Genetic Resources and Crop Evolution (2005) 52: 1069–1078
DOI 10.1007/s10722-004-6115-z

Identification of AFLP markers that discriminate between cultivated cotton and the Hawaiian island endemic, *Gossypium tomentosum* Nuttall ex Seeman

J.S. Hawkins, J. Pleasants and J.F. Wendel*
*Department of Ecology, Evolution, and Organismal Biology, 353 Bessey Hall, Iowa State University, Ames, IA 50011 USA; *Author for correspondence (e-mail: jfw@iastate.edu; fax: (515) 294-1337)*

Received 7 July 2003; accepted in revised form 10 January 2004

Key words: AFLP, Cotton, Gene flow, Genetic diversity, Genetic marker, Introgression, *Gossypium*

Abstract

Interspecific gene flow from crop species into wild relatives may compromise the genetic integrity of native species and in the case of transgenic crops lead to the escape of transgenes into natural populations. Approximately 72% of the upland cotton (*Gossypium hirsutum* L.) acreage grown in the United States in 2000 utilized transgenic cultivars. The closest relative of *G. hirsutum* is *G. tomentosum* Nuttall ex Seeman, endemic to the Hawaiian archipelago. Because these two species are fully interfertile, cultivation of transgenic *G. hirsutum* in the Hawaiian Islands is restricted. We investigated the possibility of identifying AFLP genetic markers that are diagnostic for each species and thus could be used in future studies to detect introgression between them. In addition, we were interested in comparing levels and geographic patterns of AFLP diversity in *G. tomentosum* to previous estimates using allozyme data. AFLP analysis led to the detection of 11 and 16 species-specific markers for *G. tomentosum* and *G. hirsutum*, respectively. These species-specific AFLP markers will be useful for detecting gene flow between *G. hirsutum* and *G. tomentosum* that has occurred in the past and thus might occur in the future if the restrictions on cultivation of transgenic *G. hirsutum* are relaxed in the Hawaiian Islands. Little genetic diversity and limited geographic patterning were discovered using AFLP markers, consistent with data from previous allozyme studies.

Introduction

A major concern with transgenic crop plants is the possibility of gene flow into wild relatives (Dale 1992; Gray and Raybould 1998). Interspecific gene flow between closely related species may compromise the genetic integrity of the wild relative and lead to the "escape" of transgenes into natural populations. In principle, introgression of transgenes into wild relatives may alter intra- and interspecific competitive relationships, lead to the origin of new weeds, or in some other way alter evolutionary trajectories (Raybould 1999; Storfer 1999; Auberson 2000). These issues are also relevant for non-transgenic crop species and their wild relatives (Arias and Rieseberg 1994; Arriola and Ellstrand 1996, 1997; Ellstrand et al. 1999; Jarvis and Hodgkin 1999), but interest is magnified when transgenic crops are involved because of the involvement of genetic material that could potentially alter wild relatives in unknown ways. Given these concerns, it is important to assess the possibility of gene transfer from cultivated plants into neighboring wild relatives, particularly wild relatives that are rare and/or endangered species.

For many major crop species grown in the United States, gene flow is not an issue because the wild relatives do not exist where the crop species are grown (e.g. corn, soybeans). For others, such as cotton, there are areas where native species

exist within the historical or modern regions of commercial cultivation. Ninety percent of the international commercial cotton crop consists primarily of Upland cotton (*Gossypium hirsutum* L.) and to a lesser extent Pima cotton (*G. barbadense* L.) (Wendel and Cronn 2003). Transgenic cotton cultivars have gained wide acceptance by growers, accounting for approximately 72% of the year 2000 cotton acreage grown in the United States (Bowman et al. 2003; USDA-AMS 1996; USDA-AMS 2000). There are three places in the United States where both commercial cotton and native cotton species are found and thus where gene flow between them could occur. One is in southernmost Florida, where both feral and putatively wild populations of *G. hirsutum* are found (Brubaker et al. 1999). A second is in southern Arizona where *G. thurberi* Todaro exists, although in this case the possibility for introgression with commercial cotton is limited due to differences in ploidy level. A third is in the Hawaiian Islands, where there is an endemic native *Gossypium* L. species, *G. tomentosum* Nuttall ex Seeman, that is closely related to and fully interfertile with both *G. barbadense* and *G. hirsutum* (Stephens 1964; DeJoode and Wendel 1992). In October 2000, the Scientific Advisory Panel supported the EPA's decision to prohibit cultivation of commercial Bt-cotton in Florida south of Highway 60 or in the state of Hawaii. Additionally, test plots and breeding nurseries located in Hawaii are under strict regulations regarding border crops for trapping pollinators and the proximity of these test plots to *G tomentosum* (see http://www.epa.gov/oppbppd1/biopesticides/pips/bt_brad.htm). A prelude to possible relaxation of these restrictions in Hawaii would be an evaluation of the possibility of gene flow between cultivated species and *G. tomentosum*.

