

# Genetic Diversity in Peach (*Prunus persica* L. Batch) Revealed by Randomly Amplified Polymorphic DNA (RAPD) Markers and Compared to Inbreeding Coefficients

Marilyn L. Warburton<sup>1</sup> and Fredrick A. Bliss<sup>2</sup>

Department of Pomology, University of California–Davis, Davis, CA 95616

Additional index words. PCR, random amplified polymorphic DNA, UPGMA, cluster analysis

**Abstract.** Previous studies of peach germplasm using pedigree information and isozyme polymorphism data have shown limited diversity in the U.S. gene pool. To further investigate the genetic diversity among peach cultivars grown in different regions of the United States, 94 RAPD markers were used to estimate the genetic distances among 136 cultivars. Of the 12 clusters formed in a dendrogram, the 90 U.S. cultivars and breeding lines and most of those from Europe and Latin America grouped to only three clusters, while the 23 peach entries from India, Pakistan, Russia, Okinawa, and China, as well as the almond cultivar used as an outgroup, were distributed among the other nine clusters. Therefore, the genetic diversity within temperate U.S. peach germplasm is quite limited, and to expand the variability, additional germplasm should be obtained, especially from Asia. Comparison of genetic similarity based on inbreeding coefficients with similarity coefficients based on the RAPD data produced a correlation of 0.395, which is comparable to values in similar investigations in other crops. Thus, similar conclusions can be drawn from these two sources of information. RAPD data are useful particularly when pedigree information is incomplete, there has been substantial selection within breeding populations, and a high proportion of alleles are identical in state but not by descent.

Peaches are grown extensively throughout the United States, and in some states, such as California, fresh-market and canning peaches are economically important. Although there are many peach cultivars, virtually all those in the United States are derived from relatively few accessions that were introduced into North America from Europe in the 18th century or from a single introduction from China known as ‘Chinese Cling’ or ‘Shanghai’ (Faust and Timon, 1995; Hedrick, 1917). While many of the peaches grown in different regions and used for different purposes share similar pedigrees, they have been subjected to natural and artificial selection for several decades. It would be informative to learn whether cultivars from different areas of the United States [i.e., the northern (including Canada), eastern (including mid-Atlantic), southeastern (including Georgia, Texas, and Florida), and western (primarily California)] have diverged genetically because of intercrossing and selection. If so, this divergence would indicate more genetic diversity in U.S. peach germplasm than expected from pedigree data. The apparent limited genetic diversity in the U.S. germplasm has been documented by Arulsekhar et al. (1986) and Byrne (1990) who found that many of the isozyme loci were monomorphic among the cultivars studied.

For continued improvement of peaches through breeding and for timely response to threats from diseases, insect pests, and abiotic stress, a diverse gene pool is essential. When there is insufficient diversity, the genetic variability should be increased by adding distinct, new accessions. Molecular markers can be

used to facilitate quantification of existing genetic diversity and uncovering duplicate or very similar genotypes and for identification of unique variants or genotypes for expanding the useful variation.

Among the different types of molecular markers available, RAPD markers are attractive because of their simplicity, versatility, modest cost, and ability to detect relatively small amounts of genetic variation (Ragot and Hoisington, 1993). They have been used to measure genetic diversity of many crop plants including cereals (Heun et al., 1994; Hilu, 1994; Yu and Nguyen, 1994), legumes (Abo-elwafa et al., 1995; Haley et al., 1994; Halward et al., 1992), brassicas and celery (dos Santos et al., 1994; Yang and Quiros, 1993), and others (Orozco-Castillo et al., 1994; Russell et al., 1993; Williams and St. Clair, 1993). Despite their favorable characteristics and wide experimental use, questions have been raised about the reliability of RAPD data due to their variable nature under different test conditions and the amplification of unrelated fragments of similar sizes leading to the unwarranted conclusion that they are genetically similar (e.g., Bakalinsky et al., 1994). These potential shortcomings can be minimized by demonstrating that the polymorphic RAPD bands scored in a diversity study segregate according to Mendelian expectations, are homologous to each other, and are repeatable for the experimental conditions used. Our objective was to assess the genetic variability among peach cultivars and breeding lines from different regions of the United States compared to 18 accessions from China, the center of diversity, and 19 other countries using RAPD markers.

## Materials and Methods

*Sources of materials.* One-hundred thirty-six peach cultivars from 20 countries and one almond cultivar were studied. Foreign sources were included as available, and those from the U.S. were chosen based on region of origin and use, with about equal numbers of cultivars from the western, southeastern, eastern, and northeastern regions (Table 1). Three internal checks in the form of DNA from the same named cultivar from different sources were analyzed.

Received for publication 15 Mar. 1996. Accepted for publication 14 June 1996. This research was submitted by M.L.W. in partial fulfillment of the requirements for the Ph.D. degree. We thank S. Arulsekhar for technical assistance and Paul Gepts and Carlos Quiros for their critical reading of the dissertation. This work was supported financially in part by the California Agricultural Expt. Station, a grant from the Univ. of California Systemwide Biotechnology Research and Education Program, and grant no. USDA 92-37300-7514 from the NRICG Program. The cost of publishing this paper was defrayed in part by the payment of page charges. Under postal regulations, this paper therefore must be hereby marked *advertisement* solely to indicate this fact.

<sup>1</sup>Postdoctoral researcher, Dept. of Vegetable Crops.

<sup>2</sup>Professor. To whom reprint requests should be addressed.

**DNA extraction and quantification.** Five grams of leaf material from a single tree were collected from each cultivar from different sources within the United States and Canada. DNA was extracted from the fresh, washed leaves according to Doyle and Doyle (1990), except that an SDS-Tris-EDTA extraction buffer was used instead of a CTAB buffer and the final precipitation in ethanol was eliminated. DNA was digested with RNase and stored in 10 mM Tris: 1 mM EDTA (pH 8.0) at 4 °C until ready for use. DNA was quantified either using a fluorometer or comparing staining intensity of DNA with ethidium bromide to a standard (lambda cut with *Hind* III) using the gel documentation and image analysis system (model IS-1000; Alpha Innotech Corp., San Leandro, Calif.).

