorphic DNA (RAPD) genetic markers with 80 10-mer random primers. Of the 80 primers screened, 33 did not amplify, 19 were monomorphic, and 28 gave reproducible, polymorphic DNA amplification patterns. Eleven primers were selected from the 28 for the study. The number of bands generated was primer- and genotype-dependent, and ranged from 1 to 10. No primer gave unique banding patterns for each of the 25 accessions; however, ten different combinations of 2 primer banding patterns produced unique fingerprints for each accession. A maternal half-sib (MHS) family was included among the 25 accessions to see if genetic relationships could be detected. RAPD data were used to generate simple matching coefficients, which were analyzed phenetically and by means of principal coordinate analysis (PCA). The MHS clustered together in both the phenetic and the PCA while the randomly selected accessions were scattered with no apparent pattern. The uses of RAPD analysis for *Mangifera* germ plasm classification and clonal identification are discussed.

Key words Random amplified polymorphic DNA · Mango · Genetic diversity · Cultivar identification

Introduction

Genetic fingerprinting has been accomplished traditionally through the use of isozymes, and more recently through restriction fragment length polymorphisms (RFLPs), variable number tandem repeats (VNTRs), or a combination of these. While these methods have been very useful in cultivar identification, they have a number of disadvantages, including a limited number of isozyme loci, and the time, expense, and use of [^{32}P] for labeling with RFLPs and VNTRs. Polymerase Chain Reaction (PCR) based methods overcome these disadvantages. Random amplified polymorphic DNA (RAPD) uses arbitrary 10-base primers to amplify random portions of the genome (Welsh and McClelland 1990; Williams et al. 1990). The fragments produced are easily visualized on an ethidium bromide-stained gel, and polymorphisms between genotypes reflect heritable differences in the genome. Large numbers of bands, or loci, can be generated with relative ease. Due to the arbitrary nature of the primers, RAPD markers, unlike RFLPs, VNTRs, and isozymes, represent a random sample from the entire genome. They are, however, inherited as dominant markers, requiring larger numbers of loci to be identified and screened to glean the same information as from RFLPs, VNTRs, and isozymes. Amplification of bands in the progeny which were not amplified in either parent have also been reported (Riedy et al. 1992).

The use of RAPDs to determine genetic relationships has been demonstrated in maize (Welsh et al. 1991), conifers (Carlson et al. 1991), cacao (Wilde et al. 1992), and various leguminous species (Chalmers et al. 1992, Echt et al. 1992). Within *Mangifera* (mango) species RAPDs have been used to determine phylogenetic relationships (Schnell and Knight 1993). On the basis of the classification of Kostermans and Bompard (1993), the genus is divided into the sub-genera Mangifera and Limus. Cluster analysis of nine species, five in the Mangifera and four in Limus, gave results inconsistent with the classification of Kostermans and Bompard, which was based on traditional taxonomic characters including flower morphology, leaf type, and fruit characteristics. The sub-genera did not form two distinct groups in the cluster analysis. Species from each group were intermixed; however, closely related species within a subgroup did cluster together.

Within the species *M. indica* L. (the common mango), genetic relationships are difficult to determine because usually only the maternal parent is known. Controlled pollinations are difficult to make because of the small flower

Communicated by A. L. Kahler

R. J. Schnell (✉) · C. M. Ronning · R. J. Knight Jr.
U.S. Department of Agriculture,
National Clonal Germplasm Repository,
Subtropical Horticulture Research Station, 13601 Old Cutler Rd.,
Miami, FL 33158, USA

size. Trees with synchronous flowering must be placed together in a fine 32 mesh cage with insects, usually flies, for successful pollination. Two types of flowers occur on the panicles of mango, male and hermaphroditic, and the ratio of one type to the other is cultivar dependent. Because of the difficulty of making controlled crosses horticulturists have traditionally collected open-pollinated seed and evaluated maternal half-sib families (MHS) for selection.

Isozymes have been used to determine parentage for some caged crosses and self-pollinated trees, but only six loci with 17 allelomorphs have been found (Degani et al. 1991). These markers are useful but are limited in their ability to discern genetic relationships because there are so few of them.