The Hawaiian archipelago is the most isolated major group of islands in the world, separated by approximately 4,000 km from the nearest continent. *Gossypium tomentosum* is found as isolated coastal populations and scattered individuals on all of the major islands in the archipelago except the big island of Hawaii (Stephens 1964; DeJoode and Wendel 1992). Molecular data show convincingly that the closest relative of *G. tomentosum* is *G. hirsutum* (DeJoode and Wendel 1992; Wendel et al. 1995; Cronn et al. 1996; Seelanan et al. 1997; Small et al. 1998). Previous data support the notion that *G. tomentosum* originated following transoceanic dispersal from the New World tropics and that the founding population comprised only 1 to several individuals (DeJoode and Wendel 1992).

To detect possible gene flow between commercial cotton and *G. tomentosum* it is necessary to establish genetic markers that can unequivocally identify genomic fragments unique to each species. AFLP (Amplified Fragment Length Polymorphism) analysis (Vos et al. 1995) has found broad applicability to analyses of patterns of genetic variation within and among populations (Gaudeul et al. 2000; Diaz et al. 2001; Hedren et al. 2001; Zawko et al. 2001; Bottini et al. 2002; Rivera-Ocasio et al. 2002), studies of crop origins (Sasanuma et al. 1996; Lubberstedt et al. 2000; Soleimani et al. 2002) and relationships among cultivars (Mace et al. 1999; He and Prakash 2001; Hagen et al. 2002; Kim et al. 2002; Simioniuc et al. 2002). AFLP markers are highly repeatable (Jones et al. 1997; Blears et al. 1998), provide broad genomic coverage and a virtually limitless number of genetic markers. These are particularly useful features for conservation genetic and rare/endangered species applications (Vos et al. 1995; Mueller and Wolfenbarger 1999).

In addition to addressing the question of possible interspecific gene flow, we were interested in what AFLPs could tell us about the patterns of genetic diversity within *G. tomentosum*. A previous study of genetic diversity within *G. tomentosum* using 14 different allozyme systems (DeJoode and Wendel 1992) found low levels of variation and no discernable patterns of geographic differentiation among populations from different islands. We were particularly interested in determining whether using a more powerful molecular marker methodology would provide greater resolution than previously employed methods, and if this presumed increase in resolving power would enable detection of diagnostic markers for populations, groups of local populations, or populations from different islands.

Methods

Plant material

Accessions of *G. tomentosum* collected from six of the eight major islands were obtained from the

Table 1. Accessions of *G. tomentosum* and *G. hirsutum* studied.

Species	Collection site	Accession(s) or HPDL #
G. tomentosum Nuttall ex Seeman	Niihau	AD3-16
	Oahu	AD3-1, AD3-2, AD3-3, AD3-5, AD3-7, AD3-21, AD3-22, AD3-25, AD3-26, AD3-28, AD3-Ma and Pa Wendel 1, AD3- Ma and Pa Wendel 2, AD3-A80-1239
	Molokai	AD3-14, AD3-15
	Maui	AD3-12, AD3-13, AD3-A82-1247
	Kauai	AD3-23*
	Kahoolawe	3685, 3686, 3687, 3688, 3689, 3690, 3691, 3692, 3693
	Lanai	AD3-10, AD3-11, 3694, 3695, 3696, 3697, 3698, 3699, 3700, 3701, 3702, 3703, 3704, 3705, 3706
	Unknown	AD3-A80-1238, Old Tom 9/1/80, AD3-17, AD3-7X
G. hirsutum L.	Wake Atoll	Wilke's Island
	Guatemala	TX184, TX99, TX379, TX180, TX188, TX115
	Mexico	TX58, TX746, TX461, TX745, TX1046, TX2094, TX2089, TX481
	Honduras	TX1045
	Belize	TX794
	Cultivated	Tamcot SP-215, Cascot 4, Delcot 344

*It has come to our attention that the correct locality for accession AD3-23 is Maui, not Kauai.

National Collection of *Gossypium* L. Germplasm, maintained in College Station, Texas (Table 1). DNAs for all individuals from Kahoolawe and all individuals with the exception of AD3–10 and AD3–11 from Lanai were obtained from Clifford Morden, University of Hawaii, Honolulu (Morden et al. 1996; Randell and Morden 1999). A total of 44 *G. tomentosum* individuals spanning the archipelago were distributed as follows: Niihau, $N = 1$; Oahu, $N = 13$; Molokai, $N = 2$; Maui, $N = 3$; Kahoolawe, $N = 9$; Lanai, $N = 15$; Kauai, $N = 1$. In addition, four *G. tomentosum* accessions of unknown geographic origin were included in the AFLP analysis. *G. hirsutum* L. accessions were also obtained from the National Collection of *Gossypium* Germplasm. These accessions span a wide geographic range and represent wild, feral, and cultivated forms (Wendel et al. 1992). Seeds were germinated in 1 mM gibberellic acid (Ga3) and seedlings were planted and grown to maturity in the Richard W. Pohl Conservatory at Iowa State University. Voucher specimens for all accessions were placed in the Ada Hayden Herbarium (ISC).