**RAPD amplification.** DNA (5 µl with a concentration of 20 µg·mL<sup>-1</sup>) was amplified in a total volume of 25 µl according to Yu and Pauls (1992) in a thermal cycler (Perkin-Elmer-Cetus model; Perkin-Elmer, Foster City, Calif). DNA concentration was optimized before analysis to determine which concentration provided the best amplification for most of the primers tested. Primers for amplification were obtained from Operon Technologies (Alameda, Calif.) and the Univ. of British Columbia. Amplified DNA was run on 2% agarose gels and stained with ethidium bromide. RAPD primers were chosen based on either their expression in a population (PMP3) derived from a cross between two heterozygous peach parents (B8-23-16 × A104-155) or because they had been identified as markers linked to expressed loci in *Prunus* in a previous study (Warburton et al., 1996). Bright, repeatable bands that were amplified in one or both parents and segregated according to the expected Mendelian ratios in the progeny were used in the diversity study. Some bands that were monomorphic in PMP3 but were amplified by the same primers also were scored in the diversity study. Many of these also proved to be monomorphic in the diversity study.

**Diversity study.** DNA from all 136 peach and the almond entries was amplified twice with each primer and scored for presence or absence of each band independently by two people to reduce the possibility of subjective scoring. By repeating all measurements, an error rate for RAPD amplification and scoring was calculated as follows: error rate = [number of bands not scored the same in both amplifications ÷ total number of bands scored (equal to the number of bands × 137 individuals)] × 100. The bands that demonstrated the highest error rates were excluded from the study; thus, increasing the repeatability of the data set used for analysis.

Three sets of cultivars, presumed to be identical, were included in this study as internal checks to provide an estimate of repeatability of RAPD banding patterns. These cultivars were 'Belle' and 'Georgia Belle' (same clone under two names), 'Springcrest' from two sources, and duplicates of 'Bai Mang Pen' from China, which were included serendipitously after it was realized that two numbers (PI552925 and quarantine no. 25935) referred to the same introduction (i.e., one number referring to the plant introduction number and one to the quarantine station number of that cultivar).

**Statistical calculations.** The NTSYS-pc version 1.7 (Rohlf, 1992) computer program was used to generate a similarity matrix based on the simple matching algorithm (Sokal and Sneath, 1963). This algorithm considers two individuals, each of which do not have a band as a match (i.e., genetically similar). With RAPDs, lack of amplification is generally due to mutation of the primer site, which is only 20 base pairs, and so, two individuals lacking a certain band probably share the same mutation. The SAHN clustering program then was used to group the entries based on the similarity coefficients using the unweighted pair group method with arithmetic averages (UPGMA). The cophenetic correlation coefficient was calculated between the similarity matrix and the

cophenetic matrix using the MXCOMP procedure to deduce how well the tree represented the data matrix. Dendrograms constructed from the complete data set and subsets as described later in this paper also were compared using MXCOMP.

An R-analysis, which is a cluster analysis of the RAPD bands, rather than of the cultivars in the study, was performed on the data using NTSYS to determine informativeness of each band in the final analysis. This procedure is a cluster analysis of rows, and in this case, different bands produced by each primer, rather than of columns or genotypes of peaches (Romesburg, 1984). Subsets of the most informative data, as determined by this procedure, were used to generate dendrograms, which then were compared to dendrograms generated using the entire data set.

**Co-ancestry.** The coefficient of co-ancestry was calculated for all pairs of entries where pedigree information was known (80 individuals) using the procedure INBREED of SAS (SAS Institute, 1987). This produced a matrix that was compared to the similarity matrices generated from RAPD data for the same 80 individuals using the MXCOMP procedure of NTSYS. This program did not assume that open-pollinations were strictly self-pollinated, which would cause lower co-ancestry values between individuals than would have been calculated using a model of self-pollination and could lead to a lower correlation between co-ancestry values and similarity values based on RAPDs.

## Results and Discussion

**Amplification conditions.** For the diversity study, experimental conditions for amplification were optimized according to Yu and Pauls (1992). A concentration of 20 µg·mL<sup>-1</sup> of DNA was satisfactory and small changes in concentration did not affect the banding patterns of the primers (data not shown). However, large changes in concentrations (i.e., an order of magnitude) did affect the amplification, with too little DNA causing either reduced or no amplification of some bands and too much DNA producing a smearing effect, probably due to nonsequence-specific binding of the primers to the DNA. Therefore, accurate quantification of the DNA was essential for optimal amplification. Spectrophotometry did not produce satisfactory quantitative estimates because the DNA isolated from peach leaves appeared to have abundant polysaccharide contamination that sometimes increased the apparent values of DNA concentrations. Fluorometry provided a more accurate measurement of DNA concentrations. Analysis of DNA stain intensity in agarose gels and comparison to a standard sample (special character cut with *Hind* III) also gave satisfactory quantification as well as showing the quality of the DNA (e.g., degree of degradation and presence of contaminants).

**Expression of RAPDs.** Twenty-one RAPD primers were chosen for measuring diversity, and from these primers, 94 bands amplified in a repeatable manner and were scored. Of the 94 bands, 25 segregated in PMP3 according to expected Mendelian ratios (data not shown), and 67 bands were monomorphic and amplified in both parents and all progeny of PMP3. Two of the 94 bands did not amplify in this cross (did not occur in the parents and, thus, not in the progeny) but amplified in some of the cultivars in the diversity study and were scored in the diversity study because they were polymorphic, highly repeatable, and very bright.

