

A Numerical Taxonomic Study of the Mango *Mangifera indica* L.¹

A. M. Rhodes,² Carl Campbell,³ Simon E. Malo,³ and S. G. Carmer^{2,4}

Abstract. Variability among 40 cultivars of *Mangifera indica* L. and one specimen each from *M. odorata* Griff. and *M. zeylanica* Hook f. were measured and classified using the unweighted pair group method of cluster analysis of distance coefficients based on 73 characters. Neither *M. odorata* nor *M. zeylanica* showed a close overall similarity to any of the cultivars of *M. indica*. Most cultivars clustered into 1 of 4 major groups. One group contained the polyembryonic cultivars with oblong fruit common to Southeast Asia. Another group consisted of monoembryonic cultivars with roundish fruit common to India. A third group, intermediate in fruit shape, included cultivars from India and one from the West Indies. The final group involved several hybrids developed in Florida and Hawaii. This group, as a rule, has large fruit and is designated as the Sandersha-Haden complex. A tentative pedigree based on both reported parentage and distance coefficients is given for this group. A few cultivars from Indo-china, the West Indies, and Réunion did not show close enough affinity to be placed into any of the above groups.

MATERIALS AND METHODS

The trees used in this study are located at the Sub-Tropical Experiment Station, Institute of Food and Agricultural Sciences, University of Florida, Homestead. The trees are grafted on 'Turpentine' mango rootstocks, which are uniform because of their nucellar origin.

MANGO genotypes are commonly divided into 2 general classes, polyembryonic and monoembryonic. The polyembryonic class is generally referred to as the Indo-chinese type and the monoembryonic class as the Indian type. Popenoe (5), after a study of mango genotypes being grown in Florida, further divided the monoembryonic class into 3 groups: 1) Mulgoba, 2) Alphonse, and 3) Sandersha. Popenoe's study was of particular importance because his classification was based on natural resemblance among genotypes. This resemblance was determined from characters taken from various parts of the plant, although main emphasis was given to the fruit and inflorescence. Similar classifications, also based on both reproductive and vegetative characters, have been reported by Mukherjee (3), Naik and Gangolly (4), Singh and Singh (8), and Valmayor (11). A summary of mango classification studies, as well as descriptions of cultivars, can be found in Singh (7). These studies, however, do not quantify the relationship of a given genotype or group of genotypes to the rest of the germ plasm of the species. If the relationships among genotypes were better understood, a more natural classification could be devised.

One aid in studying genetic relationships is the use of the methods of numerical taxonomy techniques. These can be employed to compare the overall phenetic similarities among a set of entities and to order these entities into groups. The overall similarity is generally determined by comparing a large number of characters common to the entities under study. In numerical taxonomic literature, an entity is generally referred to as an operational taxonomic unit (OTU). The same terminology will be followed in this paper. Further details of the theory and practices of numerical taxonomy are described by Sokal and Sneath (10).

The purpose of this paper is to report a study using the methods of numerical taxonomy to measure and classify the variability among 40 OTU's of mangos. Some of the OTU's have been developed in Florida, and the others have been introduced from various parts of the world. As a group, these 40 OTU's represent much of the genetic variability among mangos being grown in Florida. In addition to the 40 OTU's of *M. indica*, 2 other species, *M. odorata* and *M. zeylanica*, were included in this study.

Table 1. List of code numbers, names, embryonic types, and sources of OTU's used in this study.