DNA isolation and multi-fluorophore AFLP analysis

Total genomic DNA was isolated from young expanding leaves using the Nucleon Plant DNA Isolation and Purification Kit (Amersham Corp., Piscataway, NJ) according to the kit protocol. AFLP analysis was performed as described (Vos et al. 1995) with some modifications. Approximately 200 ng total genomic DNA was digested with 10 U *Eco*RI and 10 U *Mse*I in a 20 μL reaction and incubated at 37 °C for 3 h. Subsequently, 20 μL of ligation master mix containing 75 pmol each *Mse*I and *Eco*RI adapter (Table 2) with 20 U T4 DNA ligase in 1X T4 DNA ligase buffer was added to the digested fragments and incubated overnight at 16 °C. The digestion-ligation mixture was diluted with 160 μL sterile dH_2O. Pre-selective AFLP amplification was performed by using a single selective nucleotide (+1) at the 3′ end of both the Mse and Eco primers (Table 2). Each 50 μL +1 AFLP reaction contained 10 μL dilute digestion-ligation reaction, 1X PCR buffer, 1.5 mM $MgCl_2$, 200 uM dNTP,

Table 2. AFLP primer and adapter sequences and amplification combinations.

	Sequence or description
Adapters	
*Mse*I forward adapter	5′-GAC GAT GAG TCC TGA G-3′
*Mse*I reverse adapter	5′-TAC TCA GGA CTC AT-3′
*Eco*RI forward adapter	5′-CTC GTA GAC TGC GTA CC-3′
*Eco*RI reverse adapter	5′-AAT TGG TAC GCA GTC-3′
+1 preselective primers	
*Mse*I + C	5′-GAC GAT GAG TCC TGA GTA A **C**-3′
*Eco*RI + A	5′-GAC TGC GTA CCA ATT C **A**-3′
+3 selective primers (combination)	
*Mse*I + CAA (I, II)	5′-GAC GAT GAG TCC TGA GTA A **CAA**-3′
*Mse*I + CAC (III)	5′-GAC GAT GAG TCC TGA GTA A **CAC**-3′
*Eco*RI + AGC (I, III)	5′-GAC TGC GTA CCA ATT C **AGC**-3′ (FAM)
*Eco*RI + ACA (II)	5′-GAC TGC GTA CCA ATT C **ACA**-3′ (FAM)
*Eco*RI + ACG (I, III)	5′-GAC TGC GTA CCA ATT C **ACG**-3′ (TET)
*Eco*RI + AAC (II)	5′-GAC TGC GTA CCA ATT C **AAC**-3′ (TET)

Six +3 selective AFLP amplifications were performed in three multiplexed reactions labeled I, II, and III. Roman numerals in parentheses indicate inclusion in a specific multi-plexed reaction. Bold-typed nucleotides designate the selective nucleotides (either 1 for preselective amplifications of 3 for selective amplifications) and fluorescent labeling is specified by either FAM or TET at the end of the primer sequence.

40 pmol of each +1 primer, and 2.5 U Taq DNA polymerase. The amplification profile was 75 °C for 2 min, followed by 20 cycles of 94 °C for 30 s, 56 °C for 30 s and 75 °C for 2 min, ending with 60 °C for 30 min. Forty μL of each 50 μL + 1 reaction was diluted with 720 μL dH_2O.

Selective AFLP amplifications were performed by adding three selective nucleotides (+3) to the end of the Mse and Eco primers (Table 2). Six +3 selective AFLP amplifications were performed using 3 primer combinations (combinations I, II, and III; see Table 2). Eco primers were 5′ fluorescently labeled with either 6-carboxyfluorescein (FAM) or 4,7,2′,7′-tetrachloro-6-carboxyfluorescein (TET). Mse primers were unlabeled. Each 25 μL +3 reaction contained 5 μL dilute +1 reaction, 1X PCR buffer, 1.5 mM $MgCl_2$, 300 μM dNTP, 4 pmol each Eco +3 primer, 25 pmol Mse +3 primer, and 1 U Taq DNA polymerase. The amplification profile was 94 °C for 2 min, 100 cycles of 94 °C for 30 s, 65 °C for 30 s, 72 °C for 2 min, reducing the annealing temperature by 1 °C per cycle, followed by 35 cycles of 94 °C for 30 s, 56 °C for 30 s, 72 °C for 2 min, ending with 60 °C for 30 min.

