# Genetic diversity of Carica papaya as revealed by AFLP markers

## Author

Kim, MS, Moore, PH, Zee, F, Fitch, MMM, Steiger, DL, Manshardt, RM, Paull, RE, Drew, RA, Sekioka, T, Ming, R

## Published

2002

## Journal Title

Genome

## DOI

10.1139/G02-012

# Genetic diversity of *Carica papaya* as revealed by AFLP markers

**M.S. Kim, P.H. Moore, F. Zee, M.M.M. Fitch, D.L. Steiger, R.M. Manshardt, R.E. Paull, R.A. Drew, T. Sekioka, and R. Ming**

**Abstract**: Genetic relationships among *Carica papaya* cultivars, breeding lines, unimproved germplasm, and related species were established using amplified fragment length polymorphism (AFLP) markers. Seventy-one papaya accessions and related species were analyzed with nine *Eco*RI–*Mse*I primer combinations. A total of 186 informative AFLP markers was generated and analyzed. Cluster analysis suggested limited genetic variation in papaya, with an average genetic similarity among 63 papaya accessions of 0.880. Genetic diversity among cultivars derived from the same or similar gene pools was smaller, such as Hawaiian Solo hermaphrodite cultivars and Australian dioecious cultivars with genetic similarity at 0.921 and 0.912, respectively. The results indicated that self-pollinated hermaphrodite cultivars were as variable as open-pollinated dioecious cultivars. Genetic diversity between *C. papaya* and six other *Carica* species was also evaluated. *Carica papaya* shared the least genetic similarity with these species, with an average genetic similarity of 0.432; the average genetic similarity among the six other species was 0.729. The results from AFLP markers provided detailed estimates of the genetic variation within and among papaya cultivars, and supported the notion that *C. papaya* diverged from the rest of *Carica* species early in the evolution of this genus.

*Key words*: DNA fingerprinting, germplasm, genetic relationship, molecular phylogeny, polymorphism.

**Résumé** : Les relations génétiques au sein du *Carica papaya* (entre cultivars, lignées en sélection, ressources génétiques non-améliorées) et avec des espèces voisines ont été examinées à l'aide du polymorphisme de longueur des fragments amplifiés (AFLP). Soixante et onze accessions de la papaye ou d'espèces voisines ont été analysées à l'aide de neuf combinaisons d'amorces *Eco*RI-*Mse*I. Au total, 186 marqueurs informatifs ont été produits et analysés. Une analyse de groupement a suggéré qu'il existe peu de variation génétique chez la papaye, l'indice moyen de similitude génétique étant de 0,880 parmi 63 accessions de la papaye. La diversité génétique entre cultivars dérivés d'un germoplasme identique ou semblable était plus faible encore. Par exemple, les cultivars hermaphrodites 'Hawaiian Solo' et les cultivars dioïques australiens présentaient des indices de similitude génétique de 0,921 et 0,912, respectivement. Les résultats indiquent que les cultivars hermaphrodites à autofécondation sont aussi variables que les cultivars dioïques à fécondation croisée. La diversité génétique entre le *C. papaya* et six autres espèces du genre *Carica* a également été évaluée. Le *C. papaya* était l'espèce la moins semblable aux autres, montrant un indice de similitude génétique de 0,432, tandis que l'indice moyen parmi les autres espèces était plutôt de 0,729. Les résultats obtenus avec les marqueurs AFLP permettent d'estimer avec précision la variation génétique parmi les cultivars de papaye et viennent appuyer l'hypothèse selon laquelle le *C. papaya* aurait divergé des autres espèces du genre *Carica* relativement tôt au cours de l'évolution.

*Mots clés* : empreintes génétiques, germoplasme, relations génétiques, phylogénie moléculaire, polymorphisme.

[Traduit par la Rédaction]

## Introduction

The tropical fruit crop papaya is a dicotyledonous, polygamous, diploid species with a small genome and nine pairs of chromosomes (Arumuganathan and Earle 1991; Purseglove 1968; Storey 1976). It belongs to the small family Caricaceae, which consists of five genera and 34 species. The family includes 21 species of *Carica* L. (the genus to which papaya belongs), seven species of *Jacaratia* A. DC., three species of *Jarilla* Rusby, two species of *Cylicomorpha*

Received 8 November 2001. Accepted 5 February 2002. Published on the NRC Research Press Web site at http://genome.nrc.ca on 5 April 2002.

Corresponding Editor: G.J. Scoles.

**D.L. Steiger and R. Ming**[1]. Hawaii Agriculture Research Center, Aiea, HI 96701, U.S.A.
**P.H. Moore, F. Zee, and M.M.M. Fitch.** USDA–ARS, Pacific Basin Agricultural Research Center, Hilo, HI 96720, U.S.A.
**M.S. Kim, R.M. Manshardt, R.E. Paull, and T. Sekioka.** Department of Tropical Plant and Soil Sciences, University of Hawaii, Honolulu, HI 96822, U.S.A.
**R.A. Drew.** School of Biomolecular and Biomedical Sciences, Nathan Campus Griffith University, Queensland 4111, Australia.

[1]Corresponding author (e-mail: rming@harc-hspa.com).

 DOI: 10.1139/G02-012 

Urban, and one species of *Horovitzia* Badillo (Badillo 1993). Papaya is believed to have originated in Central America, probably in the regions of southern Mexico and Costa Rica (Candolle 1908; Purseglove 1968; Storey 1976), and is now distributed throughout tropical and subtropical regions worldwide. The Caricaceae belong to a group of plant families characterized by production of glucosinolates. Morphological, biochemical, and molecular evidence from this clade has demonstrated a common phylogenetic origin with the Brassicaceae, and consequently with the completely sequenced model plant *Arabidopsis* (Bremer et al. 1998, Rodman et al. 1996).

Papaya germplasm shows considerable phenotypic variation for many horticulturally important traits, including fruit size, fruit shape, flesh color, flavor and sweetness, length of juvenile period, plant stature, stamen carpellody, and carpel abortion. In addition, commercial papaya cultivars may be inbred gynodioecious lines, typified by the Hawaiian Solo lines (Solo refers to a group of small-fruited, high sugar content, commerical cultivars developed in Hawaii and originating via introduction from Barbados in 1910 (Storey 1969)); out-crossing dioecious populations, such as the Australian papaws from southern Queensland; $F_1$ hybrids, including the Tainung series (Taiwan), Eksotica II (Malaysia), and Rainbow (Hawaii); or occasionally even clones, such as Hortus Gold in South Africa.

The only published study of genetic diversity in papaya germplasm examined a small group of seven Hawaiian Solo cultivars and three unrelated lines, using random amplified polymorphic DNA (RAPD) markers (Stiles et al. 1993). The objective was to test the validity of relationships revealed by RAPD data in germplasm of known pedigree, rather than to provide a broad overview of papaya diversity.

