

# Inheritance of AFLP Markers and Their Use for Genetic Diversity Analysis in Wild and Domesticated Pawpaw [*Asimina triloba* (L.) Dunal]

Ying Wang

Department of Plant Breeding, Cornell University, Ithaca, NY 14853

Gregory L. Reighard and Desmond R. Layne<sup>1</sup>

Department of Horticulture, Clemson University, Clemson, SC 29634

Albert G. Abbott

Department of Genetics, Biochemistry and Life Science Studies, Clemson University, Clemson, SC 29634

Hongwen Huang<sup>2</sup>

Wuhan Institute of Botany, The Chinese Academy of Sciences, Wuhan, Hubei 430074, P.R. China

ADDITIONAL INDEX WORDS. genetic resources, biodiversity, wild populations, marker clustering, genetic linkage groups

**ABSTRACT.** Pawpaw (*Asimina triloba*) produces the largest fruit native to the United States. Six linkage groups were identified for *A. triloba* using the interspecific cross [PPF1-5 (*A. triloba*) × RET (*A. reticulata* Shuttlw. ex Chapman)], covering 206 centimorgans (cM). A total of 134 dominant amplification fragment length polymorphism (AFLP) markers (37 polymorphic and 97 monomorphic) were employed for estimating the genetic diversity of eight wild populations and 31 cultivars and advanced selections. For the wild populations, the percentage of polymorphic loci over all populations was 28.1% for dominant markers and Nei's genetic diversity ( $H_e$ ) were 0.077 estimated by 134 dominant markers. Genetic diversity and the percentage of polymorphic loci estimated using only polymorphic dominant AFLPs were 0.245 and 79%, respectively, which are comparable with other plant species having the same characteristics. Estimated genetic diversity within populations accounted for 81.3% of the total genetic diversity. For cultivars and advanced selections, genetic diversity estimated by 134 dominant markers was similar to that of wild pawpaw populations ( $H_e = 0.071$ ). Thirty-one cultivars and advanced selections were delineated by as few as nine polymorphic AFLP dominant loci. Genetic relationships among wild populations, cultivars and advanced selections were further examined by unweighted pair group method with arithmetic mean (UPGMA) of Nei's unbiased genetic distance. The genetic diversity estimated for wild populations using the clustered polymorphic markers was lower than the result estimated using the nonclustered polymorphic markers. Therefore, this study indicates that the number of sampled genomic regions, instead of the number of markers, plays an important role for the genetic diversity estimates.

The North American pawpaw produces the largest fruit by any tree native to the United States, but its fruit remains largely under-exploited. Pawpaw has long been appreciated for its oblong-shaped, green-skinned fruit with buttery, yellow flesh and unique tropical flavor that can be described as a blend of banana (*Musa xparadisica* L.), mango (*Mangifera indica* L.), and pineapple [*Ananas comosus* (L.) Merr.] with a mild, pleasant fruit aroma. Pawpaw fruit is an excellent food source that exceeds many common fruits in vitamins, minerals, amino acids, and food energy values (Jones and Layne, 1997). Pawpaw trees also produce natural defense compounds (annonaceous acetogenins) in leaf, bark, and twig tissues that have highly effective anti-tumor and pesticidal properties (McLaughlin, 1997; Oberlies et al., 1997a,

1997b). Harvesting the leaves and twigs for extraction of these compounds may also generate a new crop for small farmers in the future (Pomper et al., 1999). In addition, pawpaw is an ideal tree for the residential landscape due to its lush tropical appearance, attractive conical tree form, size, fall color, and delicious fruit (Layne, 1997). Therefore, pawpaw has great potential for commercial development as an alternative tree fruit crop in the United States and elsewhere (Jones and Layne, 1997).

Pawpaw is a diploid ( $2n = 2x = 18$ ) and has had limited study of the reproduction biology (Willson and Schemske, 1980). The species reproduces predominantly by asexual means through root suckering. The flowers are strongly protogynous, strictly outcrossing, and mainly insect-pollinated (Faegri and van der Pijl, 1971). Low fruit set has been observed in nature and it is primarily attributed to inefficient dispersal of pollen and/or a lack of pollinators (Lagrange and Tramer, 1985; Willson and Schemske, 1980). Seed dispersal and germination is also generally poor in the wild (Peterson, 1991). Given these characters, it is plausible that within-population genetic variation may be low. On the other hand, species whose populations are distributed over a wide geographic range, such as *A. triloba*, may maintain significant genetic variation between populations (Hamrick and Godt, 1989; Rogstad et al., 1991). Because pawpaw is an outcrosser, seed reproduction must have played an important role in extending the

Received for publication 19 July 2004. Accepted for publication 4 Dec. 2004. This material is based upon work supported by the CSREES/USDA, under project no. SC-4960 and the CAS-project KSCX2-SW-104. We thank Dr. T.L. Kubisiak and his group in USDA (Saucier, MS) for DNA extraction and purification.

<sup>1</sup>Former Principal Investigator of Horticulture and Curator, USDA National Clonal Germplasm Repository For *Asimina* spp., Kentucky State Univ., Frankfort, KY 40601.

<sup>2</sup>Corresponding Author. Professor of Plant Genetics and Breeding, Wuhan Botanical Garden/Wuhan Institute of Botany, The Chinese Academy of Sciences, Wuhan, Hubei 430074, P.R. China. E-mail: hongwen@wbcas.cn; phone: 86-27-87510232; Fax: 86-27-887510331.

species range and in adapting to environmental-ecological changes through genetic recombination (Peterson, 1991). Another source of genetic variation in *A. triloba* may occur by crossing with other tropical *Asimina* Adans. species near the southern extremities of its distribution range (Norman et al., 1992).

Efforts toward the domestication of pawpaw began early in the 20th century (Peterson, 1991). Considerable progress has been made in selecting and breeding commercial cultivars for orchard production (Peterson, 1991, 2003), resulting in a wide array of advanced selections and cultivars available for commercialization. Currently, 46 cultivars and selections are commercially available in different nurseries [Jones et al., 1998; Kentucky State Univ. (KSU), 2004]. Recent efforts at germplasm collection and breeding of pawpaw, led by KSU and the PawPaw Foundation (PPF), have framed current research strategies. This includes germplasm collection, variety trials, cultural recommendations, clonal propagation, and breeding commercial cultivars (Layne, 1996). However, some concern to plant breeders and custodians of germplasm is the genetic variation found among cultivars and germplasm collections held at the National Clonal Germplasm Repository for *Asimina* species at KSU.