Code no.	Name	Embryo type	Reported parentage	Geographical origin
01	Adams	Mono	Unknown	Florida
02	Ameeri	Mono	Unknown	India
03	Amini	Mono	Unknown	India
04	Aroemanis	Poly	Unknown	Java
05	Bennett	Mono	Unknown	India
06	Bombay Green	Mono	Unknown	India
07	Borsha	Mono	Unknown	India
08	Brooks	Mono	Seedling of Sandersha	Florida
09	Carabao	Poly	Unknown	Philippines
10	Golek	Poly	Unknown	Java
11	Edward	Mono	Haden × Carabao	Florida
12	Fascell	Mono	Seedling of Brooks	Florida
13	Florigon	Poly	Unknown	Florida
14	Gola	Mono	Unknown	India
15	Haden	Mono	Seedling of Mulgoba	Florida
16	Irwin	Mono	Seedling of Lippens	Florida
17	Itamaraca	Mono	Unknown	Brazil
18	Julie	Mono	Unknown	Réunion
19	Keitt	Mono	Seedling of Mulgoba	Florida
20	Kent	Mono	Seedling of Brooks	Florida
21	Madame Francis	Poly	Unknown	Haiti
22	Madoe	Poly	Unknown	Java
23	Micongensis ^a	Poly	Unknown	Indochina
24	Mulgoba	Mono	Unknown	India
25	Mun	Poly	Unknown	Thailand
26	Nam Dank Mai ^b	Poly	Unknown	Thailand
27	Number Eleven	Poly	Unknown	West Indies
28	Paheri	Mono	Unknown	India
29	Palmer	Mono	Unknown	Florida
30	Pettigrew	Mono	Unknown	Florida
31	Pope	Mono	Seedling of Irwin	Hawaii
32	Rajpuri	Mono	Unknown	India
33	Rockdale Saigon	Poly	Unknown	Indochina
34	Safeda Lucknow	Mono	Unknown	India
35	Sandersha	Mono	Unknown	India
36	Simmonds	Poly	Haden × Carabao	Florida
37	Smith	Mono	Seedling of Haden	Florida
38	Tommy Atkins	Mono	Unknown	Florida
39	Turpentine	Poly	Unknown	Cuba
40	White Langra	Mono	Unknown	India
41	M. Odorata	Mono		Southeast Asia
42	M. Zeylanica	Mono		Ceylon

¹Received for publication October 6, 1969.

²University of Illinois, Urbana.

³University of Florida, Homestead.

⁴The authors express their appreciation to R. A. Conover, R. A. Hamilton and W. B. Storey for their assistance in reviewing the manuscript.

^aSyn. Mekongensis.

^bSyn. Nam Doc Mai.

Data for fruit and seed characters were recorded during the 1968 season, and the other characters were recorded during the period from December 1968 to April 1969. Table 1 lists the code number, name, embryo type, reported parentage, and reputed geographical origin for each OTU. Table 2 lists the characters, character states for each character, and integer code for each character state.

The following computations were carried out on the coded data matrix: 1) Q-correlation, distance (9), and divergence (2) coefficients were computed for all pairs of OTU's. 2) The resulting similarity matrices were clustered using the unweighted pair group method with arithmetic averages (UPGMA) described in Sokal and Sneath (10), and the results were summarized in the form of phenograms. 3) Before computing Q-correlation and distance coefficients, characters were standardized to have zero means and unit variances. 4) Before computing divergence coefficients, the coded character scores, which were in arithmetic sequence (e.g. 1,2,3) for all characters, were changed to a geometric sequence (e.g. 1,2,4) to stabilize the effects of the denominator in the divergence equation (6). The scores were not standardized when computing the divergence coefficients.

Computations for the 3 types of similarity coefficients and cluster analysis were performed using University of Illinois Agronomy Statistical Laboratory programs on an IBM 360 Model 75 computer.

Three additional diagrams were constructed to support the OTU relationships as found in the cluster analysis.

RESULTS AND DISCUSSIONS

Phenograms and hybrids. The phenograms, based on cluster analyses of the 3 similarity matrices of 42 OTU's, show essentially the same arrangement in the first or primary clusters, but at lower levels the clusters are not so congruent. Part of this discrepancy may be due to the hybrid nature of the Florida OTU's, some of which have been developed from crosses between diverse types of mangos introduced from India, Southeast Asia, and the American Tropics.

In cluster analysis, a hybrid can cluster with only one parent when the parents are in different clusters, and because the hybrid is usually intermediate with respect to its parents, it will exert an oblique effect on subsequent clustering arrangement of other OTU's. For this reason, an additional analysis was made on the 27 introduced OTU's with hybrids from Florida and Hawaii, which are descendants of the introduced OTU's, omitted.