The long juvenility period of mangos (up to 5 years) would make RAPD markers an extremely useful tool for the identification of cultivars during propagation and growth. The ability to identify cultivars using RAPD markers would also aid in the management of germ plasm collections of mango, as identical cultivars often have different names. We report here on the use of RAPD markers to distinguish cultivars of mango. To examine the usefulness of RAPDs for determining genetic relatedness, a sample of 16 random cultivars and a group with known maternal pedigree were chosen for analysis. The maternal group consisted of the parents, F_1, and a MHS family of six seedlings derived from the F_1. Our objective was to determine if known genetic relationships between the parents, F_1, and MHS would become apparent and differ from the 16 random cultivars using these molecular markers.

Materials and methods

Plant material

Twenty-five clones of mango were selected from the germ plasm collection at the National Clonal Germplasm Repository, Miami, Fla. (Table 1). The clones included a group with a known maternal pedigree consisting of 'Mulgoba' (♀), 'Turpentine' (♂), seedlings, ('Eldon', 'Lippens', 'Osteen', 'Glenn', 'Tommy' Atkins, and 'Zill' (Campbell 1992)) as well as 16 randomly selected clones.

DNA extraction

DNA was extracted from young, flushing leaves using the hexadecyltrimethylammonium bromide (CTAB) method described by Doyle and Doyle (1990), with modifications. Fresh tissue 3 g was ground to a fine powder in liquid nitrogen, followed by the addition of 22.5 ml preheated (60°C) extraction buffer with further grinding. The extraction buffer consisted of 4% (w/v) CTAB, 1.4 *M* NaCl, 0.1% (v/v) β-mercaptoethanol, 20 m*M* EDTA, 100 m*M* TRIS-HCl, pH 8.0, and 1% (w/v) polyvinylpolypyrrolidone (PVPP). The homogenate was incubated at 60°C for 30 min and extracted one time with chloroform:isoamyl alcohol (24:1), after which the nucleic acids were precipitated in cold isopropanol. After treatment with RNase A, DNA concentration was determined spectrophotometrically. In addition, the DNA from 8 clones was purified through a CsCl gradient (Sambrook et al. 1989) for comparison with the CTAB method.

Primer screening

Eighty primers were initially screened using 5 genotypes to determine the suitability of each primer for the study. The genotypes selected were the south Florida cultivars 'Keitt', 'Tommy Atkins',

Table 1 Mango germ plasm analyzed for RAPDs

Clone	M-number	Source	Origin
Mulgoba	MIA9554	Poona India	India
Langra	MIA17546	Poona India	Pakistan
Vanraj	MIA17697	Poona India	India
Fajri	MIA17549	Pakistan	Pakistan
Madras	MIA6596	India	India
Himsagar	MIA17568	Dacca Pakistan	India
Manila	MIA25601	Chiapas Mexico	S.E. Asia
Haden	MIA9555	Miami Fla., USA	South Florida USA
Eldon	MIA19833	Homestead Fla., USA	South Florida USA
Lippens	MIA19680	Homestead Fla., USA	South Florida USA
Osteen	MIA26403	Merritt Is. Fla., USA	South Florida USA
Gleen	MIA34818	Miami Fla., USA	South Florida USA
Tommy Atkins	MIA24791	Broward Co. Fa., USA	South Florida USA
Zill	MIA17453	Lake Worth Fla., USA	South Florida USA
Keitt	MIA17415	Homestead Fla., USA	South Florida USA
Irwin	MIA17452	Miami Fla., USA	South Florida USA
Ruby	MIA19836	Miami Fla., USA	South Florida USA
Van Dyke	MIA23259	Miami Fla., USA	South Florida USA
Kent	MIA17450	Coconut Grove Fla., USA	South Florida USA
Pope	MIA34827	Homestead Fla., USA	South Florida USA
Edgehill	MIA19047	Bonsall Calif. USA	Unknown
Turpentine	MIA25158	Miami Fla., USA	West Indies
Mamita	MIA24217	St. Croix VI USA	Puerto Rico
Tyler	MIA17858	Nyasaland Zimbabwe	Zimbabwe
Zelanica	MIA6921	Peradeniya Sri Lanka	Sri Lanka

'Irwin', 'Van Dyke', and 'Turpentine'. Of the 80, 33 did not amplify, 19 amplified but were monomorphic, and 28 amplified and were polymorphic. From the 28 primers that were polymorphic, 11 were chosen for screening of the experimental population.