Pawpaw is at an early stage of domestication, therefore maintaining a high level of genetic diversity is essential for the long-term success of a breeding program and for sustaining a diverse source of useful genes in the germplasm repository. Several molecular-based analyses have been recently used to assess genetic diversity in pawpaw. These include a minisatellite probe (M13) (Rogstad et al., 1991), allozymes (Huang et al., 1997, 1998), rapid amplification of polymorphic DNA (RAPDs) (Huang et al., 2000, 2003), and intersimple sequence repeats (ISSR) (Pomper et al., 2003). These previous studies determined that the genetic variation in both cultivated and wild pawpaws is comparable to those of other long-lived, temperate woody perennials. These perennials are fully characterized by a widespread geographic range, insect-pollinated outcrossing breeding systems, secondary asexual reproduction, and animal-dispersed seed.

A comprehensive understanding of genetic diversity among and within pawpaw populations and molecular characterization of cultivated pawpaw cultivars or advanced selections using the same set of molecular markers are needed for formulating appropriate sampling and management strategies for the KSU *Asimina* spp. repository and for furnishing useful genetic information for future pawpaw breeding efforts. A more detailed analysis with a larger number of genetic markers would provide a better understanding of the genetic diversity that exists in both wild populations and germplasm collections, which will contribute to useful gene conservation for a long-term breeding program. AFLP markers are reliable, high throughput, cost-effective, and have wide genome coverage (Hansen et al., 1999; Prashanth et al., 2002). The objectives of the present study were to determine the inheritance of the AFLP markers in the interspecific crossed family and construct the genetic linkage groups for pawpaw, and then use the same set of markers to assess genetic diversity among eight wild populations and characterize 31 pawpaw cultivars and advanced selections.

## Materials and Methods

**PLANT MATERIALS AND DNA EXTRACTION.** Dormant stem cuttings from various pawpaw trees located at the USDA National Clonal Germplasm Repository for *Asimina* species in Frankfort,

Ky., were harvested in late Fall 1996 and early 1997. In addition, several pawpaw cultivars not conserved at the repository were collected from commercial nurseries in the United States. The cuttings included 34 progeny from an interspecific cross [PPF 1-5 (*A. triloba*) × RET (*A. reticulata*)], seedlings grown from open-pollinated seeds collected from one specific location in each of the eight states within the natural range of *A. triloba* (4 to 18 trees from each site in a certain state), and 31 pawpaw cultivars or advanced selections (Tables 1 and 2) representing almost the entire gene pool of currently available cultivated pawpaws (Peterson, 1991). All stem cuttings were placed in a greenhouse, and leaf tissue collected from newly emerged shoots was used for DNA extraction.

Total nucleic acids were isolated from  $\approx 2$  g of leaf tissue using a modification of the cetyltrimethylammonium bromide (CTAB) based procedure outlined in Wagner et al. (1987). Working solutions of genomic DNA at  $100 \text{ ng} \cdot \mu\text{L}^{-1}$  in TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0) were prepared for AFLP analyses.

**AFLP ANALYSIS AND GENETIC LINKAGE MAP CONSTRUCTION.** The AFLP analysis was performed essentially as described by Vos et al. (1995). Primer pairs for the selective amplification were first tested using *EcoRI* primers carrying two or three selective nucleotides (E+2 or M+3) and *MseI* primers carrying two or three selective nucleotides (M+2 or M+3). Because E+2 and M+2 primer combinations generated too many bands and it was hard to score the unambiguous markers, four E+3 and M+3 primer combinations were used for fingerprinting: EACGMCAC, EAAGMCAT, EACAMCCT, and EACTMCAA. An AFLP DNA ladder ranging from 30 to 330 base pairs (Invitrogen Corp., Carlsbad, Calif.) in length of 10-bp increments was used to determine the size of amplified DNA fragments.

Only intense and unambiguous bands were manually scored as present or absent. The AFLP markers were designated by *EcoRI/MseI* primer combination responsible for their amplification, followed by a three-digit number indicating the product size in base pairs. All polymorphic AFLP markers were evaluated by chi-square tests for goodness-of-fit against a 3:1, 1:2:1, and 1:1 ratio ( $P \leq 0.05$ ) in the interspecific cross family. Codominant markers were identified with 1:2:1 segregation ratio in the cross family. Dominant markers segregating with 1:1 ratio and contributed from *A. triloba* were used for developing genetic linkage map for *A. triloba*. Genetic linkage groups for *A. triloba* were constructed using MapMaker software (V2.0) for Macintosh (Lincoln et al., 1992), a 4.0 minimum LOD score, 0.25 maximum theta (recombination frequency), and the Kosambi mapping function.

**DATA ANALYSIS.** For the eight wild populations, both 37 polymorphic and 97 monomorphic markers were used for estimating a set of intra- and inter-population genetic statistics: the percentage of polymorphic loci (with 99% criterion), expected heterozygosity ( $H_e$ ; Nei, 1973), observed heterozygosity ( $H_o$ ), Shannon and Weaver's (1949) index (I), and F statistics. For the 31 cultivars and selections, Nei's (1973) gene diversity ( $H_g$ ) and Shannon and Weaver's (1949) index (I) were calculated for polymorphic markers. For both the wild and cultivated pawpaws, Nei's (1978) genetic distances and identities were calculated for all possible pairwise comparisons. Dendrograms for wild populations and domesticated pawpaw were constructed based on the matrix of the distances using UPGMA. All genetic parameters were calculated with assistance of POPGENE version 1.32 (Yeh and Boyle, 1997).