Data analysis

Selective +3 AFLP amplification products were resolved using automated sequencing gels on an ABI Prism® 373 DNA sequencer at the DNA synthesis and sequencing facility at Iowa State University. Image analysis was performed using Genescan® software, version 2.0.2 and Genotyper®, version 2.0 (PE/Applied Biosystems), and by visual inspection. Similarity of fragment size was assumed to indicate homology. A replicate amplification was performed for each accession, and samples and/or fragments lacking reproducible banding patterns were removed from the data set. Fragment data were recorded as + (present) or − (absent).

Genetic identity and distance between *G. tomentosum* and *G. hirsutum* and among populations of *G. tomentosum* were determined using Nei's unbiased measures of genetic identity and genetic distance (Nei and Li 1979). NTSYS-pc (Rohlf 1998) was employed to perform PCA analysis in an attempt to visualize the genetic differences between *G. tomentosum* and *G. hirsutum* and to detect geographical patterns among *G. tomentosum* populations. Genetic diversity within *G. tomentosum* was hierarchically partitioned using Analysis of Molecular Variance (AMOVA) (Excoffier et al. 1992) in the Arlequin program version 2.000 (Schneider et al. 2000). Expected heterozygosity within and among populations of both species was calculated using PopGene version 1.31 (Yeh and Boyle 1997).

Results

Using the six primer combinations, a total of 351 readily scored and reproducible AFLP markers were generated for each *G. tomentosum*

Table 3. Species-specific markers revealed by AFLP analysis.

Species/Marker	Notes	Location(number of accessions positive for marker)
G. tomentosum		
EMF1-303.5	Polymorphic (85.4% +)	O(13), N(1), Mo(2), Ma(3), Ku(1), Kh(4), L(13), U(4)
EMF3-169.0	Monomorphic	Present in all accessions
EMF3-170.0	Monomorphic	Present in all accessions
EMF3-188.5	Polymorphic (85.4% +)	O(9), N(1), Mo(2), Ma(3), Kh(9), L(13), U(4)
EMF3-189.5	Monomorphic	Present in all accessions
EMF3-204.0	Polymorphic (92.0% +)	O(13), N(1), Mo(2), Ma(3), Ku(1), Kh(6), L(14), U(4)
EMF3-224.5	Monomorphic	Present in all accessions
EMT1-85.5	Polymorphic (4.2% +)	Kh(1), L(1)
EMT2-99.0	Monomorphic	Present in all accessions
EMT3-118.6	Monomorphic	Present in all accessions
EMT3-368.0	Polymorphic (2.1% +)	Kh(1)
G. hirsutum		
EMF1-156.0	Polymorphic (65% +)	G(4), M(5), H(1), B(1), C(2)
EMF1-195.5	Polymorphic (75% +)	W(1). G(5), M(7), H(1), B(1)
EMF2-122.6	Monomorphic	Present in all accessions
EMF2-165.0	Monomorphic	Present in all accessions
EMF2-166.5	Polymorphic (80% +)	W(1), G(6), M(5), B(1), C(3)
EMF3-54.0	Monomorphic	Present in all accessions
EMF3-183.5	Polymorphic (60% +)	W(1), G(2), M(2), C(3)
EMF3-184.5	Monomorphic	Present in all accessions
EMT1-153.0	Polymorphic (45% +)	G(2), M(6), B(1)
EMT2-120.0	Monomorphic	Present in all accessions
EMT3-78.5	Monomorphic	Present in all accessions
EMT3-136.0	Polymorphic (5% +)	M(1)
EMT3-143.7	Polymorphic (40% +)	W(1), M(5), H(1), B(1)
EMT3-149.5	Monomorphic	Present in all accessions
EMT3-337.0	Monomorphic	Present in all accessions
EMT3-350.6	Monomorphic	Present in all accessions

Markers are named as follows: E = *Eco*RI, M = *Mse*I, F = FAM labeled, T = TET labeled; 1, 2, or 3 = primer combination I, II, or III, respectively; subsequent number = size of the fragment in base pairs. *G. tomentosum* accession locations are abbreviated O = Oahu, N = Niihau, Mo = Molokai, Ma = Maui, Ku = Kauai, Kh = Kahoolawe, L = Lanai, U = Unknown. *G. hirsutum* accession locations are abbreviated W = Wilke's Island, G = Guatemala, M = Mexico, H = Honduras, B = Belize, C = cultivated accession.

individual. Of these, 22 (6.3%) were polymorphic but only 11 of the 22 polymorphic loci attained a frequency greater than 0.10 for the rare allele. A total of 356 AFLP markers were generated for *G. hirsutum*, of which 41 were polymorphic (11.5%). A total of 11 species-specific AFLP markers were detected for *G. tomentosum* (Table 3). Six of these were present in all *G. tomentosum* accessions studied, whereas the remaining five were present only in some of the accessions. Of these latter five markers, three were present in all but a few accessions, while the remaining two were present only in one or two accessions. Sixteen species-specific markers were generated for *G. hirsutum*, of which nine were uniformly present within the species. Thus, a total of 27 AFLP markers distinguish the two species (half of which are monomorphic within species), and can be used to diagnose introgression via observation of *G. hirsutum* specific markers in *G. tomentosum*, and vice versa.