The 21 primers were scored in the 137 entries in the diversity study, and of 94 bands total, 37 were polymorphic among the peach cultivars and an additional 26 were polymorphic among the peach and almond. The remaining 31 were completely monomorphic but were included in the diversity calculations because using only the polymorphic bands could provide a misleading representation of

Table 1. Peach and almond cultivars included in the diversity study. Regions where grown are included. E = eastern, N = northern, SE = southeastern, and W = western.

Cultivar	Region	Parent		Cultivar	Region	Parent	
		Maternal	Paternal			Maternal	Paternal
Admiral Dewey	E U.S.	Alexander	OP	Jerseyqueen	E, N U.S.	NJ67239	NJN17
Aguascalientes 12-12	Mexico			June Lady	W U.S.	Fortyniner	Gemfree
Andross	W U.S.	Dix 5A-1	Fortuna	Junegold	E, SE U.S.	Flamingo	Springtime
Appia	Italy	(Southland x	Pesco Noce #1) OP	Klamt	W U.S.	Dixon 1	Wiser
Armkling	W U.S.	Palomar X	Palomar X	Loadel	W U.S.	Lovell	OP
		Springtime	Springtime	Lola	USSR		
Babygold's #5	N U.S.	PI35201	(J.H.Hale x Goldfinch) OP	Loring	E, N U.S.	Frank	Halehaven
				Maravilha	SE U.S.	Sunred	(Okinawa x Hiland) OP
Baladi No. 1	Israel			May Grand	W U.S.	Red Grand X	
Baronesa	Brazil					Early Sungrand	F2
Belle	SE U.S.	Chinese Cling	OP	Maygold	E U.S.	Sunhigh	Southland
Bienvenida No. 65	Morocco			Monroe	E U.S.	Rio Oso Gem	Shippers Late Red X Sunhigh
Blake	E, SE U.S.	J.H.Hale	Primrose	No. 0195	Turkey		
Bolivian Cling	Bolivia			O'Henry	W U.S.	Merrill Bonanza	OP
Bowen	W U.S.	Dix 22A-5	Str14-24B	Okinawa	W U.S.		
Canadian Harmony	N U.S.	Redskin	Sunhaven	Panamint	W U.S.	Babcock X Boston	Gold Mine X Rio Oso Gem
Candor	E U.S.	Redhaven	Erlly-Red-Fre	Proskauer	Germany		
Carlos	Mexico			Quetta	Pakistan		
Carolyn	W U.S.	Libbee	Lovell	Redglobe	E, SE U.S.	W3-16	Fireglow
Carson	W U.S.	Maxine	Leader	Redhaven	NE, SE U.S.	Halehaven	Kalhaven
Chalpachu	Pakistan			Redskin	N, E U.S.	J.H.Hale	Elberta
Chinese Cling	China			Redtop	W U.S.	Sunhigh	July Elberta OP
Chinese Flat	Unknown			Rio Oso Gem	E U.S.		
Corona	W U.S.	Libbee	Lovell	Ross	W U.S.	GH8-14	D30-3E
Coronet	E, SE U.S.	FV 5-56	Dixigem	Rutger's Red Leaf	SE U.S.		
Cresthaven	N U.S.	Kalhaven	SouthHaven 309	Saharanpur No. 1	India		
Criollo	Guatemala			Southland	E U.S.	Halehaven	self
De Coosa	Italy			Springcrest (2nd)	W U.S.	FV 89-14	Springtime
Desertgold	W U.S.	(FV178-55 x FV244-2)	(FV131-48 x FV244-4)	Springtime	W U.S.	Luken's Honey X July Elberta	Robin
Dixired	E U.S.	Halehaven		St. John	E U.S.		
Dixon 1	W U.S.	Orange Cling	Australian Muir	Starn	W U.S.	Paloro	OP
Dr. Davis	W U.S.	D25-9E	G40-5E	Sundowner	SE U.S.	Columbina	Sunred OP
Early Crawford	E U.S.			Sungold	SE U.S.	NJ5107397	Okinawa X
Early Sungrand	W U.S.	Sungrand	OP	Panamint			
Elberta	E U.S.	Chinese Cling	OP	Sunhaven	E U.S.	Redhaven	SH 50
Elegant Lady	W U.S.	Early O'Henry	July Lady	Sunhigh	E U.S.	J.H.Hale	NJ40C5
Fantasia	W U.S.	Gold King	P101-24	Sunhome	SE U.S.	Sunred	FLA 15-97
Fay Elberta	W U.S.	Elberta	OP	Sunland	SE U.S.	FV 323-12	FV 9-6345
Firebite	W U.S.	Flavortop	F66-90	Sunprince	SE U.S.	Redglobe	FV9-6288
Flavorcrest	W U.S.	P53-68	FV 89-14	Sunred	SE U.S.	Panamint	FLA R9T10
Flavortop	W U.S.	Fairtime	OP	Tennessee Natural	SE U.S.		
Flordagold	SE U.S.	RioGrande	OP	Topaz	SE U.S.	Loring	self
Flordaking	SE U.S.	FLA 9-67	Early Amber	TropicSweet	SE U.S.	FLA 46-95	KGGold
Flordaprince	SE U.S.	FLA 2-7	Maravilha	Tufts	W U.S.	Dix 10A-4	Everts
Flordasun	SE U.S.	L 1-15	Springtime bud mut.	Walgant	South Africa		
Garnet Beauty	N U.S.	Redhaven		Washington	E U.S.	Freeman Fr. OP	Sunhigh
George IV	E U.S.			Ying Tsui Tao	China		
Georgia Belle	SE U.S.	Chinese Cling	OP	Yumyeong	Rep. of Korea		
Glohaven	E U.S.	SH 20	Kalhaven	552907	France		
Greensboro	SE U.S.	Connett	OP	552925 Bai Mang Pen	Hangzhou, China		
Hakuto	Japan			61-00-87	Czechoslovakia		
Halehaven	E U.S.	J.H.Hale	South Haven	101668 Ta Tao No. 6	Shantung, China		
Halford	W U.S.			62602 Pi Tao	Haitzu, China		
Harbelle	E, N U.S.	Sunhaven	self	55776	Yunnan, China		
Harbinger	N U.S.	Cherryred	NJ5601519	63850 Shail Seed	India		
Harblaze	N U.S.	Stark Delicious	Hardi	77876 Tos China #1	China via Italy		
Harbrite	N U.S.	Redskin	Sunhaven	240928 Mao Tao Seed	Shantung, China		
Harcrest	N U.S.	Redskin	H4219	95501	Shansi, China		
Hardi Red	N U.S.	Lexington	NJN32	134401 PI80089 Seed	Hangzhou, China		
Harken	E, N U.S.	Redskin	Sunhaven	25935 Bai Mang Pen	Hangzhou, China		
Harko	N U.S.	Lexington	NJN32	25938 Ta QiaoYi Hao	Hangzhou, China		
Harland	N U.S.	V37016	Earltred	25939 Xin Dai Jiu Ba	Xiangshen, China		
Harrow Beauty	N U.S.	Cresthaven	Harken	28934 Feicheng	Shantung, China		
Harrow Diamond	N U.S.	Redskin	Harbinger	505-8	China		
Harson	N U.S.	Redskin	Sunhaven	533 Chui Lum Tao	China		
Harvester	E, SE U.S.	Redskin	Southernlo	737 PI414385	Yunnan, China		
Hermasillo	SE U.S.	FLA 3-4N	FLA 5-5	705 PI55776	Yunnan, China		
Indian Cling	Unknown	Bl. Clingstone	OP	664 Tzim Pee Tao	China		
Indian Freestone	Unknown			6A-11 (Almond)	Unknown		
J.H.Hale	E U.S.	Elberta	self				
Jerseyglo	E U.S.	Jefferson	Loring				