Table 1. Pawpaw cultivars and advanced selections used for genetic diversity analysis.<sup>z</sup>

Cultivar/selection	Breeding pedigree and origin	Selected year
1. PPF <sup>v</sup> 1-23	open-pollinated seedling of 'Taytwo'	1988–90
2. PPF 2-54	open-pollinated seedling of GAZ <sup>w</sup> -VA	1988–90
3. PPF 8-58	open-pollinated seedling of BEF <sup>w</sup> -30	1988–90
4. PPF 9-47	open-pollinated seedling of BEF-49	1988–90
5. PPF 9-58-1 <sup>v</sup>	open-pollinated seedling of BEF-50	1988–90
6. PPF 10-35	open-pollinated seedling of BEF-49	1988–90
7. PPF 11-5	open-pollinated seedling of BEF-53	1988–90
8. PPF 11-13	open-pollinated seedling of BEF-53	1988–90
9. Middletown	wild seedling from Middletown, Ohio	1951
10. Mitchell	wild seedling from Iuka, Ill.	1979
11. NC-1	'Davis' × 'Overleese'	1985
12. Overleese	cultivated seedling (open-pollinated) from Rushville, Ind.	1950
13. PA-Golden	wild seedling from eastern Pennsylvania	unknown
14. Sunflower	wild seedling from Chanute, KS	1970
15. Taylor	wild seedling from Eaton Rapids, Mich.	1968
16. Taytwo	wild seedling from Eaton Rapids, Mich.	1968
17. Wilson	wild seedling from Cumberland, Ky.	1985
18. PPF 1-68	open-pollinated seedling of 'Overleese'	1988–90
19. PPF 2-10	open-pollinated seedling of BEF-30	1988–90
20. PPF 3-11	open-pollinated seedling of BEF-33	1988–90
21. PPF 3-21	open-pollinated seedling of BEF-43	1988–90
22. PPF 4-2	open-pollinated seedling of BEF-53	1988–90
23. PPF 5-5	open-pollinated seedling of BEF-54	1988–90
24. PPF 7-90	open-pollinated seedling of RS <sup>u</sup> -2	1988–90
25. PPF 8-20	open-pollinated seedling of 'Sunflower'	1988–90
26. PPF 2-49	open-pollinated seedling of GAZ <sup>w</sup> -VA	1988–90
27. Rebecca's Gold	open-pollinated seed from Corwin Davis of Bellevue, Mich.	1974
28. PPF 9-58-2 <sup>v</sup>	open-pollinated seedling of BEF-50	1988–90
29. Prolific	cultivated seedling (open-pollinated) from Bellevue, Mich.	1985
30. SAA-Zimmerman	open-pollinated seedling of 'Zimmerman'	1982
31. Wells	cultivated seedling (open-pollination) from Salem, Ind.	1990

<sup>z</sup>More details about background information of the cultivars and selections can be found in Peterson (1991, 2003) and Layne (1997).

<sup>v</sup>PPF = Pawpaw Foundation orchards; numerous "wild" selections from the remnant collections of Allard, Blandy, Buckman, Hershey, Schlaanstine, and Zimmerman, plus some from truly wild trees and some from named cultivars; assembled by R.N. Peterson and H.J. Swartz at the experiment stations of the Univ. of Maryland, Keedysville, and Queenstown.

<sup>w</sup>GAZ = G.A. Zimmerman collection; containing most, if not all, of the named cultivars of the time, plus numerous wild selections and interspecific hybrids; assembled by G.A. Zimmerman of Linglestown, Pa., circa 1920 to 1940.

<sup>w</sup>BEF = Blandy Experimental Farm collection; numerous wild selections plus much of G.A. Zimmerman's collection, donated posthumously; assembled by O.E. White and staff of the Blandy Experimental Farm, Boyce, Va., from 1926 to 1955.

<sup>v</sup>PPF 9-58-1 and PPF 9-58-2 were provided by PPF and were thought to be half siblings from the original selection of PPF 9-58.

<sup>u</sup>RS = R. Schlaanstine collection; material descending from G.A. Zimmerman's collection via J. Hershey; assembled by R. Schlaanstine of West Chester, Pa., date uncertain (circa 1960?).

## Results and Discussion

### INHERITANCE OF AFLP MARKERS AND GENETIC LINKAGE ANALYSIS.

Four AFLP primer combinations, EACGMCAC, EAAGMCAT, EACAMCCT, and EACTMCAA, generated 138 readable bands with size ranging from 85 to 450 bp. Two codominant markers, EACGMCAC-213 and EAAGMCAT-220, were identified in the progeny of the interspecific cross and segregated as 1:2:1 confirmed by chi-square test. Sixty-six dominant AFLPs were polymorphic in the interspecific crossed family, with 54 of them segregating as 1:1 and 12 of them inherited from *A. triloba* segregating as 3:1 according to the chi-square test (data not shown). These 66

AFLPs were genetically confirmed and are very valuable for the following and future genetic diversity studies. Thirty-three of these 66 AFLPs were polymorphic in the wild populations and the cultivated pawpaws. EACTMCAA-170 and EACGMCAC-105 were monomorphic in the progeny, but were polymorphic in the wild populations and cultivars. EACGMCAC-122 and EACTMCAA-111 could not be amplified from any individuals of the progeny, but were polymorphic among the wild populations and cultivars and selections.

Fifty-four AFLP markers and 20 RAPD markers (Huang et al., 2000) segregating as 1:1 and contributed from *A. triloba* were used to construct the genetic linkage groups for *A. triloba*.

Table 2. Genetic diversity statistics of wild populations and domesticated *Asimina triloba* evaluated by dominant AFLP markers.

	Sample size (no.)	Polymorphic loci (%)	Gene diversity (H) <sup>z</sup>	Shannon's index (I)
<i>For pawpaw wild populations</i>				
Indiana	17	19.4	0.061	0.093
Maryland	5	12.2	0.048	0.071
Pennsylvania	4	8.6	0.039	0.055
New York	16	15.8	0.052	0.078
Illinois	5	18.0	0.067	0.100
Georgia	11	17.3	0.053	0.081
West Virginia	18	19.4	0.074	0.109
Virginia	5	9.4	0.039	0.058
Mean	10.1	15.0	0.054	0.081
Species total	80	28.1	0.077	0.120
<i>For domesticated pawpaw</i>				
Species total	31	28.0	0.071	0.112

<sup>z</sup>Nei (1973).