To explore overall patterns of genetic differences between the two species, multivariate techniques were employed. PCA analysis on *G. tomentosum* and *G. hirsutum* combined AFLP data show a clear distinction between these two closely related species (Figure 1A). There are no accessions of *G. tomentosum* that occupy an intermediate

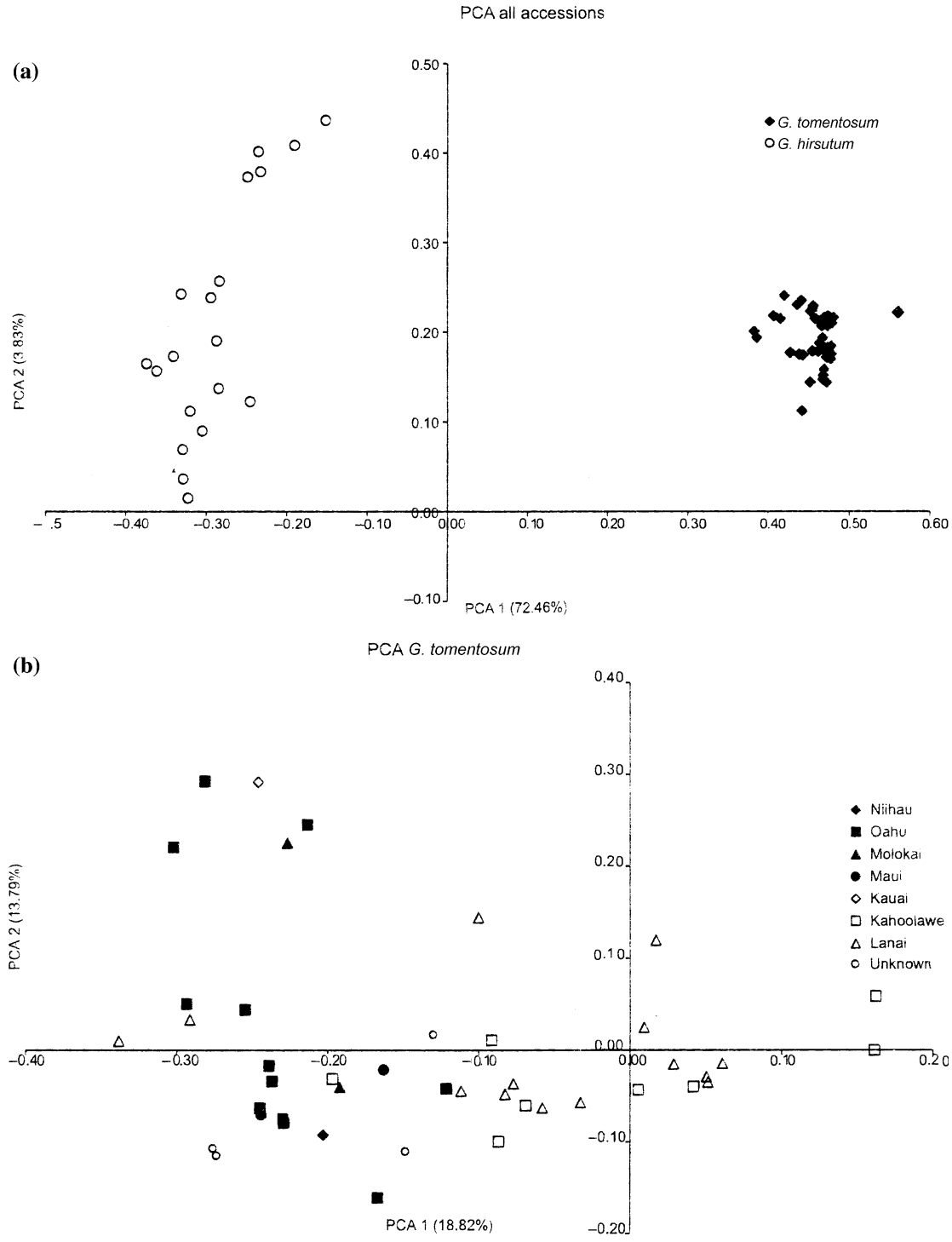

Figure 1. Principal component analysis (PCA) of AFLP data for *G. tomentosum* and *G. hirsutum* accessions based on the variance-covariance matrices of AFLP fragment frequencies. Percent of total variance explained by the first two principal components is shown on the abscissa and ordinate axes, respectively. A. PCA of both *G. tomentosum* and *G. hirsutum* accessions. B. PCA of *G. tomentosum* accessions. All *G. tomentosum* accessions were included in the PCA analysis. However, accessions with identical banding patterns plot to the same position and are indistinguishable from one another.