DNA amplification

Using RAPD primer kits A, B, C, D, and F (Operon Technologies, Alameda, Calif.) DNA was amplified following a modified protocol of Williams et al. (1990). The reactions contained 10 m*M* TRIS-HCl, pH 8.3; 50 m*M* KCl; 3.0 m*M* $MgCl_2$; 100 µ*M* each dATP, dCTP, dTTP, and dGTP; 0.3 µ*M* primer; 0.5 U Ampli *Taq* DNA polymerase (Perkin Elmer Cetus, Norwalk, Conn.); and 25 ng genomic DNA, in a total volume of 25 µl. Reactions were cycled 45 times at 92°C for 1 min, 35°C for 1 min and 72°C for 2 min in a DNA thermocycler 480 (Perkin Elmer). Upon completion of the amplification, a 10 µl aliquot was loaded onto a 1.4% agarose/0.5×TRIS Borate EDTA gel and electrophoresed at 4.4 V/cm. Gels were stained using ethidium bromide. Each amplification product was identified with reference to the primer used in the reaction followed by its size in base pairs (bp). Amplification was repeated twice.

Data analyses

Data were scored as discrete variables, using 1 to indicate presence and 0 to indicate absence of a band. A pairwise difference matrix between genotypes was determined using the simple matching coefficient (Sokal and Michener 1958) and the average taxonomic distance. Unweighted pair-group method using arithmetic average (UPGMA) cluster analysis was performed using this matrix. Principal coordinate analysis (PCA) of the difference matrix was used to further estimate relationships between individuals. All computations were performed using the microcomputer program NTSYS (Numerical Taxonomy System ver. 1.8).

Results and discussion

Optimization

To determine the optimal amplification conditions, a factorial experiment varying $MgCl_2$, DNA, and primer concentrations was conducted. The design included three levels of $MgCl_2$ (2.5, 3.0, and 4.0 m*M*), three levels of primer (0.1,0.2, and 0.3 µ*M*) and two levels of DNA (0.5 and 1.0 ng/µl). Operon primer OPF-13 was used with 2 different clones in the optimization experiment.

$MgCl_2$ and primer concentrations each had an effect on RAPD banding patterns, while no differences were found to be due to DNA concentrations. More bands were produced using 4.0 m*M* $MgCl_2$, but the patterns were not reproducible. Both 2.5 and 3.0 m*M* $MgCl_2$ gave identical, consistent patterns. Primer concentration had the greatest effect on band number and reproducibility. The number of bands increased with increasing primer concentration over all $MgCl_2$ levels and DNA concentrations. No interactions between the three factors were observed. On the basis of these results, 3.0 m*M* $MgCl_2$, 0.3 µ*M* primer, and 1.0 ng/µl gDNA were used in all subsequent amplification reactions. Three primers, OPA15, OPA18, OPF13, were used to test for differences between crude CTAB preparations and CsCl purified samples. Under the optimized conditions described above, no differences were found between crude and pure samples.

Table 2 Primers and basepair lengths of RAPD generated markers

Primer	basepair length (bp)									
OPA8	1900[a]									
OPA15	1600	1500	1300	1000						
OPA18	2100	1800	1600	1500	1300	1150	1100	1000	950	800
OPA19	2000	1800	1600	700	600	500				
OPA20	2000	1700	1500	1150	1000	850	650[a]			
OPF6	1150[a]	1050	900	800	600					
OPF12	2000	1800	1300	1100	550					
OPF13	2100	2000	1700	1600	1500	1150	950	850[a]	650[a]	500
OPF15	1350	1100[a]	1000[a]	950	800	550				
OPF16	1050[a]	950[a]	800							
OPF20	2100	1900[a]	1700	1500	1200	1050	950	800	700	400

[a] Monomorphic

Problems with reproducibility of RAPD banding patterns have been reported by Penner et al. (1993) and observed in our laboratory. Using our optimized conditions, we performed two sets of reactions separately for each primer. During primer screening replications were loaded on separate gels and compared. Since no differences were detected between reps run on separate gels, replicate amplifications thereafter were run on the same gel. Only strong, reproducible bands were scored, while variable and faint bands were excluded.