Genetic relationships between papaya and related wild species have been investigated using RAPD and isozyme markers (Jobin-Décor et al. 1997) and restriction fragment length polymorphism (RFLP) in a cpDNA intergenic-spacer region (Aradhya et al. 1999), as well as amplified fragment length polymorphism (AFLP) markers (Van Droogenbroeck et al. 2002). These three reports suggested that *C. papaya* diverged from the rest of the genus early in its evolution. Aradhya et al. (1999) noted that the wild South American *Carica* species are more closely allied with a member of the related genus *Jacaratia* than to *C. papaya*. More pertinently, no attempt was made to survey genetic variation within *C. papaya*.

AFLP has been recognized as a reliable and efficient DNA marker system, compared with RFLP, RAPD, or microsatellites (Vos et al. 1995; Powell et al. 1996; Russell et al. 1997; Pejic et al. 1998). AFLP markers have been used extensively for studying genetic diversity in different plant species because of their high reproducibility and multiplex ratio (Maughan et al. 1996; Ellis et al. 1997; Breyne et al. 1999; Erschadi et al. 2000). The objectives of this project were as follows: (*i*) to determine patterns of genetic relationship in a large collection spanning different geographic origins and mating systems (inbreeding and (or) out-crossing); (*ii*) to compare AFLP results with previous, smaller papaya germplasm characterization research using chloroplast RFLPs, RAPDs, and isozymes; and (*iii*) to examine intravarietal variation in papaya.

## Materials and methods

### Plant materials

Seventy-one accessions were collected from different countries for DNA fingerprinting, including 63 accessions of *Carica papaya*, 6 accessions of related species, and 2 accessions of interspecific hybrids. Among the 63 papaya accessions, 18 were commercial cultivars, 15 were improved (but not released) breeding lines, and 30 were unimproved germplasm. Multiple samples of selected accessions were collected to evaluate the genetic variation within each of those accessions, and each sample was collected from a single tree. One-hundred and nine samples were collected from 71 accessions. This collection represents cultivars and breeding lines adapted to different climates in tropical, subtropical, and temperate regions and unimproved germplasm from the center of origin. Because the genetic relationships of *Carica* species have been reported previously (Aradhya et al. 1999; Van Droogenbroeck et al. 2002), the focus of this study was the genetic relationships of commercial cultivars, breeding lines, and unimproved germplasm within *C. papaya*. Six *Carica* species were included as out groups. The detailed collection sites and geographic origins of these samples are listed in Table 1.

### DNA extraction

Young papaya leaf tissue from a single plant in each sample was lyophilized and homogenized to a fine powder. Genomic DNA was extracted based on the procedure described by Chittenden et al. (1994) with minor modifications. DNA samples were purified and quantified before restriction digestion.

### AFLP analysis

#### *Genomic DNA digestion*

The protocol outlined by Vos et al. (1995) was used with the following adjustment in the digestion conditions: 250 ng of DNA at 37°C for 3 h incubated with 5 U of *Eco*RI and 5 U of *Mse*I.

#### *Adapter ligation and preamplification*

The reaction conditions described by Vos et al. (1995) were adapted with the exception that 1 U *Taq* polymerase was used.

#### *Primer labeling and selective amplification*

IRDye™-700 and -800 labeled *Eco*RI primers (listed in Table 2) were purchased from LI-COR (Lincoln, Nebr.). *Mse*I primers were obtained from Operon Technologies (Alameda, Calif.). The reaction conditions described in the manufacturer's instructions were followed. A preliminary survey consisting of five papaya accessions was performed with $^{32}$P-labeled primers to select the most polymorphic primer sets for fingerprinting the entire collection.

#### *Gel analysis*

An equal volume of loading dye (95% v/v formamide and 0.08% w/v bromophenol blue in 20 mM EDTA) was added to each sample, which was then denatured at 95°C for 3 min and placed on ice for 2 min before loading. The sample volume was set to 1.0 μL and loaded with a Hamilton multi-

channel pipette (Hamilton, Reno, Nev.) onto 25-cm gels prepared with 6.5% KB Plus Polyacrylamide Gel Matrix (LI-COR, Lincoln, Nebr.). Electrophoresis and detection of the AFLP fragments were performed on a LI-COR IR$^2$ automated DNA sequencer (LI-COR, Lincoln, Nebr.). The electrophoresis parameters were set to 1500 V, 40.0 mA, 40.0 W, 50°C, and a scan speed of 4. The run time was set to 2.0 h and gel images were saved as TIF files for further analysis.

**Data analysis**

The gel images were scored using a binary scoring system that recorded the presence and absence of bands as 1 and 0, respectively. Semi-automated scoring was performed, initially with AFLP-Quantar™ (version 1.0, KeyGene Product, Wageningen, The Netherlands) and then manually to make adjustments to the automated score sheet when necessary. The copy-score function in the software allowed for scoring of subsequent images with the same primer combinations. The data were exported into a spreadsheet and formatted for the NTSYSpc (v. 2.1) cluster analysis software (Exeter Software Co., New York). Monomorphic markers were not scored. The data for 109 samples from 71 accessions were used to compute pair-wise simple matching coefficients (Sokal and Michener 1958). Cluster analysis was performed on the similarity matrix using the unweighted pair group method using arithmetic means (UPGMA) algorithm (Sneath and Sokal 1973) provided in the software package NTSYSpc. For comparison among accessions with multiple samples, the average similarity score for a particular accession was used. The cophenetic correlation coefficient was calculated to test the goodness of fit between the similarity and cophenetic matrices.

## Results

The initial screen for polymorphisms was carried out on five papaya cultivars, including four closely related Solo cultivars (‘Kapoho’, ‘SunUp’, ‘Sunset’, and ‘Higgins’), and the unrelated, dioecious ‘Washington’. The number of polymorphic markers generated by each primer set ranged from 0 to 9 with an average of 3.2. Among the 64 sets of *Eco*RI–*Mse*I primers with three nucleotides extension, the eight most polymorphic were selected for genotyping 109 samples. Each primer set generated 11–36 markers. An additional primer set consisting of an *Eco*RI primer with three-nucleotides extension and *Mse*I primer with only two-nucleotides extension was used to generate 20 markers. A total of 186 polymorphic markers (42%) from 445 readable bands were detected by nine primer sets among 109 samples (Table 2).

The genetic variation among 63 papaya accessions was estimated using pair-wise comparison of genetic similarity (Fig. 1). The average pairwise genetic similarity was 0.880 and ranged from 0.741 to 0.978. About 82% of the pair-wise comparisons among papaya accessions exhibited genetic similarity greater than 0.85; less than 4% showed genetic similarity lower than 0.80. The most closely related cultivars sharing genetic similarity of 0.978 were ‘Kamiya’, ‘Line 8’, ‘Waimanalo’, and ‘Maunawili Sweet’. The least related were the dioecious ‘Coimbatore 7’ and a temperate–subtropical dioecious cultivar from Australia, ‘AU25’, with genetic similarity of 0.741.