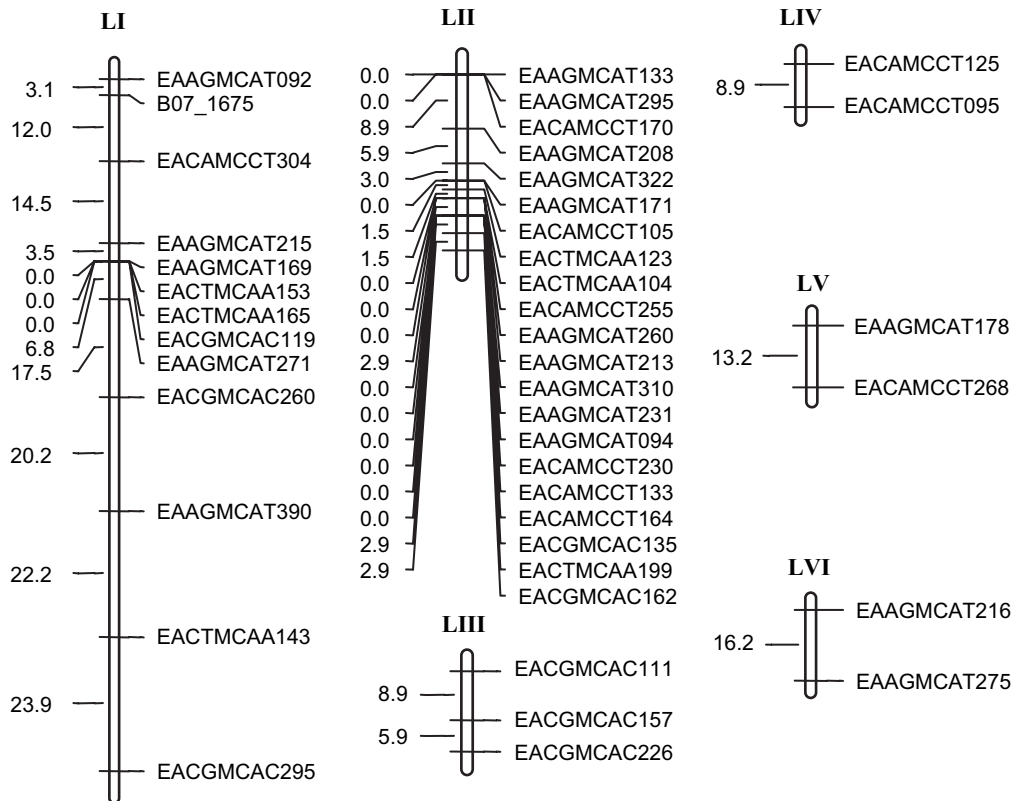


Fig. 1. Genetic linkage groups LI to LVI constructed for pawpaw (PPF 1-5, *A. triloba*) using the interspecific cross PPF 1-5 (*A. triloba*) × RET (*A. reticulata*). Genetic distance in centimorgans and loci were listed on the left and the right side, respectively. The designation of the AFLPs is based on the primer combination followed by three digits indicating the size of the band as base pairs.

Forty-three markers were not linked with any other marker. Six linkage groups including 42 AFLPs and one RAPD marker were constructed covering 206 cM (Fig. 1), with three linkage groups having more than three markers. As the pawpaw genome has nine chromosomes, these six linkage groups covered only part of the pawpaw genome. The other unlinked markers were either from the other chromosomes or from segments of the chromosomes that are not closely linked to these six linkage groups. In order to have better coverage of the whole genome, more markers need to be developed. These linkage groups could serve as a good

starting point of the future mapping project, which might greatly facilitate the pawpaw molecular breeding program.

Three AFLP clusters were observed on linkage groups I and II, with four, four, and seven markers co-segregating, respectively. AFLP clusters were also found in *Arabidopsis thaliana* (L.) Heynh. (Alonso-Blanco et al., 1998; Peters et al., 2001), *Zea mays* L. (Vuylsteke et al., 1999), *Petunia hybrida* (Hook.) Vilm. (Strommer et al., 2002), and *Coffea arabica* L. (Ky et al., 2000). Such clusters are known to be located at centromeric regions where heterochromatin locates and many repetitive frag-



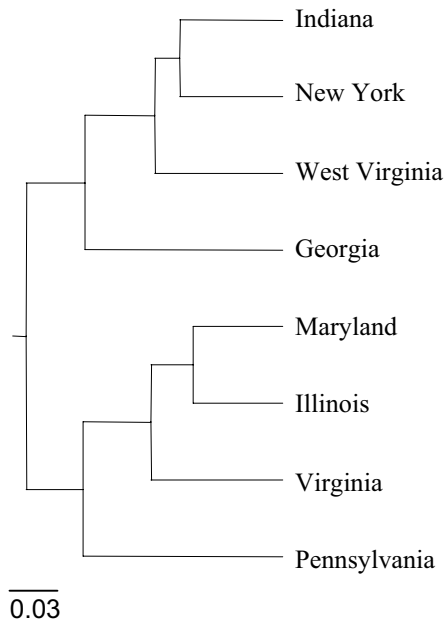


Fig. 2. The UPGMA dendrogram of Nei's genetic distances among eight wild pawpaw populations based on dominant AFLP markers.

ments are found. A total of 15 AFLPs (22.7%) were clustered in pawpaw genetic linkage groups. Although AFLP markers have broad distribution throughout the genome, this high percentage of clustered markers might affect the degree of whole genome coverage when a small number of AFLPs are employed for the population and conservation study.