In comparing the phenograms of the 27 introduced

Table 2. List of characters, character states, and arithmetic code for phenetic classification of mangos.

No.	Character	Character states with arithmetic code
01	Tree size.	(1) dwarf, (2) small, (3) medium, (4) large.
02	Branching habit.	(1) erect, (2) intermediate, (3) spreading.
03	Bark texture.	(1) smooth, (2) intermediate, (3) rough.
04	Relative shoot thickness.	(1) thin, (2) medium to thick.
05	Shoot color.	(1) green, (2) yellow, (3) pink, (4) red.
06	Bearing habit.	(1) consistent, (2) variable.
07	Relative maturity.	(1) early, (2) mid season, (3) late.
08	Average annual yield.	(1) light, (2) medium, (3) heavy.
09	Mature leaf color (dorsal surface).	(1) light, (2) intermediate, (3) dark green.
10	Leaf orientation.	(1) upheld, (2) outheld, (3) downheld.
11	Leaf habit.	(1) flat, (2) slightly folded, (3) folded.
12	Leaf halves usually symmetrical.	(1) no, (2) yes.
13	Leaf margin undulated.	(1) no, (2) slightly, (3) moderately.
14	Leaf twisted.	(1) no, (2) slight, (3) yes.
15	Leaf shape.	(1) lanceolate, (2) elliptic, (3) oblanceolate.
16	Leaf apex.	(1) acuminate, (2) subacuminate, (3) acute, (4) obtuse.
17	Tip of apex.	(1) pointed, (2) rounded.
18	Average leaf length (mm).	(1) under 250, (2) 250 to 350, (3) over 350.
19	Ratio of leaf length to width.	(1) under 3.8, (2) 3.8 to 4.5, (3) over 4.5.
20	Number of pairs of secondary veins, basal half of leaf.	(1) under 11, (2) 11 to 14, (3) over 14.
21	Length of panicle (mm).	(1) under 300, (2) 300 to 350, (3) over 350.
22	Ratio of panicle width to length.	(1) .50, (2) .50 to .75, (3) over .75.
23	Leafy bracts persist.	(1) no, (2) yes.
24	Density of hairs on panicle.	(1) none to few, (2) moderate, (3) many.
25	Color of panicle.	(1) pink tint, (2) dull red, (3) bright red.
26	Flower fragrant.	(1) no, (2) yes.
27	Color of wilted petals.	(1) brown, (2) faint pink, (3) pink to red.
28	Position of fertile stamen to pistil.	(1) oblique, (2) parallel.
29	Relative height of staminodia to fertile stamen.	(1) less than .25, (2) .25 to .50, (3) over .50.
30	Shape of disc (receptacle).	(1) swollen, (2) narrow.
31	Length of fruit (mm).	(1) over 125, (2) 100 to 125, (3) under 100.
32	Ratio of length to mid-breadth of fruit.	(1) over 1.70, (2) 1.40 to 1.70, (3) under 1.40.
33	Weight of fruit (oz.).	(1) under 8, (2) 8 to 16, (3) over 16.
34	Fruit surface at stalk attachment.	(1) elevated, (2) level, (3) depressed.
35	Dorsal shoulder of fruit.	(1) abruptly falling, (2) moderating falling, (3) level.
36	Ventral shoulder of fruit.	(1) rising, (2) level, (3) falling.
37	Beak of fruit.	(1) prominent, (2) not prominent or missing.
38	Apex of fruit (breadth view).	(1) generally curved, (2) generally symmetrical.
39	Apex of fruit (thickness view).	(1) acute, (2) obtuse to rounded.
40	Fruit surface.	(1) smooth, (2) undulating.
41	Ground color of fruit skin.	(1) green, (2) greenish yellow, (3) bright yellow, (4) pink.
42	Amount of bloom on fruit.	(1) none, (2) slight, (3) medium, (4) heavy.
43	Color of bloom on fruit.	(1) gray, (2) purple.
44	Blush present on fruit.	(1) no, (2) yes.
45	Blush color of fruit.	(1) orange to orange red, (2) pink, (3) crimson to dark red.
46	Relative size of lenticels on skin.	(1) small, (2) large.
47	Relative frequency of lenticels on skin.	(1) none to few, (2) numerous.
48	Color of lenticels on skin.	(1) green, (2) white or yellow, (3) russet.
49	Relative skin thickness.	(1) thin, (2) intermediate, (3) thick.
50	Skin toughness.	(1) tender, (2) intermediate, (3) tough.
51	Skin adhesive to pulp.	(1) easily separating, (2) intermediate, (3) adheres.
52	Pulp color.	(1) pale to lemon yellow, (2) deep yellow to orange.
53	Pulp texture.	(1) soft, (2) intermediate, (3) firm.
54	Pulp tender.	(1) yes, (2) no.
55	Pulp melting.	(1) yes, (2) no.
56	Amount of fiber in pulp.	(1) none to scanty, (2) intermediate to abundant.
57	Relative texture of fiber.	(1) fine, (2) intermediate, (3) coarse.
58	Relative length of fiber.	(1) short, (2) intermediate, (3) long.
59	Aroma strength of fruit.	(1) none to weak, (2) medium to strong.
60	Aroma repugnant.	(1) no, (2) yes.
61	Carrot odor of fruit.	(1) yes, (2) no.
62	Fruit taste sweet.	(1) yes, (2) no.
63	Fruit taste spicy.	(1) yes, (2) no.
64	Fruit taste acid.	(1) no, (2) yes.
65	Fruit taste resinous.	(1) no, (2) yes.
66	Fruit has aromatic "flavor".	(1) no, (2) yes.
67	General flavor of fruit.	(1) rich, (2) mild.
68	General quality rating of fruit.	(1) excellent, (2) good, (3) fair, (4) poor.
69	Anthracnose infection on fruit.	(1) none to slight, (2) moderate, (3) severe.
70	Weight of stone (oz.).	(1) under 1.25, (2) 1.25 to 1.50, (3) over 1.50.
71	Endocarp texture.	(1) thin and papery, (2) intermediate, (3) thick and woody.
72	Percent seed size to stone cavity.	(1) under 50, (2) 50 to 90, (3) over 90.
73	Embryo type.	(1) usually polyembryonic, (2) usually monoembryonic.