Pair-wise comparisons of genetic similarity among Hawaiian Solo papaya cultivars and breeding lines revealed narrow genetic diversity reflecting the consequence of inbreeding from a limited gene pool. The overall average genetic similarity of the 15 Solo accessions was 0.921, ranging from 0.859 to 0.978 (Fig. 2). The least similar Solo cultivars were ‘Line 2’ and ‘Sunset’, with the most similar being ‘Kamiya’, ‘Line 8’, ‘Waimanalo’, and ‘Maunawili Sweet’ (Table 3). Among the Solo cultivars with multiple samples, the least variable appeared to be ‘SunUp’, with an average genetic similarity of 0.976 ranging from 0.940 to 1. The most variable cultivar was ‘Sunset’, with an average genetic similarity of 0.939 ranging from 0.903 to 0.973. The hybrid ‘Rainbow’ showed a higher level of variability than either of its parents, ‘Kapoho’ and ‘SunUp’, with an average genetic similarity of 0.950 ranging from 0.914 to 0.978.

The level of genetic variation among dioecious cultivars was similar to that of the hermaphrodite cultivars. Among eight dioecious cultivars collected from tropical, subtropical, and temperate regions in Australia, the average genetic similarity was 0.914 and ranged from 0.839 to 0.962. The average similarity between Israeli dioecious accession AU9 and eight Australian cultivars was 0.883, less than the average similarity among Australian cultivars. The genetic similarity between two individuals of ‘AU1’ was 0.978 (Table 4).

To test the genetic relationship of closely related Hawaiian Solo cultivars and breeding lines, cluster analysis was conducted on 15 Solo accessions with multiple samples from five well-established cultivars. The resulting dendrogram showed a clear separation of two major clusters (Fig. 3). The first cluster included ‘Kapoho’, ‘Higgins’, ‘Line 40’, ‘Line 116’, ‘Kamiya’, ‘Line 8’, ‘Waimanalo’, ‘Maunawili Sweet’, and 4-16-35. Within this major cluster, 12 ‘Kapoho’ samples, mostly collected from eight different papaya farms on the island of Hawaii, formed a sub-cluster, suggesting a close genetic relationship among these samples. ‘Kamiya’, ‘Line 8’, ‘Waimanalo’, and ‘Maunawili Sweet’ formed another subcluster with equal genetic similarity (0.978) among them. Another four ‘Kapoho’ samples, collected on Oahu, formed a subcluster that differed slightly from samples collected on the island of Hawaii. The second major cluster included ‘SunUp’, ‘Sunset’, ‘Sunrise’, and ‘Rainbow’. ‘SunUp’, ‘Sunset’, and ‘Sunrise’ are sister lines from the same parentage, whereas ‘Rainbow’ is a hybrid between ‘Kapoho’ and ‘SunUp’. Three groups of samples (‘SunUp’, Nos. 2 and 4; ‘Kapoho’, Nos. 10, 11, and 14; and ‘Kapoho’, Nos. 12 and 13) were judged in an identical manner based on 186 DNA markers. Lines 2 and 5 formed a small cluster close to the two major clusters. Two ‘Sunset’ samples formed a fourth cluster that deviated from the second major cluster, to which the other five ‘Sunset’ samples belong. ‘Coimbatore 7’ was used as an out group and was shown to be clearly different from the Solo cultivars.

Cluster analysis of 71 papaya accessions and related species illustrated the genetic relationship among individual cultivars developed in different regions. Two samples of ‘AU1’ are presented here because non-Solo accession with multiple samples was not shown in Fig. 3. There are discrete clusters separating cultivars derived from the same or a simi-

**Table 1.** List of commercial papaya cultivars, breeding lines, unimproved germplasm, and related species analyzed by ALFP markers.

| (*a*) Commercial cultivars | | | |
|---|---|---|---|
| Accession | Samples | Source | Mating system |
| Kapoho | 16 | Hawaii | Gynodioecious |
| Rainbow | 9 | Hawaii | Gynodioecious |
| SunUp | 6 | Hawaii | Gynodioecious |
| Sunset | 8 | Hawaii | Gynodioecious |
| Sunrise | 2 | Hawaii | Gynodioecious |
| Waimanalo | 1 | Hawaii | Gynodioecious |
| Kaimya | 1 | Hawaii | Gynodioecious |
| Higgins | 1 | Hawaii | Gynodioecious |
| Eksotica I | 1 | MARDI, Malaysia | Gynodioecious |
| Khag Naun | 1 | Thailand | Gynodioecious |
| Coimbatore 7 | 1 | Coimbatore, India | Gynodioecious |
| AU16 | 1 | S.E. Queensland, Australia | Dioecious |
| AU25 | 1 | S.E. Queensland, Australia | Dioecious |
| AU28 | 1 | S.E. Queensland, Australia | Dioecious |
| AU35 | 1 | S.E. Queensland, Australia | Dioecious |
| AU70 | 1 | Central Qnslnd., Australia | Dioecious |
| AU97 | 1 | N. Queensland, Australia | Dioecious |
| AU2001 | 1 | S.E. Queensland, Australia | Dioecious |
| (*b*) Breeding lines (improved, not released) | | | |
| Accession | Samples | Source | Mating system |
| Line 2 | 1 | Hawaii | Gynodioecious |
| Line 5 | 1 | Hawaii | Gynodioecious |
| Line 8 | 1 | Hawaii | Gynodioecious |
| Line 40 | 1 | Hawaii | Gynodioecious |
| Line 116 | 1 | Hawaii | Gynodioecious |
| 4-16-35 | 1 | Hawaii | Gynodioecious |
| Maunawili Sweet | 1 | Maunawili, Oahu, Hawaii | Gynodioecious |
| Wan Peng | 1 | MARDI, Malaysia | Gynodioecious |
| KL Yellow | 1 | Kuala Lumpur, Malaysia | Gynodioecious |
| Saipan Red | 1 | Saipan, Micronesia | Gynodioecious |
| AU1 | 2 | S.E. Queensland, Australia | Dioecious |
| AU9 | 1 | Israel | Dioecious |
| UH 356-3 | 1 | Colombia, via Florida U.S.A. | Dioecious |
| 9-1-D | 1 | Coimbatore, India | Dioecious |
| Washington | 1 | Coimbatore, India | Dioecious |
| (*c*) Unimproved germplasm | | | |
| Accession | Samples | Source | Mating system |
| UH 370 | 1 | Oahu, Hawaii | Gynodioecious |
| Waikane | 1 | Waikane, Oahu, Hawaii | Gynodioecious |
| HAES 7836 | 1 | Malaysia | Gynodioecious |
| Mutant Yellow | 1 | Malaysia | Gynodioecious |
| UH 721 | 1 | Thailand | |
| UH 728 | 1 | Mexico | |
| UH 775 | 1 | P. Belem, Peru | |
| UH 778 | 1 | Tamshiyaku, Peru | |
| UH 892 | 1 | San Sebastian, Guatemala | |
| UH 897 | 1 | San Sebastian, Guatemala | |
| UH 899 | 1 | La Lima, Honduras | |
| UH 903 | 1 | San Pedro Sula, Honduras | Gynodioecious |
| UH 905 | 1 | Motrique, Honduras | |
| UH 918 | 1 | Orotina, Costa Rica | Gynodioecious |
| UH 926 | 1 | Rio El General, Costa Rica | |
| UH 927 | 1 | Rio El General, Costa Rica | |
| UH 928 | 1 | Rio El General, Costa Rica | |
| UH 929 | 1 | El Brujo, Costa Rica | Gynodioecious |
| N-90-66 | 1 | Central America | Gynodioecious |