**GENETIC DIVERSITY OF WILD POPULATIONS OF *A. TRILOBA*.** Ninety-seven monomorphic and 37 polymorphic dominant AFLPs were used to analyze the genetic diversity among the eight wild populations. The percentage of polymorphic loci over all populations was 28.1%, ranging from 8.6% to 19.4% in the individual population (Table 2). Nei's genetic diversity (Nei, 1973) across all the populations was 0.077 (ranging from 0.039 to 0.074) estimated by dominant markers (Table 2). The genetic diversity estimated using dominant AFLP markers among wild populations was lower than those estimated by RAPDs (0.249 for  $H_e$  and 64% for percentage of polymorphic loci) (Huang et al., 2000) and isozymes (0.172 for  $H_e$  and 43.5% for percentage of polymorphic loci) (Huang et al., 1998). The low level of  $H_e$  estimated using dominant AFLP markers was due to the large portion (72%) of monomorphic markers. The genetic diversity of 0.225 estimated using 37 polymorphic dominant AFLPs, with four AFLPs cosegregated with others and three tightly linked with others (<3 cM), were lower than 0.245 estimated using 30 non-clustered polymorphic AFLPs. Thus, the genetic diversity might be decreased due to the effect of clustered markers. This genetic diversity level is comparable to other outcrossing, long-lived, woody species: 0.27 for *Cedrela odorata* L., 0.1441 for *Prunus mahaleb* L., 0.192 for *Hippophae rhamnoides* L. (Bartish et al., 1999; Cavers et al., 2003; Jordano and Godoy, 2000).

Partitioning of the population genetic diversity showed that genetic diversity within populations accounted for 81.3% of the total genetic diversity based on dominant AFLPs. This result agrees with findings of Huang et al. (1998, 2000) using isozymes and RAPDs that most of the genetic diversity of pawpaw resides within the wild population. Furthermore, the proportion of 18.7% genetic diversity observed among *A. triloba* populations is slightly lower

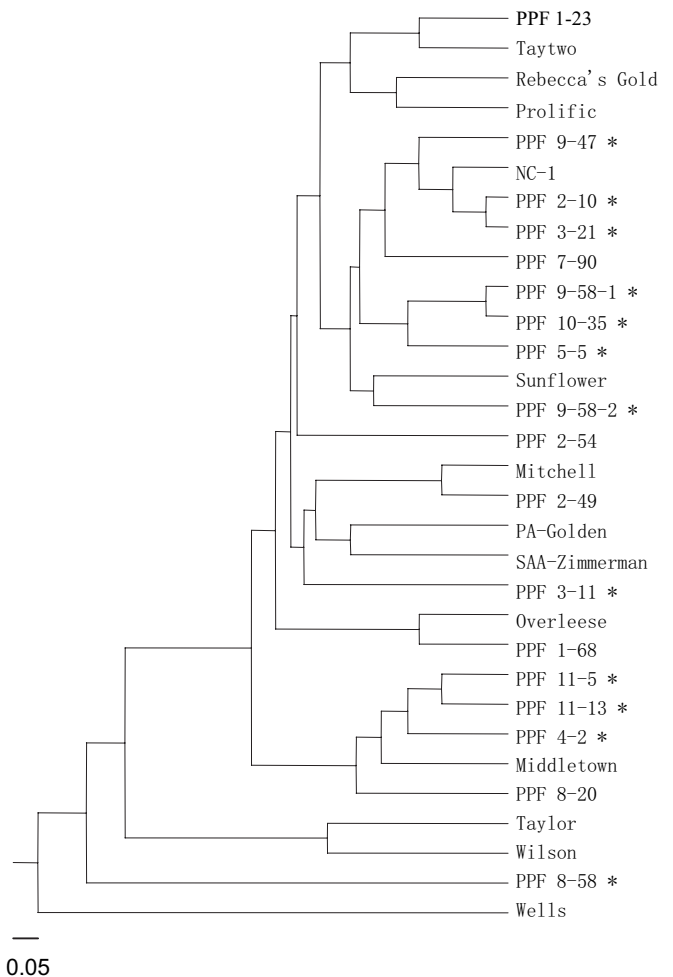


Fig. 3. The UPGMA dendrogram of Nei's genetic distances among 31 pawpaw cultivars and advanced selections based on dominant AFLP markers. \*Denotes the BEF (Blandy Experimental Farm) selections.

than the average for species with wide geographic range (21.0%), insect outcrossing species (24.3%), species with animal-ingested seed dispersal mechanisms (12.4% to 27.7%) and reproduction modes (21.3% to 22.5%), and temperate species (24.6%) (Hamrick and Godt, 1989). Gene flow rate was indirectly estimated as  $Nm = (1-F_{ST})/4F_{ST} = 2.17$  ( $F_{ST}$ , a measure of genetic differentiation over subpopulations) using the dominant AFLPs, which indicated a high level of genetic divergence among populations due to the lack of gene flow. *Asimina triloba*'s reproductive biology has demonstrated that selfed seeds are rare, seeds germinate poorly and the resulting seedlings do not compete well in growth with cross-pollinated seedlings (Faegri and van der Pijl, 1971; Willson and Schemske, 1980). It is also known that *A. triloba* has low efficiency of cross-pollination; insect pollinators may be scarce and poor fruit production often results in the wild (Norman and Clayton, 1986; Walker, 1971). Therefore, low efficiency of pollen and seed dispersal are likely the major factors contributing to such population differentiations in *A. triloba*, although it is continuously and widely distributed in the eastern United States.

Nei's (1978) unbiased genetic identity and distance were estimated for 28 pairwise population comparisons using 134 dominant AFLPs. Genetic relationships among wild populations were further examined by UPGMA of Nei's unbiased genetic distance (Fig. 2). The genetic identity estimated using AFLP markers ranged

from 0.9572 to 0.9933. Only three individuals from New York and three individuals from Indiana had the DNA fragments of EACAMCCT-149, indicating that this allele might be a rare or private allele endemic to these two populations. Therefore, the future conservation plan or sampling strategy should consider the importance of these local alleles especially in the extreme northern or southern regions of the natural range for *A. triloba*.

**GENETIC DIVERSITY IN DOMESTICATED PAWPAPS.** The genetic diversity of cultivated and wild pawpaws determined using dominant markers was 0.071 (Table 2). However, this estimation was much lower than that estimated using the RAPDs ( $H_c = 0.28$  for cultivated, Huang et al., 2003) and isozymes ( $H_c = 0.166$  for cultivated, Huang et al., 1997). Because most cultivars were directly selected and clonally propagated from the wild trees or selected as seedlings from open-pollinated wild trees (Peterson, 1991), the overall genetic composition has remained close to that observed in the wild populations (Huang et al., 2000). For a plant species like pawpaw in its early stage of domestication, extensive exploration for germplasm from wild populations in the natural range of *A. triloba* might be very helpful to provide more diversified breeding resources. Therefore, discovery and maintaining the genetic diversity is essential for future pawpaw improvement.