Table 4. Nei's unbiased measure of genetic identity and genetic distance among islands.

pop ID	1	2	3	4	5	6	7
1	*****	0.996	0.989	0.991	0.980	0.990	0.992
2	0.004	*****	0.996	0.996	0.992	0.993	0.996
3	0.011	0.004	*****	0.997	0.993	0.994	0.994
4	0.009	0.004	0.003	*****	0.993	0.995	0.994
5	0.019	0.009	0.007	0.008	*****	0.989	0.990
6	0.011	0.007	0.006	0.006	0.011	*****	0.998
7	0.008	0.004	0.006	0.006	0.010	0.002	*****

Nei's genetic identity is above the diagonal and genetic distance is below the diagonal. Population designations are as follows: 1 = Niihau, 2 = Oahu, 3 = Molokai, 4 = Maui, 5 = Kauai, 6 = Kahoolawe, 7 = Lanai.

position in PCA space indicating that there is no evidence of introgression in our sample.

With regard to genetic diversity within *G. tomentosum*, the low level of allelic diversity observed was reflected in the estimate of mean panmictic heterozygosity (Ht; Nei 1987) for the species as a whole. Utilizing all 351 markers, Ht was estimated at 0.012 ± 0.004. The apportionment of genetic diversity within *G. tomentosum* was explored using AMOVA. Results for a global AMOVA and locus-by-locus AMOVA employing only polymorphic markers demonstrated that there was little inter-population differentiation. Only 13.2% of the variation occurred among populations, with 86.8% of the variation residing within populations. Similarly, Nei's unbiased measure of genetic identity revealed only small differences among islands (Table 4), but hinted at a relationship between genetic and geographic distance. The lowest estimate of genetic distance was between populations from two adjacent islands, Lanai and Kahoolawe. Indeed, 12 of the 22 *G. tomentosum* polymorphic markers were shared between these islands, and seven markers were specific to these two islands. Four of the 22 markers were specific to the central islands of Oahu, Maui, Lanai and Kahoolwe and two additional markers suggested a shared history among accessions from Lanai, Kahoolawe, and Oahu. Although the first two principal components explain slightly less than one-third (32.61%) of the total variance, a projection of populations defined by these components reveals a spatial patterning among populations (Figure 1B).

Six of the polymorphic markers in *G. tomentosum* were specific to only one accession and therefore do not bear on relationships among accessions. These markers, however, may be used as island-specific markers, with the caveat that sampling was rather limited, particularly for some islands. The apparent existence of island-specific markers suggests that these might be used to identify the origin of accessions in germplasm collections that lack locality information. As an example, two markers suggest that "Old Tom" is from Oahu while one marker supports the same conclusion that AD3–7x and AD3-A80–1238 are from Oahu.

Discussion

Detecting gene exchange between G. hirsutum and G. tomentosum

The sister-species relationship between *G. tomentosum* and *G. hirsutum* is well-supported by allozyme, 5S ribosomal DNA, gene sequence, and cpDNA restriction site data (DeJoode and Wendel 1992; Wendel et al. 1995; Cronn et al. 1996; Seelanan et al. 1997; Small et al. 1998), as well as by the observation that interspecific hybrids between the two species are readily synthesized and have high fertility (Wendel, pers. obs.). Based originally on chloroplast DNA data (Wendel et al. 1989) and more recently by extensive analysis of sequence variation at multiple nuclear loci (Senchina et al. 2003), polyploidization in *Gossypium* occurred 1–2 mya, marking the earliest possible date for colonization of the Hawaiian islands by the forerunner of modern *G. tomentosum*. This date is consistent with the high interspecific genetic identity between *G. tomentosum* and *G. hirsutum* as inferred from the AFLP data (I = 0.90) as well as DeJoode and Wendel (1992) estimate of 0.82 for allozyme loci.

It is not surprising, therefore, that the majority of AFLP markers are shared between these two species. Not withstanding this overall similarity, the prospects for detecting gene flow are good, because numerous species-specific markers were detected. A total of 11 and 16 diagnostic markers were observed for *G. tomentosum* and *G. hirsutum*, respectively, with half of these being monomorphic within species (Table 3). These AFLP markers provide a sensitive means for evaluation of the genetic integrity of *G. tomentosum* individuals or populations. Only one of the three multiplex primer combinations (combination III) used in this study generated more than two markers diagnostic for *G. tomentosum*. Toward this end, we recommend that future monitoring activities should optimally employ *Eco*RI + *Mse*I primer combination III, as these most efficiently produced species-specific markers.