**Table 1** (*concluded*).

| Accession | Samples | Source | Mating system |
|---|---|---|---|
| N-90-98 | 1 | Guadeloupe | |
| N-90-99 | 1 | Guadeloupe | |
| N91-106 | 1 | Thailand | Polygamous |
| N91-107 | 1 | Thailand | Polygamous |
| N91-108 | 1 | Thailand | Polygamous |
| N91-7 | 1 | Thailand | Polygamous |
| Park Chong 1 | 1 | Thailand | Polygamous |
| Yuen Nong No. 1 | 1 | China | Dioecious |
| Brash Panama | 1 | Panama | Gynodioecious |
| Mexican No. 2 | 1 | Mexico | Dioecious |
| BDX 584 | 1 | Puerto Rico | |
| (*d*) *Carica* species- and inter-specific hybrids | | | |
| Accession | Samples | Source | Mating system |
| Hybrid 1* | 1 | South America | Dioecious |
| Hybrid 2** | 1 | South America | Dioecious |
| *C. goudatiana* | 1 | Columbia | Dioecious |
| *C. horovitziana* | 1 | Ecuador | Dioecious |
| *C. stipulata* | 1 | Ecuador | Dioecious |
| *C. parviflora* | 1 | South America | Dioecious |
| *C. pubescens* | 1 | South America | Dioecious |
| *C. monoica* | 2 | South America | Monoecious |

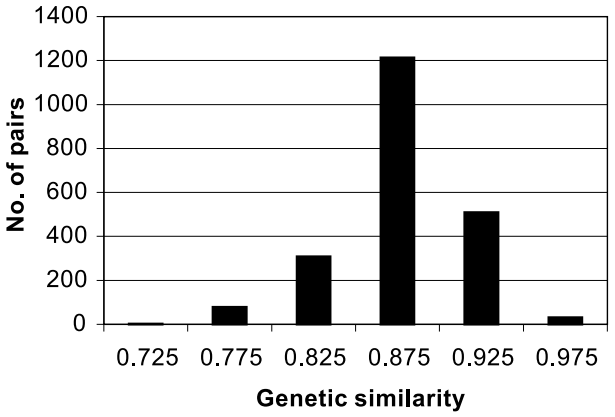
*(*C. parviflora* × *C. goudatiana*) × *C. goudatiana*
***C. parviflora* × *C. goudatiana***

**Table 2.** List of AFLP primers used in DNA fingerprinting of papaya.

| Code | Primer combination | Polymorphic Bands |
|---|---|---|
| E01M01 | E-AAC, M-CAA | 13 |
| E01M08 | E-AAC, M-CTT | 36 |
| E03M01 | E-ACA, M-CAA | 11 |
| E03M04 | E-ACA, M-CAT | 27 |
| E03M22 | E-ACA, M-CC | 20 |
| E05M07 | E-ACG, M-CTG | 15 |
| E05M08 | E-ACG, M-CTT | 22 |
| E06M04 | E-ACT, M-CAT | 26 |
| E06M07 | E-ACT, M-CTG | 16 |
| Total | | 186 |

**Note:** E, *Eco*RI; M, *Mse*I.

**Fig. 1.** Distribution of pair-wise comparison of genetic similarity among papaya cultivars.

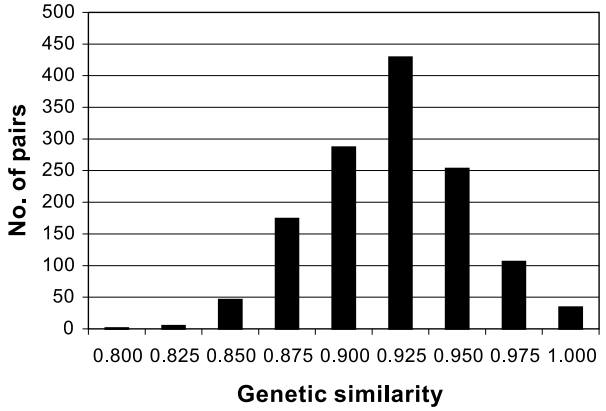



lar gene pool. The first cluster included all 15 Solo cultivars and breeding lines, plus 'Eksotica I', which was derived from a cross between Solo 'Sunrise' and a Malaysian variety, 'Subang-6', and then backcrossed to 'Sunrise' (Chan 1987) (Fig. 4). The second cluster included dioecious Australian cultivars and Indian cultivars that grow in subtropical or temperate regions. A group of cultivars originating from different countries formed the third cluster. The materials from Central America grouped into four different clusters indicating greater genetic diversity in the area where papaya originated. 'Brash Panama' seems to share the least genetic similarity with the rest of the papaya cultivars and accessions investigated.

The differences between papaya and other *Carica* species were substantial (Table 5). The average genetic similarity among papaya and the other six *Carica* species was 0.432, ranging from 0.296 to 0.545, whereas the average genetic similarity among those six *Carica* species was 0.729, ranging from 0.634 to 0.870. The most closely related *Carica* species were *C. pubescence* and *C. stipulata* with a genetic similarity of 0.870. The most distantly related species were *C. papaya* and *C. goudatiana*, showing a genetic similarity of only 0.364.

The six other *Carica* species and two interspecific hybrids formed a distinctive cluster, including seven samples from six *Carica* species and two interspecific hybrids derived from *C. parviflora* and *C. goudatiana*. Two samples from *C. monoica* shared a genetic similarity of 0.94. Two subclusters were formed among these species and hybrids. One included two hybrids, one of their parents (*C. goudatiana*), and two samples of *C. monoica*. The other included *C. parviflora*, *C. pubescence*, *C. stipulata*, and *C. horovitziana*. *Carica pubes-*

**Fig. 2.** Distribution of pair-wise comparison of genetic similarity among Hawaiian Solo papaya cultivars.




*cence* and *C. stipulata* shared the highest genetic similarity among the seven *Carica* species fingerprinted. *Carica papaya* appeared to be significantly different from other *Carica* species.