**CULTIVAR IDENTIFICATION AND GENETIC RELATEDNESS.** Thirty-one cultivars and advanced selections were discriminated by as few as nine AFLP dominant loci, comparing the identification efficiency of full discrimination by 14 RAPDs loci and 75% partial discrimination by 18 isozymes for a similar set of cultivars and advanced selections (Huang et al., 1997, 2003). Thus, the AFLP dominant markers were more effective in discriminating individuals than RAPDs and isozymes. Although codominant markers are much more informative than the dominant markers, the lack of them limited the discriminative ability in this study (data not shown).

All the isozyme (Huang et al., 1997), RAPD (Huang et al., 2000), and AFLP markers distinguished the PPF 8-20 from the maternal selection 'Sunflower'. AFLP banding profiles showed PPF 8-20 had 10 unique bands that did not appear in 'Sunflower,' which confirmed that PPF 8-20 had a different parent source than 'Sunflower'. PPF 8-58 and PPF 2-10 were recorded as open-pollinated seedlings from BEF-30, which was confirmed by the isozyme analysis with genetic distance of 0.016 (Huang et al., 1997). However, AFLP and RAPD analyses showed that PPF 8-58 was not subclustered with its half-sibling PPF 2-10, as illustrated by the genetic identity of 0.8797 and 0.520, respectively. Thus, the genetic identity results indicated that PPF 8-58 and PPF 2-10 probably originated from different maternal trees and might not be half-siblings. NC-1 was recorded as seedlings from 'Davis' × 'Overleese' (Layne, 1997). The AFLP profiles discovered seven unique marker bands in NC-1, but not in the assumed parents, which supported the results from isozymes, RAPDs, and ISSR that NC-1 was not closely related to 'Overleese' (Huang et al., 2003; Pomper, 2003). Because there was some doubt for the pollen parent (Huang et al., 2003) and NC-1 was clustered together with other PPF selections from the BEF selection group as indicated by isozymes, RAPDs, and AFLPs, NC-1 might have some genetic background from the BEF selection group. This study also clustered PPF 7-90 together with other BEF selections as indicated by RAPDs (Huang et al., 2003). Although PPF 7-90 was recorded as an open-pollinated seedling from RS-2, RS-2 might be descended from the BEF collection (Huang et al., 2003).

The genetic relatedness among the cultivars and advanced

selections was demonstrated by the UPGMA cluster analysis (Fig. 3). The dendrogram was constructed according to Nei's unbiased measures of genetic identity and genetic distance (Nei, 1978). The dendrogram developed using dominant markers (Fig. 3) separated 'Wells' from other cultivated pawpaw entries. 'Wells' was a cultivar selected in 1990 by D. Wells, from the wild in Salem, Ind. (Layne, 1997). PPF 11-5, PPF 11-13, and PPF 4-2 were clustered together with very high similarity (>0.96), confirming that they were the open-pollinated seedlings of a superior tree labeled as BEF53 (Peterson, 1991). Although sharing high similarity, they were not clustered directly in the same subgroup according to the RAPDs and isozymes analyses (Huang et al., 1997, 2003). The direct subclustering of PPF 1-23 and 'Taytwo', 'Overleese' and PPF 1-68, 'Taylor' and 'Wilson', and PPF 11-5 and PPF 11-13 were consistently discovered by AFLP and RAPD markers, which indicated parent-offspring or sibling relationships. Although 'Taylor' and 'Wilson' were wild seedlings from Eaton Rapids, Mich., and Harlan Co., Ky., respectively, the genetic identity of 90.6% by AFLPs, 86.7% by RAPDs, and 95.1% by isozymes indicated that they were possibly descended from one common ancestor. The genetic identity between two independent entries of selection PPF 9-58 was 0.9549 (Table 2) according to the AFLP dominant marker analysis results. They were thought to be half-siblings and were loosely clustered into a subgroup, with difference in six of 132 dominant AFLP markers. RAPD marker analysis showed genetic identity of 0.794 between these two entries and clustered them together, which indicated the possible half-sibling relationship between them (Huang et al., 2003). The pedigree information of 'Prolific' was not clear, as it was thought to be derived either from the pollen parent or from the female tree of 'Taytwo'. The clustering of 'Taytwo', PPF 1-23, and 'Prolific' indicated the parentage of 'Taytwo' according to both the AFLP and RAPD analysis.

**EFFECT OF MARKER CLUSTERING AND THE NUMBER OF SAMPLED GENOMIC REGIONS ON THE GENETIC DIVERSITY EVALUATION.** When estimating the genetic diversity among organisms, random and better coverage of genomes giving a maximum amount of information with minimum cost is always the best experimental strategy (Karp et al., 1996, 1997; Virk et al., 2000). Studies showed that levels of within-population diversity were only seldom correlated when their estimates were performed with different markers (Mariette et al., 2002). The number of sampled genomic regions does play an important role for the genetic diversity estimates. In this study, two codominant AFLP markers can only sample two regions in the pawpaw genome and estimate the genetic distances among the wild populations from 0.0005 to 1.000, which is less informative when compared with the genetic distances ranging from 0.0067 to 0.0476 estimated by 134 dominant markers. Gaudeul et al. (2004) also found the discrepancy between AFLPs and microsatellites for within-population genetic diversity in *Eryngium alpinum* L. may be explained by a better coverage of the genome with numerous AFLPs. In a study by Le Clerc et al. (2002) using *Daucus carota* L. accessions, the number of AFLP markers (20, 50, 60, and 70 AFLPs from 249 initial AFLP markers) required for accession separation was examined and 60 markers were adequate for separation of the genotypes examined. Thus, the number of AFLP markers does affect the genetic diversity estimates and the optimal number needed for estimating genetic diversity should be determined on the species by species basis (Barbosa-Neto et al., 1997; Le Clerc et al., 2002). When designing experiments, therefore, one should consider the clustering and co-segregating of molecular markers and the number of sampled genomic regions,

instead of only the number of markers. When small numbers of markers are used for genetic diversity estimates, the map positions of these markers are important, because randomly distributed markers have a better coverage of the genome (Karp et al., 1996, 1997; Laurie et al., 1997). In this study, 15 markers were found to cluster in three centromeric regions, which represented only three genomic regions. Nei's genetic diversity estimated for wild populations using these polymorphic clustered markers was  $H_e = 0.18$ , which was much lower than the results estimated using the same number of nonlinked polymorphic markers ( $H_e = 0.39$ ). This lower genetic diversity level might be due to the paucity of the sampled genomic regions. In addition to the number of markers used for genetic diversity estimates, the evolutionary history of the wild population and the genome heterogeneity are key points for having different experimental results of diversity estimates in a set of populations (Mariette et al., 2002).