A total of 151 bands were generated by the 11 primers. Only 67 (44%) were reproducible "loci" used for the data analysis. The primers varied in the number of bands produced, ranging from 1 to 10, with an average of 6.1 bands per primer. The size range of amplification products varied with the primer but was generally from 2000 bp to 450 bp (Table 2). Examples of the molecular profile generated by primer OPA18 are shown in Fig 1. Of the 67 loci generated, 57 (85%) were polymorphic (Table2).

Cultivar identification

While no single primer produced unique banding patterns for all 25 individuals, combinations of primers gave a range of 0–100% unique individuals (Table 3). When data from each primer was analyzed in pairs the primer combinations OPA15/OPA18, OPA18/OPA19, OPA18/OPA20, OPA18/OPAF6, OPA18/OPF13, OPA18/OPF15, OPA18/OPF20, OPA15/OPA19, OPF12/OPF13, and OPF13/OPF20 produced unique patterns for all 25 individuals. OPA18 is the most useful primer for identifying individuals differentiating all but 2 of the 25 genotypes. Thus, a standard set of primers, (OPA15, OPA18, OPA19, OPF12, OPF13, and OPF20) can be used to distinguish and characterize these mango cultivars, which represent most of the common mango cultivars grown in the USA.

Genetic relationships

The pedigree for the 'Haden' family of mango is known from historical records. 'Mulgoba' was imported as an in-

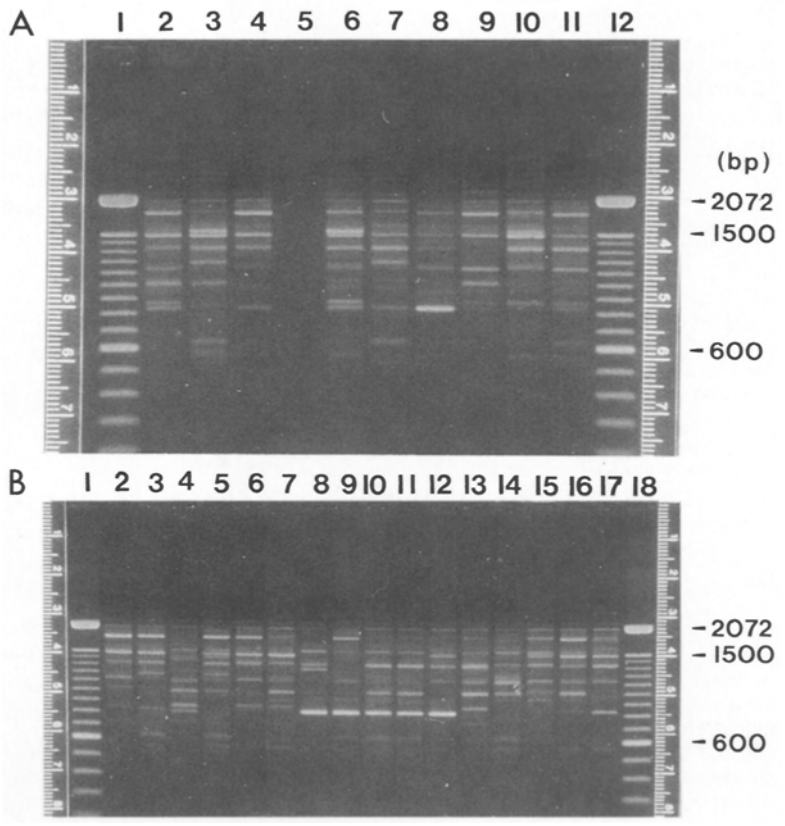

Fig. 1A, B RAPD-detected polymorphisms in 25 mango accessions using primer OPA18. **A** Parents, F_1, and maternal half-sibs (see Table 1). *Lanes 1–12* 100-bp ladder, 'Mulgoba', 'Turpentine', 'Haden' (blank), 'Eldon', 'Lippens', 'Osteen', 'Glenn', 'Tommy Atkins', 'Zill', 100 bp ladder. **B** Randomly selected clones. *Lanes 1–18* 100 bp ladder, 'Ruby', 'Pope', 'Zelanica', 'Vanraj', 'Edgehill', 'Keitt', 'Madras', 'Langra', 'Himsagar', 'Fajri', 'Van Dyke', 'Manila', 'Mamita', 'Tyler', 'Kent', 'Irwin', 100-bp ladder