the genetic diversity present in the germplasm. However, some statistical calculations were performed using only the polymorphic bands to ensure that a clear illustration of the relationships among cultivars could be gained.

The 39.4% polymorphism among peach cultivars in this study using RAPD markers was considerably higher than that reported previously with isozymes (Arulsekhar et al., 1986; Byrne, 1990). In the isozyme study by Arulsekhar et al., only one locus of 12 surveyed was polymorphic, an 8.3% polymorphism rate. That rate is too low to be practical for diversity studies, although it might have been higher if more cultivars from Asia had been studied. Percentages of polymorphism for RFLP markers in peach were 20% (four of 23) for genomic clones and 33% (three of nine) for cDNA clones (Eldredge et al., 1992). Thus, to our knowledge, the level of polymorphism we found using RAPDs is the highest reported in peach so far.

*Repeatability of RAPDs.* When duplicates of all RAPD loci had been amplified from all entries and scored independently by two people, an error rate was calculated to provide an estimate of repeatability of amplification. Of 9794 bands scored, 141 were scored differently between the first and second amplifications, an error rate of 1.4%. Errors could have been due to technician error, uneven amplification of RAPDs, or aberrant transmission of RAPDs (Pooler and Scorza, 1995). However, most of the misscored bands were from the same few loci, rather than being distributed randomly among all loci. When the data from the six most troublesome loci were removed, an error rate of 0.52% was calculated between the two amplifications. This error rate is sufficiently low to provide an unbiased representation of the relationships among the entries and demonstrates that RAPDs have a high ability to be repeated when important characteristics are controlled. Of the three internal checks, 'Belle'/'Georgia Belle' were identical for all RAPD bands scored, and the two 'Springcrest' clones were identical for all but two bands, which places the two clones in a close but not identical position on the dendrogram (Fig. 1). The two sources of 'Bai Mang Pen', however, were different for many bands, and although the two accessions fall into the same cluster, they should not be considered the same clone. Related introductions were probably mislabeled at some point. The identity of the 'Belle'-'Georgia Belle' clones and the close similarity of the 'Springcrest' clones indicate that RAPD bands have a high ability to be repeated but are subject to a low degree of error due to misscoring or uneven amplification. The overall patterns of genetic diversity measured by this method should not be affected by this low error rate, but using RAPDs for genetic fingerprinting should be subjected to more rigorous controls.

*Patterns of diversity.* Most of the diversity present was due to cultivars from Asia. Of 12 clusters formed in a dendrogram (Fig. 1), only three contain the 90 U.S. breeding lines and cultivars, as well as most European, Latin American, and Pacific Island accessions, which demonstrates the shared ancestry of peaches grown in these countries and reflects the cooperative nature of peach improvement among countries of these regions. The remaining nine clusters contain 23 clones from India, Pakistan, Russia, China, and Okinawa and the almond outgroup. In contrast to the similarity of U.S. germplasm, the diversity of the Chinese accessions is demonstrated because Chinese materials occur in all but the two small peach clusters III and IV.

The three clusters containing the U.S. cultivars are very similar, indicating that there has been limited divergence. One cluster (I) contains primarily western nonmelting clingstone cultivars, which show the highest degree of similarity of all the entries in this study,