Native populations of *G. tomentosum* have already been exposed to potential gene flow from *G. hirsutum*. Cultivars of *G. hirsutum* were introduced to the Hawaiian Islands as long ago as the late 18th or early 19th century (Stephens 1963; Stephens 1964; Bates 1990) and feral escapes from these earlier periods of cultivation survive today. This raises the possibility of historical gene flow between these closely related species. As shown before, using allozyme data (DeJoode and Wendel 1992), AFLP analysis did not produce compelling evidence of interspecific introgression for any of the accessions analyzed. However, the sample sizes for *G. tomentosum* populations were small and thus unlikely to detect low levels of gene flow. In some cases, inspection of marker distributions revealed shared single fragments between individual *G. hirsutum* accessions and *G. tomentosum*, but other evidence indicates that these cases are likely to reflect shared ancestry or size homoplasy rather than introgression. For example, we observed several differences between TX2094, a truly wild plant from the north coast of the Yucatan Peninsula, and the other *G. hirsutum* accessions. For four of the AFLP markers, the band was present in all *G. tomentosum* accessions and absent in all *G. hirsutum* accessions with the exception of TX2094. In addition, one marker was present in all *G. hirsutum* accessions except TX2094 and absent in all *G. tomentosum* accessions.

A larger scale study using the AFLP markers we have identified could detect low levels of gene flow. The detection of historical gene flow would mean that gene flow between transgenic cotton and *G. tomentosum* would be possible. The potential for introgression should be considered realistic. A recent survey (Pleasants and Wendel, unpubl.) of flowering phenology and visitation by potential pollen vectors demonstrates the possibility of gene flow given reasonable proximity between populations of commercial cotton and those of *G. tomentosum*. Most notably, non-native honeybees and carpenter bees have been observed to forage on flowers in native stands of *G. tomentosum*. Given the ability of these species to forage several kilometers from their hive, up to a maximum of about 12 km (Seeley 1985), and given that they commonly visit flowers of commercial cotton (Moffet 1983), they are potential vectors for hybridization.

Genetic diversity in Gossypium tomentosum

Genetic diversity within *G. tomentosum* has previously been assessed using 14 different allozyme systems (DeJoode and Wendel 1992). The study included 30 accessions of *G. tomentosum* from all of the major islands on which *G. tomentosum* is found. Allozyme analysis revealed low levels of panmictic heterozygosity (Ht = 0.033), no definitive evidence of introgression from either *G. barbadense* or *G. hirsutum*, and no obvious patterns of geographic differentiation among populations from different islands.

Perhaps not unexpectedly, therefore, the amount of genetic diversity detected within and among *G. tomentosum* populations using AFLP markers is extremely low. Only 6.3% of the 351 AFLP markers were polymorphic and mean panmictic heterozygosity (Ht) for this species is 0.012 ± 0.004. This low level of genetic variation is reflected in genetic identity estimates, which are uniformly high within and among populations from the different islands. If colonization of the Hawaiian Islands by the ancestor of modern *G. tomentosum* occurred within the last 1–2 million years from trans-oceanic dispersal of a single individual, these results are not surprising. In addition to experiencing an extreme genetic bottleneck at the time of formation, several life-history charac-

teristics of the species would collectively retard its subsequent accumulation of genetic diversity, including self-compatibility and probable high selfing rates and the fact that populations are small and scattered with limited recruitment (Wendel and Pleasants, pers. obs.). In addition, the long generation time of the species means that the number of generations since species origin is lower than for many taxa, lessening the opportunity for fixation of mutations that might be detected by AFLP analysis. With an estimated maximum age of 1–2 million years (Wendel et al. 1989; Senchina 2003; Wendel and Cronn 2003), therefore, little inter-island genetic differentiation is expected to have arisen, and indeed, little was observed. We note, however, that the limited genetic variation detected supports the interpretation of isolation by distance, as geographic and genetic distance are weakly associated, and because of the presence of diagnostic bands that are exclusively shared among populations from adjacent islands (as noted in Results).

Acknowledgements

The authors thank Cliff Morden for generously providing DNAs for *Gossypium tomentosum* accessions from the islands of Kahoolawe and Lanai, Chris Wozniak for helpful discussion, John Nason for providing help and advice with computer programs, and Rodney Dyer for assistance with AFLP marker analysis. This study was funded by grants from Monsanto Inc., the US Environmental Protection Agency, and the National Science Foundation.

References

Arias D.M. and Rieseberg L.H. 1994. Gene flow between cultivated and wild sunflowers. Theor. Appl. Genet. 89: 655–660.

Arriola P.E. and Ellstrand N.C. 1996. Crop-to-weed gene flow in the genus *Sorghum* (Poaceae): Spontaneous interspecific hybridization between johnsongrass, *Sorghum halepense*, and crop sorghum, *S. bicolor*. Amer. J. Bot. 83: 1153–1159.

Arriola P.E. and Ellstrand N.C. 1997. Fitness of interspecific hybrids in the genus *Sorghum*: persistence of crop genes in wild populations. Ecol. Appl. 7: 512–518.

Auberson L. 2000. Transgenic crops from another perspective. Tibtech 18: 404–405.