Table 3 Percentage of unique patterns among 25 mango accessions for each pairwise combination of primers

Primer	OPA15	OPA18	OPA19	OPA20	OPF6	OPF12	OPF13	OPF15	OPF16	OPF20
OPA15	0.04									
OPA18	1.00	0.92								
OPA19	0.76	1.00	0.48							
OPA20	0.60	1.00	0.92	0.28						
OPF6	0.52	1.00	0.48	0.52	0.12					
OPF12	0.64	0.92	1.00	0.92	0.76	0.08				
OPF13	0.52	1.00	0.92	0.92	0.72	0.68	0.52			
OPF15	0.36	1.00	0.80	0.56	0.48	1.00	0.80	0.16		
OPF16	0.04	0.92	0.48	0.40	0.28	0.40	0.52	0.02	0.00	
OPF20	0.68	1.00	0.84	0.84	0.92	0.92	1.00	0.76	0.64	0.48

arched plant from India by the USDA in 1887. In 1900, Captain Haden of Coconut Grove, Fla., bought fruit of 'Mulgoba' from Professor Gale, of West Palm Beach, Fla., and planted seedlings of 'Mulgoba' in his yard. The common turpentine mango, which had been introduced from the West Indies during the late 1800s, was widely grown in Florida at the time and probably grew in Professor Gale's yard. One of Captain Haden's 'Mulgoba' seedlings produced a fine quality fruit with red color and annual bearing. This seedling was named 'Haden' and was propagated by grafting. It became the standard commercial mango cultivar in Florida from 1910 until about 1945 (Knight and Schnell 1993). 'Haden' has been assumed to be a seedling of 'Mulgoba'×'Turpentine'; however, Schnell and Knight (1992) demonstrated that phenotypically indistinguishable trees of 'Turpentine' can be genotypically different. 'Turpentine' exists as a landrace of primarily seed-propagated clones, with most seedlings being of nucellar origin but with occasional zygotic seedlings. Thus, several genotypes apparently can occur within populations recognized as 'Turpentine'.

The phenogram (Fig. 2) and PCA (Fig. 3) generally support the known and assumed pedigree with the exception of 'Glenn' and 'Osteen'. With both analyses, 'Glenn' and 'Osteen' failed to cluster with their related MHS, cvs 'Eldon', 'Lippens', 'Tommy Atkins', and 'Zill'. 'Glenn' was selected from a dooryard planting in 1945 in Miami, Fla., and 'Osteen' was selected from a dooryard planting in 1935 at Merritt Island, Fla. While the presumed seed parent was 'Haden', this information is anecdotal, and our

data suggest that the parents may have been other varieties. 'Haden' clusters midway between 'Turpentine' and 'Mulgoba' in the PCA; the MHS cluster around 'Haden' (Fig. 3). All of the cultivars that make up the MHS have fruit more similar to 'Haden' or 'Mulgoba' than to 'Turpentine'. This is expected, since these cultivars were selected using standards which most closely approach the standards for a commercial cultivar, as seen in Haden. All cultivars flower and set fruit annually under south Florida conditions, while many cultivars selected elsewhere and imported into south Florida do not. This trait was most likely inherited from the paternal grandparent 'Turpentine', which fruits regularly and abundantly in Florida. 'Mulgoba', the maternal grandparent, flowers irregularly and bears few fruit under south Florida conditions.