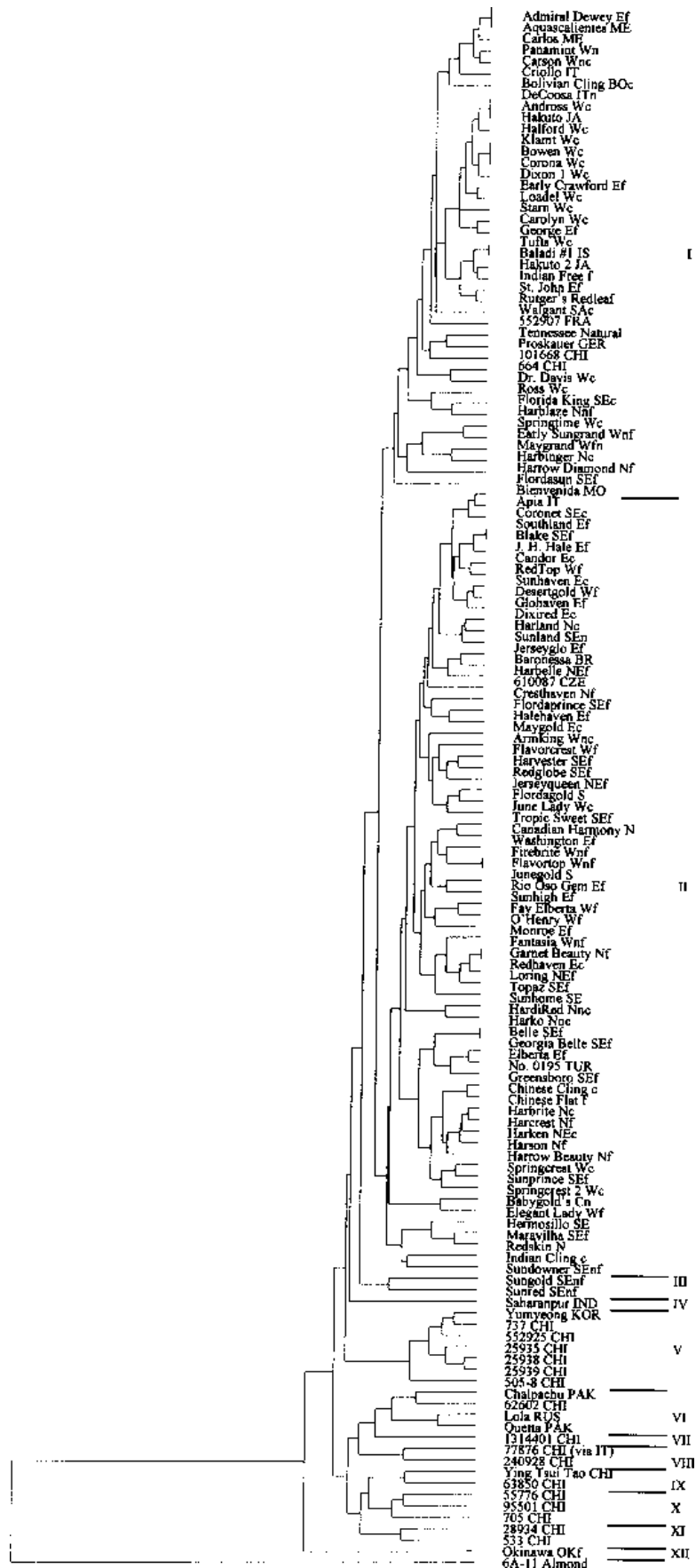
which is probably due to using only a few parents in breeding programs and the requirement for uniformity of canning peaches. To increase the genetic variability of cling peaches and to offer greater opportunities for enhanced breeding progress, the genetic base of these materials should be diversified. Most other U.S. peach germplasm is concentrated in cluster II, except for a few cultivars originating in Florida, which form cluster III. The Florida materials probably contain exotic germplasm originally included to reduce the winter chilling requirement, thus decreasing the similarity to other U.S. peach cultivars (Scorza et al., 1988). Cluster II is slightly more diverse than cluster I but this difference does not appear to be significant, which also underscores the need to expand the genetic base of U.S. peach germplasm. The U.S. materials in these three clusters are more similar to each other than to the clusters containing materials from China, India, and Pakistan (clusters IX to XII), suggesting that the original peaches introduced into the United States contained only a limited representation of the available diversity. Future introductions should concentrate on geographic areas where this and future studies indicate there would be additional diversity to complement the existing U.S. cultivars. Adding diversity also would avoid redundancy and reduce costly quarantine of duplicate or genetically similar samples.

Some cultivars could not be distinguished from each other by the RAPDs used in this study (Fig. 1). These included 'Bowen', 'Corona', and 'Dixon 1' (similarity coefficient = 1), which are western canning cultivars, with 'Bowen' and 'Dixon 1' derived from the cultivar 'Orange Cling'. The ancestors of 'Corona' are unknown before the immediate parents. 'Andross', 'Hakuto', and 'Halford' also had identical RAPD profiles, with 'Andross' and 'Halford' being western canning cultivars, but the pedigree of 'Halford' is unknown. To verify the location of the Japanese cultivar Hakuto in this cluster, DNA was re-isolated, and the analysis repeated because of the possibility that this cultivar had been mislabeled before or during the analysis. The second analysis placed 'Hakuto' in a different cluster (see 'Hakuto 2'). Thus, the results of the RAPD analysis with this entry cannot be verified. Finally, 'Admiral Dewey', 'Aguascalientes', and 'Carlos' were identical for the patterns produced by RAPD analysis. 'Aguascalientes' and 'Carlos' are from Mexico and may have a similar genetic background, but it is unclear why they were identical to 'Admiral Dewey', since pedigree information is not available.

'Chinese Cling' is a parent of numerous modern peach cultivars, many of which can be found in Cluster II very close to 'Chinese Cling' (Fig. 1). 'Belle'-'Georgia Belle', 'Elberta', 'Fay Elberta', and 'J. H. Hale' are all economically important descendants of 'Chinese Cling'. 'J. H. Hale' also is in the pedigree of many peach cultivars. It was suggested that 'J. H. Hale' could be either a seedling resulting from self-pollination or a bud mutation of 'Elberta'. From the results of our study, it is apparent that 'J. H. Hale' is not a bud mutation because if it were, it would be expected to be identical to 'Elberta' in most, if not all, of the RAPD amplification profiles. Most likely it is a seedling of 'Elberta' because it is similar to this cultivar, as are many of the cultivars containing 'J. H. Hale' in their pedigree (e.g., 'Blake', 'Halehaven', and 'Sunhigh').