OTU's, the phenogram (Fig. 1) based on distance coefficients gave a slightly better separation of the Southeast Asia and Indian germ plasm than the phenograms (not shown) based on correlation and divergence coefficients. Because we are concerned here with classifying the germ plasm of mangos rather than comparing coefficients to measure overall similarity, only phenograms based on distance coefficients will be pursued further.

Distance phenogram of introduced germ plasm. In Fig. 1 most of the Indian mangos are shown in the upper part of the diagram (OTU's 2 to 40, except for OTU 21 from Haiti). The next large cluster contains Southeast Asia mangos (OTU's 4 to 33). Another cluster contains OTU 22 (Java), OTU 34 (India), and OTU 39 (Cuba). OTU's 17 (Brazil) and 18 (Réunion) join secondary clusters separately and at low levels. OTU's 27 (West Indies) and 35 (India) cluster at a low level before joining other clusters. The 2 species *M. zeylanica* (OTU 42) and *M. odorata* (OTU 41) join individually to complete the phenogram. Because the latter 2 species do not appear to be closely related to the other OTU's, they will not be considered further.

Similarity diagram. Since a phenogram is a somewhat distorted representation of a similarity matrix, a diagram is shown in Fig. 2 to explain certain relationships among OTU's that are not apparent in the phenogram (Fig. 1). The closely related OTU's in the diagram are shown as a group. The groups were defined by the clusters of the