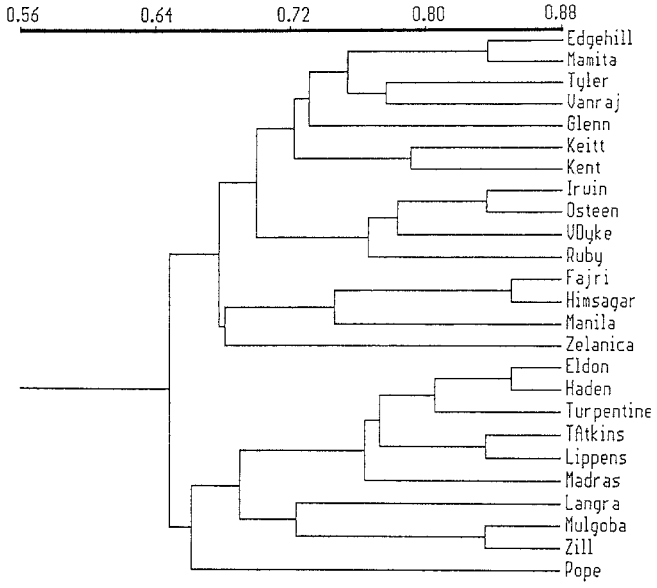

Fig. 2 Phenogram derived from analysis of 25 mango accessions using 11 random primers and a pairwise difference matrix between genotypes with UPGMA cluster analysis

Associations among the random group of south Florida cultivars can been seen between 'Irwin' (a seedling of 'Lippens'), 'Van Dyke' (unknown parentage), 'Keitt' (a seedling of 'Mulgoba'), 'Kent' (a seedling of 'Brooks'), and 'Glenn' (suspect seedling of Haden). These cultivars have many similar phenotypic traits such as fruit shape and fruit color. They also have many differences, such as fruit size: 400–600 g for 'Irwin', 'Glenn', and 'Van Dyke', to 750–2000 g for 'Kent' and 'Keitt'. 'Ruby' is also a south Florida cultivar of unknown origin, with an yellow color and a oblong shape that is distinct from many other south Florida cultivars. 'Brooks', the seed parent of 'Kent', is a seedling of 'Sandersha', and this fact may account for 'Kent's' relative distance from 'Haden' and most 'Haden' seedlings. 'Sandersha' may also be involved in the lineage of other Florida cultivars that are distant from 'Haden' and most of its seedlings, but without further investigation this possibility remains conjectural. Over the past 100 years Florida has become a secondary center of diversity for mango germ plasm (Knight and Schnell 1993), and it is not unreasonable to expect considerable diversity among the Florida cultivars. The Indian cultivars 'Mulgoba' (maternal parent of 'Haden'), 'Langra', 'Vanraj', and 'Madras', differ from each other in many traits, and from other Indian cultivars examined. 'Himsagar' and 'Fajri' are phe-