The peach commonly referred to as 'Tennessee Natural' is represented by populations of wild peaches that grow in parts of the eastern and southeastern United States. It has been suggested that these stands are remnants of peaches introduced by the Spanish into the New World from Spain, which in turn were introduced from North Africa by the Moors and, therefore, would represent a

Fig. 1. Dendrogram of 136 peach and one almond cultivars based on banding patterns of 21 RAPD primers. Names of genotypes, regions of origin, and fruit type are included (E = eastern United States; W = western United States; SE = Southeastern United States; N = northern United States and Canada; NE = northeastern United States; ME = Mexico; IT = Italy; BO = Bolivia; IS = Israel; SA = South Africa; JA = Japan; GER = Germany; CHI = China; FRA = France; MO = Morocco; B = Brazil; CZE = Czechoslovakia; TUR = Turkey; PAK = Pakistan; RUS = Russia; KOR = Korea; OK = Okinawa; c = clingstone; f = freestone; n = nectarine).



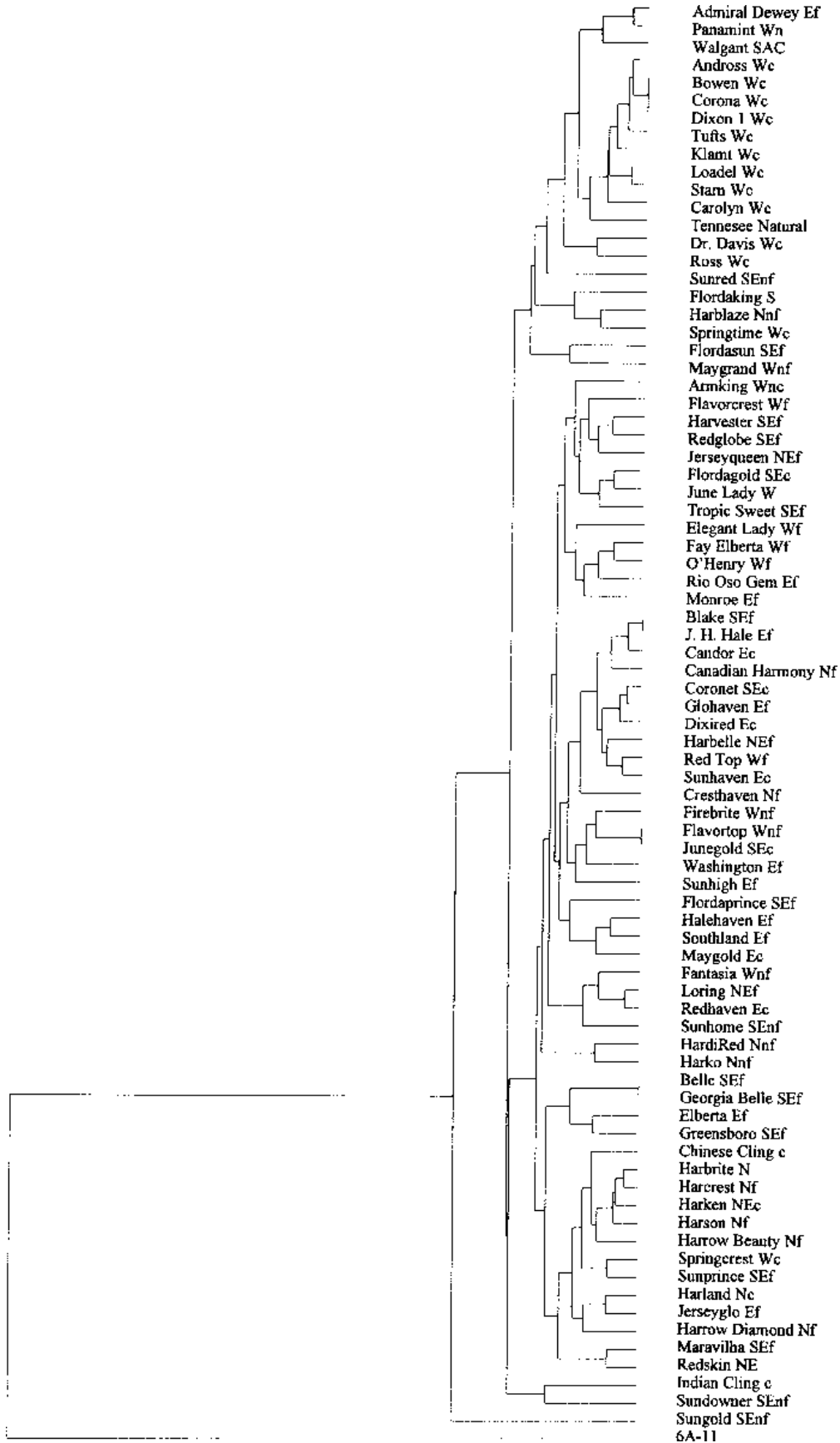


Fig. 2. Dendrogram of only those 80 peach individuals in the diversity study for which pedigree information was available. This tree was based on banding patterns of polymorphic bands scored for 21 RAPD primers. Names of cultivars and region of origin are included and are labeled as in Fig. 1.

potential source of diverse germplasm. However, the sample included in this study indicated that this population of 'Tennessee Natural' is very similar to the western canning cultivars (Cluster I) and probably consists of feral peaches that were derived from the same clingstone parents as modern peach cultivars, rather than from older, more diverged materials.