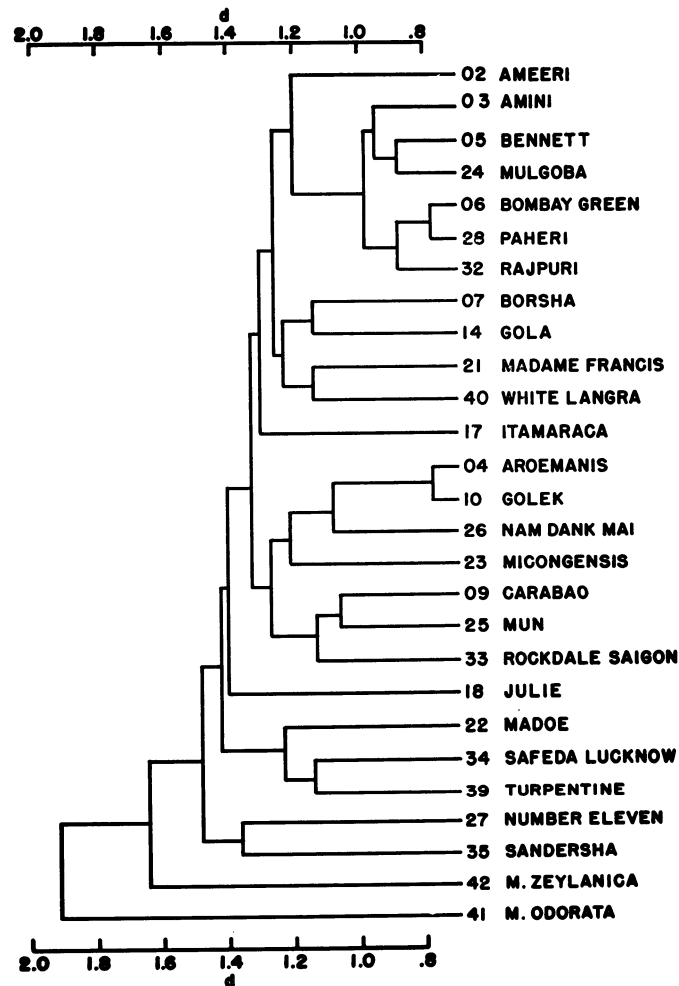


Fig. 1. Phenogram based on a UPGMA cluster analysis of a distance matrix containing 27 OTU's. See text for further explanation.

phenogram from Fig. 1. The lines with distance values taken from the distance coefficient matrix (not shown) connect the closest OTU's of each group. It is evident that certain OTU's or groups of OTU's are intermediate of the extreme types. For example, OTU 21 (Haiti) is intermediate to OTU's 4 (Java) and 40 (India). Likewise, OTU 26 (Thailand) is intermediate to OTU's 4 (Java), 7 (India), 34 (India), and 6 (India). OTU 34 (India) shows a close affinity to OTU's 39 (Cuba) and 22 (Java). Similar relationships not evident by the phenogram can be seen for OTU's 17 (Brazil) and 33 (Indochina). OTU's 35 (India), 27 (West Indies), and 18 (Réunion) are not shown because of their respective lack of close affinity to any OTU in the phenogram.

Florida complex. The relationships of the Florida hybrids and OTU 31 from Hawaii can best be shown by a tentative pedigree as given in Fig. 3. The reported parents (Table 1) of progeny are shown by solid arrows, and alternate pathways based on low distance coefficient values are shown by dashed arrows. Values alongside each line are taken from the distance coefficient matrix. Each arrow is drawn to be in accord with the location of the reported or possible parent(s) at the time of origin of a hybrid.

Two genetic complexes can be seen in Fig. 3. One complex begins with OTU's 35, 24, and 15, and the related descendants of this complex contain OTU's 8, 12, 16, 19, 20, 29, 31, 37 and 38. This complex can be generally characterized as large-fruited and late-maturity.

The second complex begins with OTU's 15 and 9 and contains OTU hybrids 11, 13, and 36, plus OTU 33. The high distance coefficient values between OTU 11 and its reported parents, OTU's 9 and 15, leave considerable doubt as to this parental relationship. No alternate parents for OTU 11 were indicated by the values of distance coefficients among the 42 OTU's.