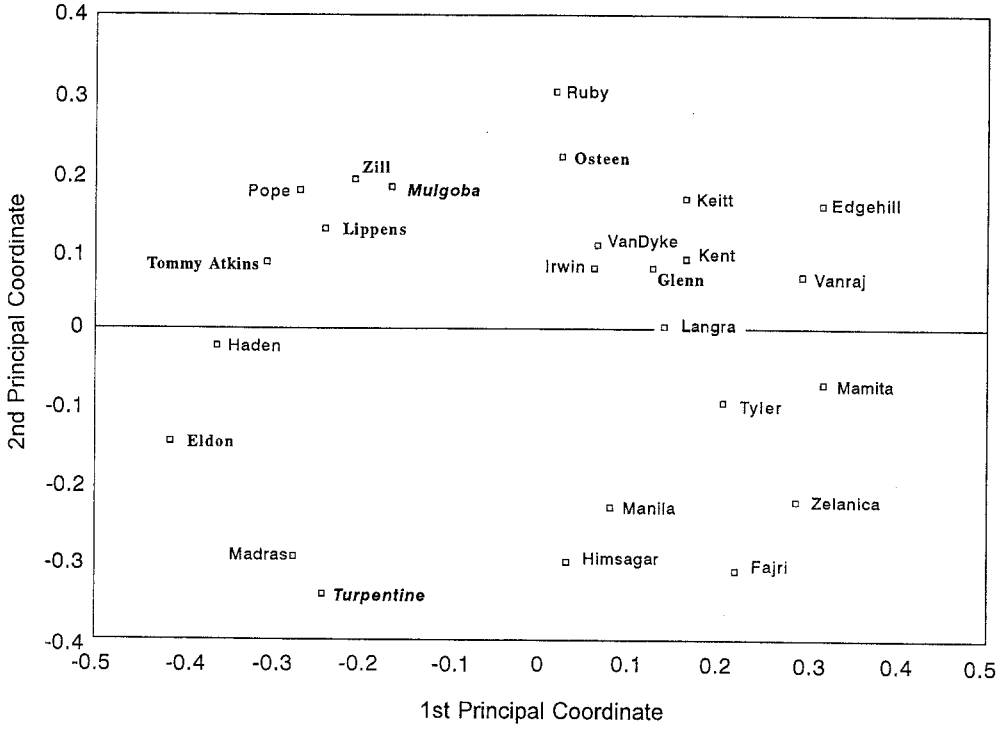

Fig. 3 Principal coordinate analysis of the pairwise difference matrix for 25 mango accessions and RAPD generated markers

notypically similar, and these two cluster together. The Indian mangos, coming from a primary center of diversity for this crop, make up a varied group. Other cultivars among the random group, including 'Manila', 'Edgehill', 'Mamita', 'Pope', 'Tyler', and 'Zelanica', are of disparate origins and unknown genetic relationships.

RAPDs will be useful for cultivar identification and for estimating genetic relationships in mango. The primers OPA15, OPA18, OPA19, OPF12, OPF13, and OPF20 provide a core set that will be useful for fingerprinting south Florida mango cultivars. In the day-to-day management of the germ plasm collection, RAPDs allow identification of redundancy and provide an additional cultivar verification method. Further genetic diversity analysis within the mango germ plasm collection will provide useful information for managing this collection and for determination of a core collection for mango.

References

Campbell RJ (1992) (Ed.) A Guide to Mangos in Florida. Fairchild Everbest Printing Co. Ltd. Hong Kong

Carlson JE, Tulsieram LK, Glaubitz JC, Luk VWK, Kauffeldt C, Rutledge R (1991) Segregation of random amplified DNA markers in F1 progeny of conifers. Theor Appl Genet 83:194–200

Chalmers KJ, Waugh R, Sprent JI, Simons AJ, Powell W (1992) Detection of genetic variation between and within populations of *Gliricidia sepium* and *G. maculata* using RAPD markers. Heredity 69:465–472

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

Doyle JJ, JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Echt CS, Erdahl LA, McCoy TJ (1992) Genetic segregation of random amplified polymorphic DNA in diploid cultivated alfalfa. Genome 35:84–87

Knight RJ, Schnell RJ (1993) Mango (*Mangifera indica* L.) introduction and evaluation in Florida and its impact on the world industry. Acta Hortic 341: 125–132

Kostermans AJGH, Bompard JM (1993) The Mangoes. Academic Press.London

Penner GA, Bush A, Wise R, Kim W, Domier L, Kasha K, Laroche A, Scoles G, Molnar SJ, Fedak G (1993) Reproducibility of random amplified polymorphic DNA (RAPD) analysis among laboratories. PCR Methods Applic 2:341–345

Riedy MF, Hamilton WJ III, Aquadro CF (1992) Excess of non-parental bands in offspring from known primate pedigrees assayed using RAPD PCR. Nucleic Acids Res 20(4):918

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Schnell RJ, Knight RJ (1992) Frequency of zygotic seedlings from five polyembryonic mango rootstocks. HortScience 27:174-176

Schnell RJ, Knight RJ (1993) Genetic relationships among *Mangifera* spp. based on RAPD markers. Acta Hortic 341:86–92

Sokal RR, Michener CD (1958) A statistical method for evaluating systematic relationships. Univ Kan Sci Bull 38:1409–1438

Welsh J, McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res 18:7213–7218

Welsh J, Honeycutt RJ, McClelland M, Sobral BWS (1991) Parentage determination in maize hybrids using the arbitrarily primed polymerase chain reaction (AP-PCR). Theor Appl Genet 82: 473–476

Wilde J, Waugh R, Powell W (1992) Genetic fingerprinting of *Theobroma* clones using randomly amplified polymorphic DNA markers. Theor Appl Genet 83:871–877

Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 18:6531–6535