## Discussion

The impact of inbreeding on narrowing the germplasm diversity was evaluated by analyzing the genetic similarity of Hawaiian Solo cultivars that have undergone more than 25 generations of inbreeding (Storey 1969). The theoretical value of heterozygosity should be nearly zero after this many generations of self-pollination. However, our data showed that 3–6% of the 186 AFLP markers scored in this survey were still polymorphic within each established Solo cultivar. This amount of variation within each cultivar or inbred line might best be explained by occasional outcrossing. Natural outcrossing in hermaphrodite papaya fields immediately adjacent to a potential source of contaminating pollen may affect about 5% of seed, as demonstrated using the β-glucuronidase (GUS) transgene as a marker in assays of seeds from non-transgenic papaya trees surrounding transgenic plants (R.M. Manshardt and M.M.M. Fitch, unpublished data).

The Solo cultivars clustered into two major groups that reflected their origins in different breeding or selection programs. Both clusters probably have 'Line 5' in common, but the 'Kapoho' cluster includes infusions from 'Line 8' and 'Betty' (from Florida), although probably not in 'Kapoho' itself. 'Kapoho' was selected in the 1940s from Solo varieties similar to 'Line 5' by a private grower, whereas the other varieties in the cluster are largely the effort of Dr. Henry Nakasone, former horticulturist at the University of Hawaii at Manoa, who was involved in papaya breeding from the 1950s to the 1980s. The 'Sunrise' cluster is derived from a cross of 'Kariya' (probably a grower selection of 'Line 5'), and red-fleshed 'Line 9', produced in the 1960s by Dr. Richard Hamilton, also a horticulturist at the University of Hawaii during the same period. There is nothing surprising about the degree of variation among these different Solo cultivars, as they have quite distinct and documented pedigrees.

**Table 3.** Average genetic similarity among Hawaiian papaya varieties.

| | Kapoho | SunUp | Sunset | Sunrise | Rainbow | Higgins | Kamiya | Line 2 | Line 5 | Line 8 | Line 40 | Line 116 | Waimanalo | 4-16-35 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| SunUp | 0.888 | | | | | | | | | | | | | |
| Sunset | 0.883 | 0.948 | | | | | | | | | | | | |
| Sunrise | 0.895 | 0.967 | 0.943 | | | | | | | | | | | |
| Rainbow | 0.921 | 0.928 | 0.918 | 0.927 | | | | | | | | | | |
| Higgins | 0.936 | 0.899 | 0.895 | 0.895 | 0.915 | | | | | | | | | |
| Kamiya | 0.936 | 0.922 | 0.917 | 0.927 | 0.928 | 0.941 | | | | | | | | |
| Line 2 | 0.882 | 0.864 | 0.859 | 0.863 | 0.886 | 0.923 | 0.918 | | | | | | | |
| Line 5 | 0.910 | 0.883 | 0.878 | 0.881 | 0.915 | 0.944 | 0.927 | 0.949 | | | | | | |
| Line 8 | 0.930 | 0.921 | 0.915 | 0.911 | 0.925 | 0.941 | 0.978 | 0.907 | 0.939 | | | | | |
| Line 40 | 0.932 | 0.907 | 0.903 | 0.897 | 0.923 | 0.950 | 0.955 | 0.949 | 0.961 | 0.967 | | | | |
| Line 116 | 0.948 | 0.883 | 0.880 | 0.874 | 0.923 | 0.957 | 0.930 | 0.902 | 0.933 | 0.952 | 0.961 | | | |
| Waimanalo | 0.925 | 0.912 | 0.906 | 0.906 | 0.922 | 0.941 | 0.978 | 0.907 | 0.922 | 0.978 | 0.950 | 0.941 | | |
| 4-16-35 | 0.921 | 0.933 | 0.926 | 0.927 | 0.938 | 0.940 | 0.946 | 0.885 | 0.905 | 0.946 | 0.922 | 0.930 | 0.946 | |
| Moanawili Sweet | 0.926 | 0.925 | 0.922 | 0.918 | 0.930 | 0.934 | 0.978 | 0.906 | 0.932 | 0.978 | 0.955 | 0.940 | 0.978 | 0.945 |

**Table 4.** Average genetic similarity among Australian papaya varieties.

| | AU16 | AU25 | AU28 | AU35 | AU2001 | AU70 | AU97 | AU9 | AU1-1 |
|---|---|---|---|---|---|---|---|---|---|
| AU25 | 0.914 | | | | | | | | |
| AU28 | 0.898 | 0.962 | | | | | | | |
| AU35 | 0.892 | 0.839 | 0.855 | | | | | | |
| AU2001 | 0.935 | 0.914 | 0.919 | 0.882 | | | | | |
| AU70 | 0.924 | 0.892 | 0.897 | 0.914 | 0.924 | | | | |
| AU97 | 0.919 | 0.919 | 0.925 | 0.876 | 0.930 | 0.941 | | | |
| AU9 | 0.870 | 0.859 | 0.876 | 0.876 | 0.881 | 0.880 | 0.930 | | |
| AU1-1 | 0.935 | 0.892 | 0.898 | 0.914 | 0.957 | 0.930 | 0.930 | 0.892 | |
| AU1-2 | 0.924 | 0.881 | 0.886 | 0.930 | 0.946 | 0.908 | 0.908 | 0.886 | 0.978 |

**Fig. 3.** Phenogram based on simple matching coefficient of similarity among Hawaiian Solo papaya cultivars, breeding lines, and selected *Carica* species. An Indian cultivar, 'Coimbatore 7', was used as out group for comparison. Cophenetic correlation coefficient = 0.99.