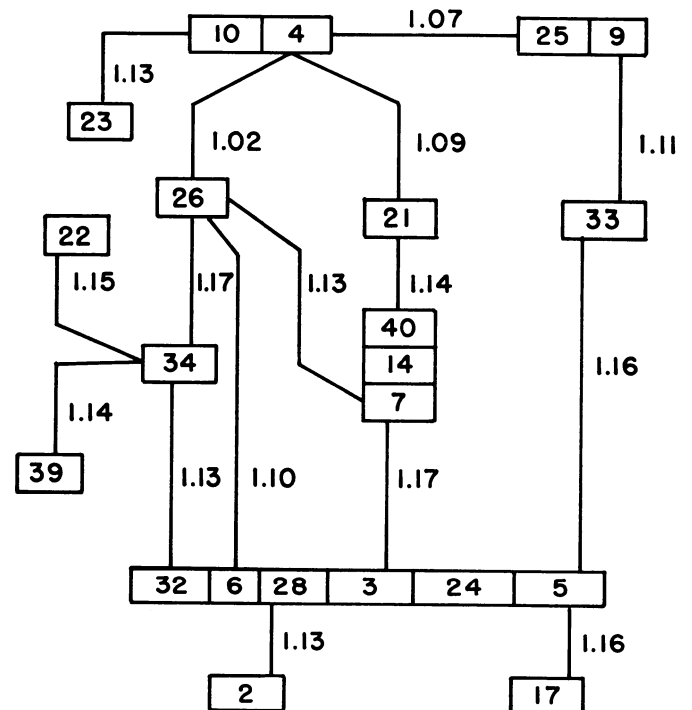


Fig. 2. Similarity diagram showing certain relationships among OTU's that are not apparent in the phenogram of Fig. 1. OTU's 18, 27, and 35 show no close relationship in the distance coefficient matrix of the 27 OTU's and they are not included in the diagram. See text for further explanation.

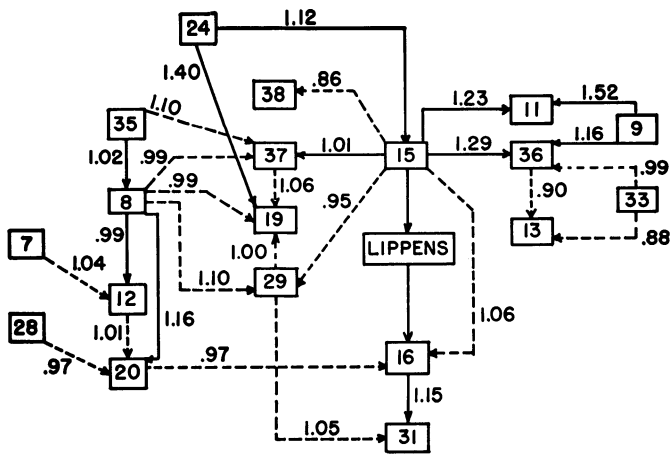


Fig. 3. Pedigree of reported parents of hybrids and alternate pathways based on a distance coefficient matrix. The enclosed code numbers represent the OTU parent(s) and progeny. The reported parents to progeny are shown by solid arrows and alternate pathways based on the distance coefficients matrix by dashed arrows. Values alongside each line are taken from the distance coefficient matrix.

Tentatively we might consider that the parents of OTU 11 should best be listed as questionable until we have data on other potential parents. Likewise, parent OTU 15 of OTU 36 and parent OTU 24 of OTU 19 might be listed as questionable because of their large distance coefficient value.

It appears that the results of the analysis can be used either to partially support or cast doubt on some reported parentages and to indicate alternate choices of suspected parents; however, methods such as Anderson's Hybrid Index (1) might be more suitable to determine probable parents. In Anderson's method, characters are so chosen that a bi-parental hybrid is intermediate in form to its parents. These characters are usually polygenetic in nature and show wide variation between suspected parents. Since the characters used in this study were not selected or recorded with this purpose in mind, we will not consider parentage further.