**Informativeness of individual RAPD markers.** Although a sufficiently large random sample of RAPD loci should provide an unbiased estimate of the genetic relationships within the germplasm, some of these markers probably will not provide new information and, thus, will be redundant for this type of study. Consequently, to use these markers most efficiently to determine genetic relationships among unknown peach genotypes and those already quantified, an R-analysis, which is a cluster analysis of the markers based on the peach entries, was performed on all the RAPD loci. Bands clustering tightly together would be expected to provide little additional information for distinguishing cultivars, whereas bands that clustered away from the others would provide unique information. Many of the RAPD markers fell into one large and nondiverse cluster, indicating that they are providing redundant information (data not shown). However, these results should not be misinterpreted, since this analysis included not only polymorphic RAPD bands but also monomorphic bands, which provide little unique information. The markers that clustered separately are those that will provide the most unique information and should be scored in future studies. The 16 most informative primers were OPA-1, OPA-4, OPA-9, OPB-17, BC-134, OPC-5, OPC-8, OPD-3, OPD-11, OPD-20, OPH-18, OPJ-1, OPJ-20, OPO-1, OPO-12, and OPO-19. The five least informative primers (OPC-10, OPD-8, OPF-5, OPO-3, and OPC-20) may not provide enough discriminating power to justify the added expense of using them in a survey. Removing the data from these five primers did not alter the final dendrogram; the similarity matrix produced using all the RAPDs and the similarity matrix produced using the 16 most informative primers were highly similar, with a correlation coefficient of  $r = 0.967$ . The Mantel- $t$  test for these data gave a value of 15.93 ( $P = 1.000$ ).

To measure the accuracy with which the dendrogram represented the similarity matrix constructed from the RAPD data, a cophenetic correlation coefficient was calculated. The correlation coefficient between the cophenetic matrix (the data in the dendrogram) and the simple matching similarity coefficient matrix for all the data was 0.894 ( $t = 14.906$ ,  $P = 1.000$ ), which indicated that the two matrices were in very close agreement.

A second dendrogram was constructed using only polymorphic RAPD markers (data not shown) to provide a larger separation between clusters and determine whether a different pattern of relationships between cultivars was depicted. A comparison of the two similarity matrices constructed from all RAPDs vs. only polymorphic RAPDs yielded a correlation coefficient of  $r = 0.992$  ( $t = 15.37$ ,  $P = 1.000$ ), indicating that the monomorphic bands neither add nor detract from the information provided by the polymorphic RAPD loci.

Two of the primers used in this study (OPO-19 and BC-134) amplified bands that segregated in a codominant fashion (i.e., 1 large molecular weight band : 2 both bands : 1 small molecular weight band). Scoring such bands can result in redundant information because scoring one band alone provides as much information as scoring all the codominant bands of a primer. In a predominantly self-pollinating species, many of the loci will be homozygous, and an individual homozygous at the RAPD locus of interest that is scored as present for one of these co-dominant alleles must be scored absent for the other allele. Therefore, a high percentage of

codominant RAPDs in a diversity study could bias the results by inadvertently covering the same locus in the genome more than once, rather than providing random coverage of the genome. Because most researchers using RAPD loci for this type of study do not know *a priori* how they segregate, this type of redundancy cannot be avoided. To determine the impact of the codominant RAPDs on the estimate of genetic diversity presented in this study, the entire statistical analysis was run with all bands scored simply as present or absent, then again using only one allele of each of the codominant RAPDs and discarding the alternate allele. Some information in the heterozygotes is lost in this manner, but as most cultivars were homozygous, the loss was minimal. When the similarity coefficient matrices for the two data sets (one with all bands and one missing one-half the codominant alleles) were compared, a correlation coefficient of  $r = 0.994$  ( $t = 15.4$ ,  $P = 1.000$ ) was obtained. In this case, no bias occurred by including both alleles of the codominant loci. Therefore, if a sufficient number of RAPD loci are scored in this type of diversity study, the issue of codominant RAPDs can be disregarded.

**Comparison of coefficients of co-ancestry with RAPD data.** The coefficients of co-ancestry for cultivars of known pedigree were calculated. Two previous inbreeding studies by Scorza et al. (1985, 1988) included two subsets of the genotypes presented in this study, one containing cultivars and breeding lines from Florida, and one containing cultivars and breeding lines from the eastern United States. Our observations agree with the results from those studies. A dendrogram was constructed using only those 80 cultivars for which pedigree information is available and is presented in Fig. 2.

When the matrix of co-ancestry coefficients was compared with the matrix of similarity coefficients based on RAPDs (i.e., the matrix that produced the dendrogram in Fig. 2), a correlation coefficient of 0.395 was obtained ( $t = 7.961$ ,  $P = 1.000$ ). While this statistic is low, it is comparable to values commonly reported for correlations between matrices based on data from different types of marker vs. pedigree information (Cox et al., 1985; Hancock et al., 1994; Heun et al., 1994; Mannen et al., 1993; Moser and Lee, 1994; Smith et al., 1992). Several factors would reduce the correlation between a similarity matrix calculated by RAPDs (or any molecular marker) vs. a matrix based on pedigree information, including, but not limited to, incomplete or inaccurate pedigree records, selection of breeding lines and cultivars, and a high percentage of alleles that are identical in state but not identical by descent. Although pedigree information would be the best determinant of genetic relationships, pedigree or passport data are often incomplete or even non-existent for peach accessions in the U.S. collections or those being introduced. Therefore, RAPD data should be useful in determining genetic similarities of unknown clones compared to peach accessions in the U.S. collections. This will allow curators of collections and breeders interested in improving existing materials to decide whether new clones that provide unique, and, therefore, desirable variability should be included in the collection.

#### Literature Cited

- Abo-elwafa, A., K. Murai, and T. Shimada. 1995. Intra- and inter-specific variations in *Lens* as revealed by RAPD markers. *Theor. Appl. Genet.* 90:335-340.
- Arulsekhar, S., D.E. Parfitt, and D. E. Kester. 1986. Comparison of isozyme variability in peach and almond cultivars. *J. Hered.* 77:272-274.
- Bakalinsky, A.T., H. Xu, D.J. Wilson, and S. Arulsekhar. 1994. Random amplified polymorphic DNA markers are inadequate for fingerprinting grape rootstocks. *HortScience* 29:528. (Abstr. no 666.)
- Byrne, D.H. 1990. Isozyme variability in four diploid stone fruits com-