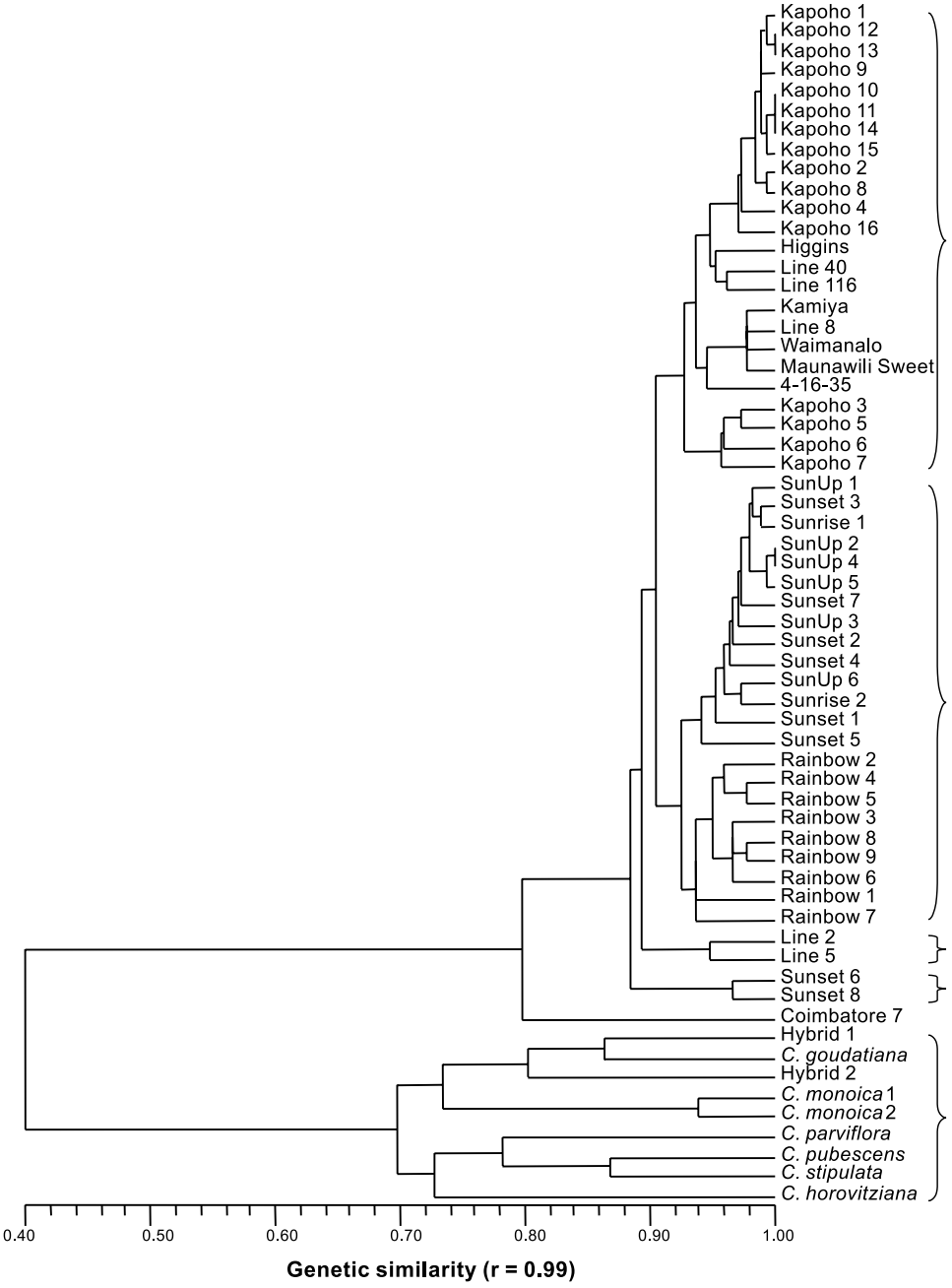
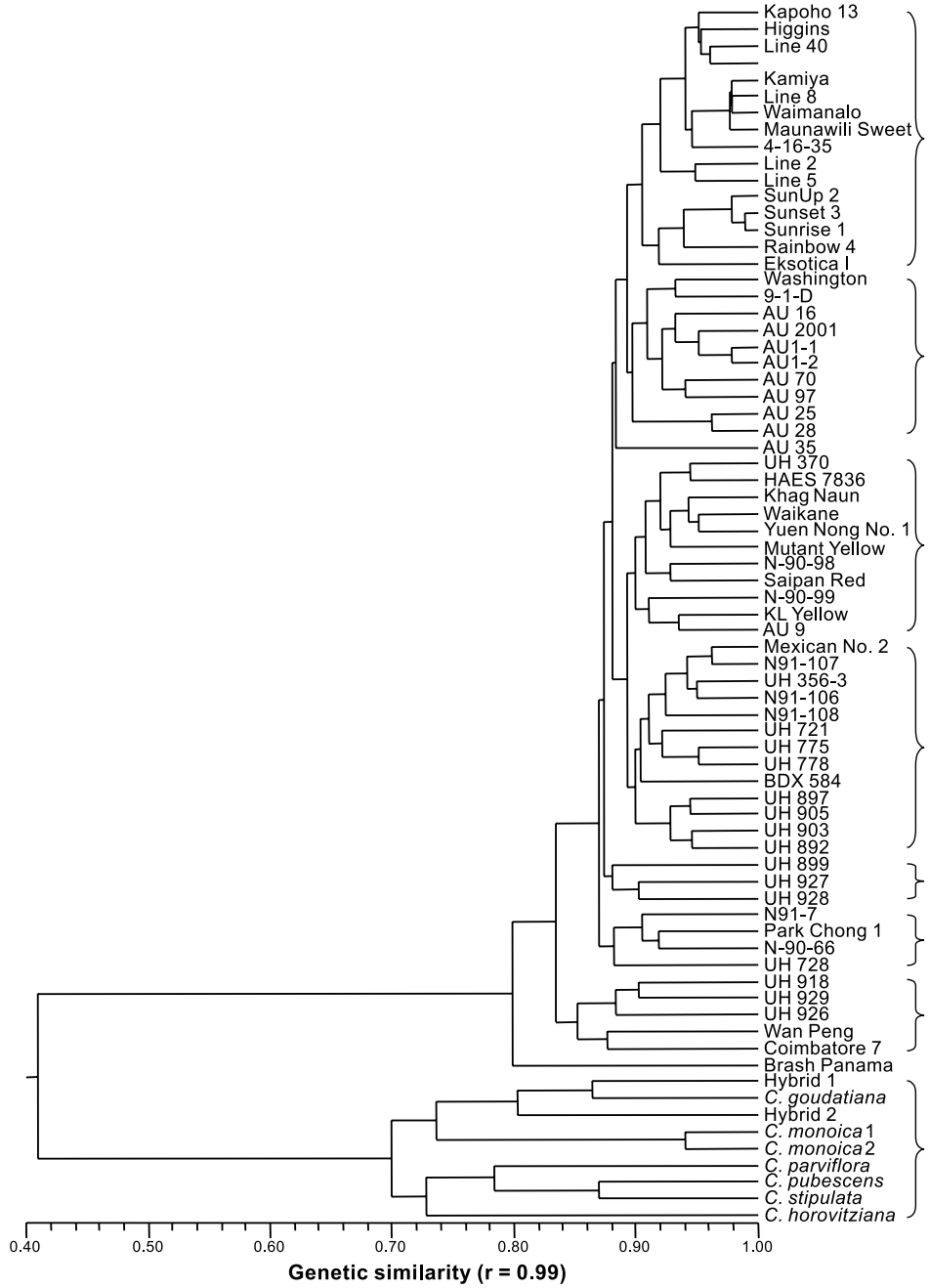


**Fig. 4.** Phenogram based on simple matching coefficient of similarity among 71 papaya accessions and related species. Cophenetic correlation coefficient = 0.99.