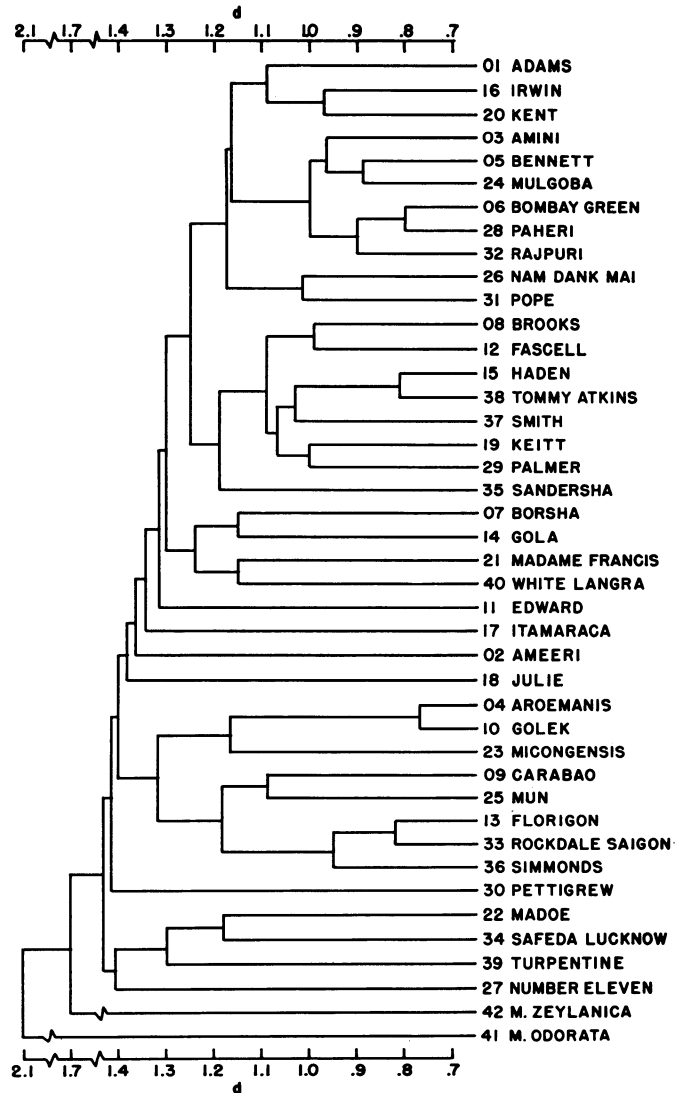


Fig. 4. Distance phenogram of 42 OTU's. See text for further explanation.

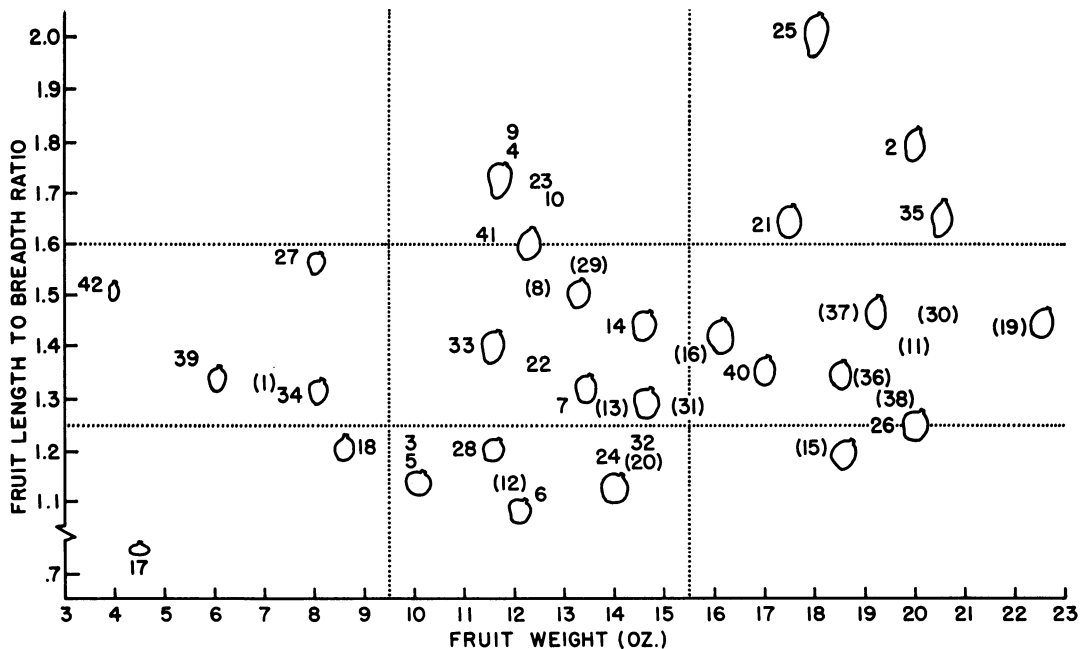


Fig. 5. Scatter diagram showing relationship of 42 OTU's as defined by character 32 (ratio of fruit length to breadth) and character 33 (weight of fruit).