In conclusion, this study provides the most comprehensive evaluation of the genetic diversity of wild and cultivated pawpaw in the United States using the same set of AFLP markers. This rich source of genetic information and the comprehensive marker profiles for major cultivars and advanced selections provided basic knowledge for future breeding and conservation programs. The genetic relatedness among the cultivars and selections could provide additional information for selecting cross parents in addition to the phenotypical traits. The high level of genetic diversity within populations indicated that the conservation strategy should focus on preserving the genetic resources in major populations. Genetic linkage groups were constructed based on 42 AFLP markers and one RAPD marker. Although this genetic linkage map was based on a limited number of markers and a small population, it could serve as the beginning of the future breeding project to identify quantitative and qualitative traits on the more saturated map. Clustering of AFLP markers might decrease the genetic diversity evaluation when using a limited number of markers. Further development of more complete and saturated genetic map would play a pivotal role in facilitating molecular breeding using marker-assisted selection (MAS) and revealing the population history based on the linkage disequilibrium.

#### Literature Cited

- Alonso-Blanco, C., A.J.M. Peeters, M. Koornneef, C. Lister, C. Dean, N. vanden Bosch, J. Pot, and M.T.R. Kuiper. 1998. Development of an AFLP based linkage map of Ler, Col and Cvi *Arabidopsis thaliana* ecotypes and construction of a LER/Cvi recombinant inbred line population. *Plant J.* 14:259–271.
- Barbosa-Neto, J.F., C.M. Hernandez, L.S. O'Donoghue, and M.E. Sorrells. 1997. Precision of genetic relationship estimates based on molecular markers. *Euphytica.* 98:59–67.
- Bartish, I.V., N. Jeppsson, and H. Nybom. 1999. Population genetic structure in the dioecious pioneer plant species *Hippophae rhamnoides* investigated by random amplified polymorphic DNA (RAPD) markers. *Mol. Ecol.* 8:791–802.
- Cavers, S., C. Navarro, and A.J. Lowe. 2003. A combination of molecular markers identifies evolutionarily significant units in *Cedrela odorata* L. (Meliaceae) in Costa Rica. *Conservation Genetics* 4:571–580.
- Faegri, K. and L. van der Pijl. 1971. The principles of pollination ecology. Pergamon Press, New York.
- Gaudeul, M., I. Till-Bottraud, F. Barjon, and S. Manel. 2004. Genetic diversity and differentiation in *Eryngium alpinum* L. (Apiaceae): Comparison of AFLP and microsatellite markers. *Heredity* 92:1–11.
- Hansen, M., T. Kraft, M. Christiansson, and N.O. Nilsson. 1999. Evaluation of AFLP in *Beta*. *Theor. Appl. Genet.* 98:845–852.
- Hamrick, J.L. and M.J.W. Godt. 1989. Allozyme diversity in plant species, p. 43–63. In: A.H.D. Brown, M.T. Clegg, A.L. Kahler, and B.S. Weir (eds.). *Plant population genetics, breeding and genetic resources*. Sinauer Assoc., Sunderland, Mass.
- Huang, H., D.R. Layne, and R.N. Peterson. 1997. Using isozyme polymorphisms for identifying and assessing genetic variation in cultivated pawpaw [*Asimina triloba* (L.) Dunal]. *J. Amer. Soc. Hort. Sci.* 122:504–511.
- Huang, H., D.R. Layne, and D. Riemenschneider. 1998. Genetic diversity and geographic differentiation in pawpaw [*Asimina triloba* (L.) Dunal] populations from nine states as revealed by allozyme analysis. *J. Amer. Soc. Hort. Sci.* 123:635–641.
- Huang, H., D.R. Layne, and T.L. Kubisiak. 2000. RAPD inheritance and diversity in pawpaw (*Asimina triloba*). *J. Amer. Soc. Hort. Sci.* 125:454–459.
- Huang, H., D.R. Layne, and T.L. Kubisiak. 2003. Molecular characterization of cultivated pawpaw (*Asimina triloba*) using RAPD markers. *J. Amer. Soc. Hort. Sci.* 128:85–93.
- Jones, S.C. and D.R. Layne. 1997. Cooking with pawpaws. Kentucky State Univ. Pawpaw Ext. Bul. 001.
- Jones, S.C., R.N. Peterson, T. Turner, K.W. Pomper, and D.R. Layne. 1998. Pawpaw planting guide: Cultivars and nursery sources. Kentucky State Univ. Pawpaw Ext. Bul. 002.
- Jordano, P. and J.A. Godoy. 2000. RAPD variation and population genetic structure in *Prunus mahaleb* (Rosaceae), an animal-dispersed tree. *Mol. Ecol.* 9:1293–1305.
- Karp, A., O. Seberg, and M. Buiatti. 1996. Molecular techniques in the assessment of botanical diversity. *Ann. Bot.* 78:143–149.
- Karp, A., K.J. Edward, M. Bruford, S. Frunk, and B. Vosman. 1997. Molecular technologies for biodiversity evaluation: Opportunities and challenges. *Nature Biotechnol.* 15:625–628.
- Kentucky State University. 2004. Pawpaw information web site. 10 Sept. 2004. <<http://www.pawpaw.kysu.edu>>.
- Ky, C.-L., P. Barre, M. Lorieux, P. Trouslot, S. Akaffou, J. Louarn, A. Charrier, S. Hamon, and M. Noirot. 2000. Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (*Coffea* sp.). *Theor. Appl. Genet.* 101:669–676.
- Lagrange, R.L. and E.J. Tramer. 1985. Geographic variation in size and reproductive success in the pawpaw (*Asimina triloba*). *Ohio J. Sci.* 85:40–45.
- Laurie, D.A., G.J. Bryan, and J.W. Snape. 1997. Genomic relationships, conserved synteny and wide-hybrids, p. 77–101. In: J.A. Callow, B.V. Ford-Lloyd, and H.J. Newbury (eds.). *Biotechnology and plant genetic resources conservation and use*. CABI, New York.
- Layne, D.R. 1996. The pawpaw [*Asimina triloba* (L.) Dunal]: A new fruit crop for Kentucky and the United States. *HortScience* 31:777–784.
- Layne, D.R. 1997. Pawpaw, p. 403–404. In: *The Brooks and Olmo, Register of Fruits & Nut Varieties*. ASHS Press, Alexandria, Va.
- Le Clerc, V., M. Briard, and P. Revollon. 2002. Influence of number and map distribution of AFLP markers on similarity estimates in carrot. *Theor. Appl. Genet.* 106:157–162.
- Lincoln, S., M. Daly, and E. Lander. 1992. Whitehead Institute technical report, 3rd ed. Whitehead Inst. for Biomed. Res., Cambridge, Mass.
- McLaughlin, J.L. 1997. Anticancer and pesticidal components of pawpaw (*Asimina triloba*). *Natl. Nut Growers Assn. Annu. Rpt.* 88:97–106.
- Mariette, S., V. Le Corre, F. Austerlitz, and A. Kremer. 2002. Sampling within the genome for measuring within-population diversity: Trade-offs between markers. *Mol. Ecol.* 11:1145–1156.
- Nei, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA.* 70:3321–3323.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89:583–590.
- Norman, E.M. and D. Clayton. 1986. Reproductive biology of two Florida pawpaws: *Asimina obovata* and *A. pygmaea* (Annonaceae). *Bul. Torrey Bot. Club.* 113:16–22.
- Norman, E.M., K. Rice, and S. Cochran. 1992. Reproductive biology of *Asimina parviflora* (Annonaceae). *Bul. Torrey Bot. Club.* 119:1–5.
- Oberlies, N.H., C.J. Chang, and J.L. McLaughlin. 1997a. Structure-