- pared with other woody perennial plants. *J. Hered.* 81:68–71.
- Cox, T.S., Y.T. Kiang, M.B. Gorman, and D.M. Rodgers. 1985. Relationship between coefficient of parentage and genetic similarity indices in the soybean. *Crop Sci.* 25:529–532.
- dos Santos, J.B., J. Nienhuis, P. Skroch, J. Tivang, and M.K. Slocum. 1994. Comparison of RAPD and RFLP genetic markers in determining genetic similarity among *Brassica oleracea* L. genotypes. *Theor. Appl. Genet.* 87:909–915.
- Doyle, J.J. and J.L. Doyle. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12(1):13–15.
- Eldredge, L., R. Ballard, W.V. Baird, A. Abott, P. Morgens, A. Callahan, R. Scorza, and R. Monet. 1992. Application of RFLP analysis to genetic linkage mapping in peaches. *HortScience* 27:160–163.
- Faust, M. and B. Timon. 1995. Origin and dissemination of peach, p. 331–380. In: J. Janick (ed.). *Hort. Rev.* vol. 17. Wiley, New York.
- Haley, S.D., P.N. Miklas, L. Afanador, and J.D. Kelly. 1994. Random amplified polymorphic DNA (RAPD) marker variability between and within gene pools of common bean. *J. Amer. Soc. Hort. Sci.* 119:122–125.
- Halward, T., T. Stalker, E. LaRue, and G. Kochert. 1992. Use of single-primer DNA amplifications in genetic studies of peanut (*Arachis hypogaea* L.). *Plant Mol. Biol.* 18:315–325.
- Hancock, J.F., P.A. Callow, and D.V. Shaw. 1994. Randomly amplified polymorphic DNA (RAPD's) in the cultivated strawberry, *Fragaria ×ananassa*. *J. Amer. Soc. Hort. Sci.* 119:862–864.
- Hedrick, U. P. 1917. The peaches of New York. J.B. Lyon, Albany, N.Y.
- Heun, M., J.P. Murphy, and T.D. Phillips. 1994. A comparison of RAPD and isozyme analyses for determining the genetic relationships among *Avena sterilis* L. accessions. *Theor. Appl. Genet.* 87:689–696.
- Hilu, K.W. 1994. Evidence from RAPD markers in the evolution of *Echinochloa* millets (*Poaceae*). *Plant Syst. Evol.* 189:247–257.
- Mannen, H., S. Tsuji, F. Mukai, N. Goto, and S. Ohtagaki. 1993. Genetic similarity using DNA fingerprinting in cattle to determine relationship coefficient. *J. Hered.* 84:166–169.
- Moser, H. and M. Lee. 1994. RFLP variation and genealogical distance, multivariate distance, heterosis and genetic variance in oats. *Theor. Appl. Genet.* 87:947–956.
- Orozco-Castillo, C., K.J. Chalmers, R. Waugh, and W. Powell. 1994. Detection of genetic diversity and selective gene introgression in coffee using RAPD markers. *Theor. Appl. Genet.* 87:934–940.
- Pooler, M. and R. Scorza. 1995. Aberrant transmission of RAPD markers in haploids, doubled haploids, and F<sub>1</sub> hybrids of peach: Observations and speculation on causes. *Scientia Hort.* 64:233–241.
- Ragot, M. and D.A. Hoisington. 1993. Molecular markers for plant breeding: Comparisons of RFLP and RAPD genotyping costs. *Theor. Appl. Genet.* 86:975–984.
- Rohlf, F.J. 1992. NTSYS-pc: Numerical taxonomy and multivariate analysis system, version 1.70. Applied Biostatistics, New York.
- Romesburg, H.C. 1984. General features of cluster analysis, p. 29–37. In: Kerk Sargent (ed.). *Cluster analysis for researchers*. Wadsworth, Belmont, Calif.
- Russell, J.R., F. Hosein, E. Johnson, R. Waugh, and W. Powell. 1993. Genetic differentiation of cocoa (*Theobroma cacao* L.) populations revealed by RAPD analysis. *Mol. Ecol.* 2:89–97.
- SAS Institute. 1987. SAS guide. SAS Inst., Cary, N.C.
- Scorza, R., S.A. Mehlenbacher, and G.W. Lightner. 1985. Inbreeding and coancestry of freestone peach cultivars of the eastern United States and implications for peach germplasm improvement. *J. Amer. Soc. Hort. Sci.* 110:547–552.
- Scorza, R., W.B. Sherman, and G.W. Lightner. 1988. Inbreeding and coancestry of low chill short fruit development period freestone peaches and nectarines produced by the University of Florida breeding program. *Fruit Var. J.* 42:79–85.
- Smith, J., C. Stephen, O.S. Smith, S. Wright, S.J. Wall, and M. Walton. 1992. Diversity of U.S. hybrid maize germplasm as revealed by restriction fragment length polymorphisms. *Crop Sci.* 32:598–604.
- Sokal, R.R. and P.H.A. Sneath. 1963. Principles of numerical taxonomy. Freeman, San Francisco.
- Warburton, M.L., V. Becerra-Velásquez, J.C. Goffreda, and F.A. Bliss. 1996. Utility and behavior of RAPD markers in identifying genetic linkages to traits of economic interest in peach. *Theor. Appl. Genet.* (In press.)
- Williams, C. E. and D.A. St. Clair. 1993. Phenetic relationships and levels of variability detected by restriction fragment length polymorphism and random amplified polymorphic DNA analysis of cultivated and wild accessions of *Lycopersicon esculentum*. *Genome* 36:619–630.
- Yang, X. and C. Quiros. 1993. Identification and classification of celery cultivars with RAPD markers. *Theor. Appl. Genet.* 86:205–212.
- Yu, K. and K. P. Pauls. 1992. Optimization of the PCR program for RAPD analysis. *NAR* 20:2602.
- Yu, L.-X. and H. T. Nguyen. 1994. Genetic variation detected with RAPD markers among upland and lowland rice cultivars (*Oryza sativa* L.). *Theor. Appl. Genet.* 87:668–672.