In the first cluster, eight ‘Kapoho’ samples, Nos. 9–16 collected from eight different farms on the island of Hawaii, formed a tight subcluster with average genetic similarity at 0.988. This result contradicts the idea that the ‘Kapoho’ cultivars grown on different farms might have diverged from each other through inbreeding, because each farmer selects and keeps his own seeds. One possible explanation for the small variation is that the original ‘Kapoho’ was highly homozygous, so that the subsequent inbreeding may not significantly change its genetic composition. However, four samples collected from Oahu deviate from most ‘Kapoho’ samples, suggesting that greater variation may exist within this variety than indicated by the samples collected from the island of Hawaii. Four cultivars in this cluster, ‘Kamiya’, ‘Line 8’, ‘Waimanalo’, and ‘Maunawili Sweet’, shared an equal genetic similarity of 0.978. It is known that ‘Line 8’ was one of the parental lines of ‘Waimanalo’, and that ‘Kamiya’ was selected from Waimanalo, Oahu (Nakasone et al. 1972). ‘Maunawili Sweet’, selected by a farmer in Maunawili, is possibly derived from one of these three cultivars. The second cluster included ‘Sunrise’, ‘SunUp’, ‘Sunset’, and ‘Rainbow’. The following four cultivars are interrelated: ‘Sunrise’ and ‘Sunset’ are sister lines; ‘Sunset’ and ‘SunUp’ are virtually the same except the *papaya ringspot virus* (PRSV) coat protein gene was added to the transgenic cultivar ‘SunUp’ (Fitch et al. 1992); ‘Rainbow’ is

**Table 5.** Average genetic similarity among *Carica* species.

| | *C. papaya* | *C. goudatiana* | *C. monoica* | *C. parviflora* | *C. pubescens* | *C. stipulata* |
|---|---|---|---|---|---|---|
| *C. goudatiana* | 0.364 | | | | | |
| *C. monoica* | 0.439 | 0.768 | | | | |
| *C. parviflora* | 0.413 | 0.634 | 0.679 | | | |
| *C. pubescens* | 0.446 | 0.714 | 0.748 | 0.773 | | |
| *C. stipulata* | 0.430 | 0.654 | 0.715 | 0.795 | 0.870 | |
| *C. horovitziana* | 0.474 | 0.686 | 0.710 | 0.730 | 0.745 | 0.708 |

a hybrid derived from the cross between 'Kapoho' and 'SunUp'.

The collection of papaya accessions used in this study differs dramatically in morphology and adaptation to different climates in tropical, subtropical, and temperate regions. The limited genetic variation (about 12%) detected among this diverse group of materials did not correspond to the wide range of morphological characteristics observed in the field. A similar low level of genetic diversity was recently documented in another self-pollinated plant species, *Coffea arabica* (Steiger et al. 2002). However, it is known that one or a few genes can sometimes significantly change plant stature (Sheridan 1988), leaf and (or) stem color (Dooner and Kermicle 1971), fruit size (Grandillo et al. 1999; Frary et al. 2000), and fruit quality (Carvalho et al. 1965; Paterson et al. 1991; Gray et al. 1999). Most of the crucial traits mentioned in the introduction are sufficiently variable, and improvement consists of selecting desirable recombinants from segregating populations. Some disease and (or) insect resistance, postharvest storage, and sex-modification (true-breeding hermaphrodite) factors are not known to exist in papaya germplasm. Our AFLP survey won't identify these, but could help point to related materials if something interesting is found in standard evaluation trials. Interspecific hybrids have been created, but are of little help owing to their high sterility (Drew et al. 1998). Genetic transformation is the only attractive route to some of these breeding objectives.

The results of genetic diversity study provide estimates on the level of genetic variation among diverse materials that can be used in germplasm management, varietal protection, and papaya improvement. Among the 31 diverse accessions collected from the USDA clonal germplasm repository, the average genetic similarity was 0.886, nearly the same as the overall average of 0.880 detected with 63 papaya accessions. These 31 accessions counted for 10% of the entire collection of the papaya germplasm at this repository, suggesting that the current collection preserved the vast majority of the natural variation in papaya. Although identical samples were found within 'Kapoho' and 'SunUp', all 63 papaya accessions were distinctively separated by 186 AFLP markers (Fig. 4). The amount of genetic diversity found with AFLP markers is sufficient to distinguish between breeding lines for varietal protection. The estimates of genetic similarity are particularly useful in choosing widely divergent parents with desirable traits for genetic mapping and selection. A genetic map of papaya has been constructed from an $F_2$ population derived from two Solo cultivars, 'Kapoho' and 'SunUp', with 507 AFLP markers (Ma et al. 2000). A quantitative trait loci (QTL) mapping project is underway to identify the loci controlling leafhopper tolerance in papaya using the same mapping population.

The genetic relationships of six other *Carica* species are much closer to each other than to *C. papaya*, thus confirming the conclusions of previous reports (Jobin-Décor et al. 1997; Aradhya et al. 1999; Van Droogenbroeck et al. 2002). The closest related *Carica* species were *C. pubescence* and *C. stipulata* as reported by Sharon et al. (1992), Jobin-Decor et al. (1997), and Van Droogenbroeck et al. (2002). Based on the phylogeny of *Carica* species derived from chloroplast DNA sequence (Aradhya et al. 1999), Badillo (2000) suggested a classification of *Carica papaya* as a separate genus, with only one species, giving a total of six genera in the family of Caricaceae (R.M. Manshardt, personal communication). Our DNA fingerprinting results support this concept.

## Acknowledgements

We thank Susan Ancheta and Ann Parisi for technical assistance; Emerson Llantero and the Papaya Administrative Committee for help collecting the papaya samples from eight papaya farms; John Stiles for papaya fruit cDNA probes; and Henrik Albert, Blake Vance, Robert Osgood, Jody Moore, and Stephanie Whalen, for reviewing the manuscript. This work was supported by a USDA–ARS Agreement for Minor Crops Research via the University of Hawaii and USDA–ARS Cooperative Agreements (Nos. CA 58-5320-3-460 and CA 58-5320-9-105) with the Hawaii Agriculture Research Center.

## References

Arumuganathan, K., and Earle, E.D. 1991. Nuclear DNA content of some important plant species. Plant Mol. Biol. Rep. **9**: 208–218.

Aradhya, M.K., Manshardt, R.M., Zee, F., and Morden, C.W. 1999. A phylogentic analysis of the genus *Carica* L. (Caricaceae) based on restriction fragment length variation in a cpDNA intergenic spacer region. Genet. Resour. Crop Evol. **46**: 579–586.

Badillo, V. 1993. Caricaecae, segundo esquema. Publicada por la asociacion de profesores, Alcance 43, Universidad Central de Venezuela, Maracay. pp. 111.

Badillo, V.M. 2000. *Carica* L. vs. *Vasconcellea* St. Hil. (Caricaceae) con la rehabilitacion de este ultim. Ernstia, **10**: 74–79.

Bremer, K., Chase, M.W., and Stevens, P.F. 1998. An ordinal classification for the families of flowering plants. Ann. Mo. Bot. Gard. **85**: 531–553.

Breyne, P., Rombaut, D., Van Gysel, A., Van Montagu, M., and Greats, T. 1999. AFLP analysis of genetic diversity within and between *Arabidopsis thaliana* ecotypes. Mol. Gen. Genet. **261**: 627–634.