- activity relationships of diverse annonaceous acetogenins against multidrug resistant human mammary adenocarcinoma (MCF-7/Adr) cells. *J. Medicinal Chem.* 40:2102–2106.
- Oberlies, N.H., V.L. Croy, M.L. Harrison, and J.L. McLaughlin. 1997b. The annonaceous acetogenin bullatacin is cytotoxic against multidrug-resistant human mammary adenocarcinoma cells. *Cancer Lett.* 115:73–79.
- Peterson, R.N. 1991. Pawpaw (*Asimina*). In: J.N. Moore and J.R. Ballington (eds.). Genetic resources of temperate fruit and nut crops. *Acta Hort.* 290:567–600.
- Peterson, R.N. 2003. Pawpaw variety development: A history and future prospects. *HortTechnology* 13:449–454.
- Peters, J.L., H. Constandt, P. Neyt, G. Cnops, J. Zethof, M. Zabeau, and T. Gerats. 2001. A physical AFLP map of *Arabidopsis*. *Plant Physiol.* 127:1579–1589.
- Pomper, K.W., D.R. Layne, and R.N. Peterson. 1999. The pawpaw regional variety trial, p. 353–357. In: J. Janick (ed.). Perspectives on new crops and new uses. ASHS Press, Alexandria, Va.
- Pomper, K.W., S.B. Crabtree, S.P. Brown, S.C. Jones, T.M. Bonney, and D.R. Layne. 2003. Assessment of genetic diversity of pawpaw (*Asimina triloba*) cultivars with intersimple sequence repeat markers. *J. Amer. Soc. Hort. Sci.* 128:521–525.
- Prashanth, S.R., M. Parani, B.P. Mohanty, V. Talame, and A. Parida. 2002. Genetic diversity in cultivars and landraces of *Oryza sativa* subsp. *indica* as revealed by AFLP markers. *Genome* 45:451–459.
- Rogstad, S.H., K. Wolff, and B.A. Schaal. 1991. Geographical variation in *Asimina triloba* Dunal (Annonaceae) revealed by the M13 DNA fingerprinting probe. *Amer. J. Bot.* 78:1391–1396.
- Shannon, C.E. and W. Weaver. 1949. The mathematical theory of communication. Univ. of Illinois Press, Urbana, Ill.
- Strommer, J., J. Peters, J. Zethof, P.d. Keukeleire, and T. Gerats. 2002. AFLP maps of *Petunia hybrida*: Building maps when markers cluster. *Theor. Appl. Genet.* 105:1000–1009.
- Virk, P.S., H.J. Newbury, M.T. Jackson, and B.V. Ford-Lloyd. 2000. Are mapped markers more useful for assessing genetic diversity? *Theor. Appl. Genet.* 100:607–613.
- Vos, P., R. Hogers, M. Bleeker, M. Reijmans, T. Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zaneau. 1995. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Res.* 23:4407–4414.
- Vuyksteke, M., R. Mank, R. Antonise, E. Bastiaans, M.L. Senior, C.W. Stuber, A.E. Melchinger, T. Lubberstedt, X.C. Xia, P. Stam, M. Zabeau, and M. Kuiper. 1999. Two high-density AFLP linkage maps of *Zea mays* L.: Analysis of distribution of AFLP markers. *Theor. Appl. Genet.* 99:921–935.
- Wagner, D.B., G.R. Furnier, M.A. Saghai-Marooof, S.M. Williams, B.P. Dancik, and R.W. Allard. 1987. Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. *Proc. Natl. Acad. Sci. USA.* 84:2097–2100.
- Walker, J.W. 1971. Unique type of angiosperm pollen from the family Annonaceae. *Science* 172:565–567.
- Willson, M.F. and D.W. Schemske. 1980. Pollinator limitation, fruit production, and floral display in pawpaw (*Asimina triloba*). *Bul. Torrey Bot. Club.* 107:401–408.
- Yeh, R.C. and T.J.B. Boyle. 1997. Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian J. Bot.* 129:157.