Distance phenogram of 42 OTU's. The clusters of OTU's in Fig. 4 can be divided into 3 or 4 major complexes. In the upper part of the phenogram, OTU's 1 to 35 make up a large group of monoembryonic mangos, except for OTU 26. Within this major group, the cluster of OTU's 3 to 32 contain the roundish-shaped fruit of India. This group appears to be a combination of Popenoe's (5) 2 groups, Mulgoba and Alphonse.

The group of OTU's 8 to 35 plus OTU's 16 and 20 might be called the Sandersha-Haden complex because of the relationships shown in Fig. 3. Sandersha and Haden appear to have given their large size and elliptic-shape to members of this complex.

OTU's 7 to 40 tend to form a third group. They appear to be intermediate in form between the roundish-shaped fruit from India and the elongated fruit from Southeast Asia.

The Southeast Asia group contains OTU's 4 to 36. All are polyembryonic. These OTU's belong to the Cambodiana group of Popenoe (5).

None of the other OTU's in the Fig. 4 appear to belong to any particular group as they join at rather low levels in the phenogram.

Scatter diagram of 42 OTU's. The shape-size relationship of the 42 OTU's in Fig. 5 are defined by character 32 (ratio of fruit length to breadth) and character 33 (fruit weight). The outlines on the diagram indicate the general shape of a fruit of near-neighbor OTU's. The fruits are not drawn to scale. This diagram shows the general trend in shape and size of the mango genotypes in this study. The numbers in parenthesis are the Florida hybrids and the one developed in Hawaii from Florida germ plasm. Most of these OTU's are found in the center

and right center sections of the diagram. Apparently the Florida mangos have been selected for increased size.

Only about one-fourth of the mango genotypes being grown in Florida were included in this study. Future studies to include other genotypes might reveal additional groups or complexes. These studies should give more attention to determining the pedigrees of hybrids so that parental materials most likely to be of value in breeding for new genotypes could be better defined.

LITERATURE CITED

1. ANDERSON, E. 1949. *Introgressive hybridization*. New York, Wiley and Sons. 109 p.
2. CLARK, P. J. 1952. An extension of the coefficient of divergence for use with multiple characters. *Copeia* 2:61-64.
3. MUKHERJEE, S. K. 1948. The varieties of mango (*Mangifera indica* L.) and their classification. *Bul. Bot. Soc. Beng.* 2.
4. NAIK, K. C., and S. R. GANGOLLY. 1950. A monograph on classification and nomenclature of South Indian mangoes. Madras, Supt. Govt. Press.
5. POPENOE, W. 1920. *Manual of tropical and subtropical fruits*. MacMillan, New York.
6. RHODES, A. M., W. P. BEMIS, T. W. WHITAKER, and S. G. CARMER. 1968. A numerical taxonomic study of Cucurbita. *Brittonia* 20:251-266.
7. SINGH, L. B. 1960. *The mango: Botany, cultivation, and utilization*. London: L. Hill; New York: Interscience Publishers. 438 p.
8. ———, and R. N. SINGH. 1956. A monograph on the mangoes of Uttar Pradesh. 2 vols. Lucknow, Supt. of Printing.
9. SOKAL, R. R. 1961. Distance as a measure of taxonomic similarity. *Syst. Zool.* 10:70-79.
10. ———, and P. H. A. SNEATH. 1963. *Principles of numerical taxonomy*. W. H. Freeman and Co., San Francisco and London. 359 p.
11. VALMAYOR, R. V. 1962. *The mango: Its botany and production*. U. of Philippines, College of Agriculture, College, Laguna.