Candolle, A de. 1908. Origin of cultivated plants. D.A. Appleton & Co. New York, N.Y.

Carvalho, A., Tango, J.S., and Monaco, L.C. 1965. Genetic control of the caffeine content of coffee. Nature (London), **205**: 314

Chan, Y.K. 1987. Backcross method in improvement of papaya (*Carica papaya* L.). Malays. Appl. Biol. Rep. **1**: 19–21.

Chittenden, L.M., Schertz, K.F., Lin, Y.R., Wing, R.A., and Paterson, A.H. 1994. A detailed RFLP map of *Sorghum bicolor* × *S. propinquum*, suitable for high-density mapping, suggests ancestral duplication of *Sorghum* chromosomes or chromosomal segments. Theor. Appl. Genet. **87**: 925–933.

Dooner, H.K., and Kermicle, J.L. 1971. Structure of the $R^r$ tandem duplication in maize. Genetics, **67**: 427–436.

Drew, R.A., O'Brien, C.M., and Magdalita, P.M. 1998. Development of *Carica* interspecific hybrids. Acta Hortic. **461**: 285–289.

Ellis, R.P., McNicol, J.W., Baird, E., Booth, A., Lawrence, P., Thomas, B., and Powell, W. 1997. The use of AFLPs to examine genetic relatedness in barley. Mol. Breed. **3**: 359–369.

Erschadi, S., Haberer, G., Schoniger, M., and Torres-Ruiz, R.A. 2000. Estimating genetic diversity of *Arabidopsis thaliana* ecotypes with amplified fragment length polymorphism (AFLP). Theor. Appl. Genet. **100**: 633–640.

Fitch, M.M.M., Manshardt, R.M., Gonsalves, D., and Slightom, J.L. 1992. Virus resistance papaya derived from tissue bombarded with the coat protein gene of papaya ringspot virus. BioTechnol. **10**: 1466–1472.

Frary, A., Nesbitt, T.C., Frary, A., Grandillo, S., van der Knaap, E., Cong, B., Liu, J., Meller, J., Elber, R., Alpert, K.B., and Tanksley, S.D. 2000. fw 2.2: a quantitative trait locus key to the evolution of tomato fruit size. Science (Washington D.C.), **289**: 85–88.

Grandillo, S., Ku, H.M., and Tanksley, S.D. 1999. Identifying the loci responsible for natural variation in fruit size and shape in tomato. Theor. Appl. Genet. **99**: 978–987.

Gray, D.A., Prestage, S., Linforth, R.S.T., and Taylor, A.J. 1999. Fresh tomato specific fluctuations in the composition of lipoxygenase-generated C6 aldehydes. Food Chem. **64**: 149–155.

Jobin-Décor, M.P., Graham, G.C., Henry, R.J., and Drew, R.A. 1997. RAPD and isozyme analysis of genetic relationship between *Carica papaya* and wild relatives. Genet. Resour. Crop Evol. **44**: 471–477.

Ma, H., Moore, P.H., Liu, Z., Kim, M.S., Yu, Q., Fitch., M.M.M., Paterson, A.H., Ming, R. 2002. Linkage mapping revealed suppression of recombination at the sex determination locus in papaya. Proceedings of the 10th International Plant and Animal Genome X Conference, 12–16 January 2002, San Diego, Calif.

Maughan, P.J., Saghai Maroof, M.A., Buss, G.R., and Huestis, G.M. 1996. Amplified fragment polymorphism (AFLP) in soybean: species diversity, inheritance, and near-isogenic line analysis. Theor. Appl. Genet. **93**: 392–401.

Nakasone, H.Y., Crozier, J.A. Jr, and Ikehara, D.K. 1972. Evaluation of 'Waimanalo', a new papaya strain. Tech. Bull. Hawaii Agr. Exp. Stn. **79**: 3–12.

Paterson, A.H., Damon, S., Hewitt, J.D., Hewitt, J.D., Zamir, D., Rabinowitch, H.D., Lincoln, S.E., Lander, E.S., and Tanksley, S.D. 1991. Mendelian factors underlying quantitative traits in tomato: comparison across species, generations, and environments. Genetics, **127**: 181–197.

Pejic, I., Ajmone-Marsan, P., Morgante, M., Kozumplick, V., Castiglioni, P., Taramino, G., and Motto, M. 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theor. Appl. Genet. **97**: 1248–1255.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S., and Rafalski, A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Mol. Breed. **2**: 225–238.

Purseglove, J.W. 1968. Tropical crops. Longman, London, pp. 45–51

Rodman, J.E., Karol, K.G., Price, R.A., and Sytsma, K.J. 1996. Molecules, morphology, and Dahlgren's expanded order Capparales. Sys. Bot. **21**: 289–307

Russell, J.R., Fuller, J.D., Macaulay, M., Hatz, B.G., Jahoor, A., Powell, W., and Waugh, R. 1997. Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theor. Appl. Genet. **95**: 714–722.

Sharon, D., Hillel, J., Vainstein, A., and Lavi, U. 1992. Application of DNA fingerprints for identification and genetic analysis of *Carica* papaya and other *Carica* species. Euphitica, **62**: 119–126.

Sheridan, W.F. 1988. Maize developmental genetics: genes of morphogenesis. Annu. Rev. Genet. **22**: 353–385.

Sneath, P.H.A., and Sokal, R.R. 1973. Numerical taxonomy. Freeman, San Francisco, Calif. pp. 573.

Sokal, R.R., and Michener, C.D. 1958. A statistical method for evaluating systematic relationships. Univ. Kans. Sci. Bull. **38**: 1409–1438.

Steiger, D.L., Nagai, C., Moore, P.H., Morden, C.W., Osgood, R.V., and Ming, R. 2002. AFLP analysis of genetic diversity within and among *Coffea arabica* cultivars. In press.

Stiles, J.I., Lemme, C., Sondur, S., Morshidi, M.B., and Manshardt, R.M. 1993. Using randomly amplified polymorphic DNA for evaluating genetic relationships among papaya cultivars. Theor. Appl. Genet. **85**: 697–701.

Storey, W.B. 1969. Papaya. *In* Outlines of perennial crop breeding in the tropics. *Edited by* F.P. Ferwerda and F. Wit. H. Veenman & Zonen N.V., Wageningen, The Netherlands. pp. 389–408.

Storey, W.B. 1976. Papaya. *In* Evolution of crop plants. *Edited by* N.W. Simmonds. Longman, London. pp. 21–24.

Van Droogenbroeck, B., Breyne, P., Goetghebeur, P., Romeijn-Peeters, E., Kyndt, T., Gheysen, G. 2002. AFLP analysis of genetic relationships among papaya and its wild relatives (Caricaceae) from Ecuador. Theor. Appl. Genet. In press.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. 1995. AFLP: a new technique for DNA finger printing. Nucleic Acids Res. **23**: 4407–4414.