Bates D.M. 1990. Malvaceae. In: Wagner W.L., Herbst D.R. and Sohmer S.H. (eds), Manual of the flowering plants of Hawaii, University of Hawaii Press Bishop Museum Press, Honolulu, HI, pp. 868–902.

Blears M.J., De Grandis S.A., Lee H. and Trevors J.T. 1998. Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. J. Indust. Microbiol. Biotechnol. 21: 99–114.

Bottini M.C.J., De Bustos A., Jouve N., and Poggio L. 2002. AFLP characterization of natural populations of *Berberis* (Berberidaceae) in Patagonia, Argentina. Plant Syst. Evol. 231: 133–142.

Bowman D.T., May O.L. and Creech J.B. 2003. Genetic uniformity of the US upland cotton crop since the introduction of transgenic cottons. Crop Sci. 43: 515–518.

Brubaker C.L., Bourland F.M. and Wendel J.F. 1999. The origin and domestication of cotton. In: Smith C.W. and Cothren J.T. (eds), Cotton; Origin, History, Technology and Production, John Wiley and Sons, NY, pp. 3–31.

Cronn R.C., Zhao X., Paterson A.H. and Wendel J.F. 1996. Polymorphism and concerted evolution in a tandemly repeated gene family: 5S ribosomal DNA in diploid and allopolyploid cottons. J. Mol. Evol. 42: 685–705.

Dale P.J. 1992. Spread of engineered genes to wild relatives. Plant Physiol. 100: 13–15.

DeJoode D.R. and Wendel J.F. 1992. Genetic diversity and origin of the Hawaiian Islands cotton, *Gossypium tomentosum*. Amer. J. Bot. 79: 1311–1319.

Diaz V., Muniz L.M. and Ferrer E. 2001. Random amplified polymorphic DNA and amplified fragment length polymorphism assessment of genetic variation in Nicaraguan populations of *Pinus oocarpa*. Mol. Ecol. 10: 2593–2603.

Ellstrand N.C., Prentice H.C. and Hancock J.F. 1999. Gene flow and introgression from domesticated plants and their wild relatives. Ann. Rev. Ecol. Syst. 30: 539–563.

Excoffier L., Smouse P. and Quattro J. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479–491.

Gaudeul M., Taberlet P. and Till-Bottraud I. 2000. Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), inferred from amplified fragment length polymorphism markers. Mol. Ecol. 9: 1625–1637.

Gray A.J. and Raybould A.F. 1998. Crop genetics: reducing transgene escape routes. Nature 392: 653–654.

Hagen L.S., Khadari B., Lambert P. and Audergon J.-M. 2002. Genetic diversity in apricot revealed by AFLP markers: species and cultivar comparisons. Theor. Appl. Genet. 105: 298–305.

He G.H. and Prakash C. 2001. Evaluation of genetic relationships among botanical varieties of cultivated peanut (*Arachis hypogaea* L.) using AFLP markers. Genet. Resour. Crop Evol. 48: 347–352.

Hedren M., Fay M.F. and Chase M.W. 2001. Amplified fragment length polymorphisms (AFLP) reveal details of polyploid evolution in *Dactylorhiza* (Orchidaceae). Amer. J. Bot. 88: 1868–1880.

Jarvis D.I. and Hodgkin T. 1999. Wild relatives and crop cultivars: detecting natural introgression and farmer selection of

new genetic combinations in agroecosystems. Mol. Ecol. 8: 159–173.

Jones C.J., Edwards K.J., Castaglione S., Winfield M.O., Sala F., van de Weil C., Bredemeijer G., Vosman B., Mattes M., Daly A., Brettschneider R., Bettini P., Buiatti M., Maestri E., Malcevschi A., Marmiroli N., Aert R., Volckaert G., Rueda J., Linacero R., Vazquez A. and Karp A. 1997. Reproducibility testing of RAPD, AFLP, and SSR markers in plants by a network of European laboratories. Mol. Breed 3: 381–390.

Kim M.S., Moore P.H., Zee F., Fitch M.M.M., Steiger D.L., Manshardt R.M., Paull R.E., Drew R.A., Sekioka T. and Ming R. 2002. Genetic diversity of *Carica papaya* as revealed by AFLP markers. Genome 45: 503–512.

Lubberstedt T., Melchinger A.E., Dussle C., Vuylesteke M. and Kuiper M. 2000. Relationships among early European maize inbreds: IV. Genetic diversity revealed with AFLP markers and comparison with RFLP, RAPD, and pedigree data. Crop Sci. 40: 783–791.

Mace E.S., Lester R.N. and Gebhardt C.G. 1999. AFLP analysis of genetic relationships among the cultivated eggplant, *Solanum melongena* L., and wild relatives (Solanaceae). Theor. Appl. Genet. 99: 626–633.

Moffet J.O. 1983. Pollination of entomophilous hybrid seed parents. In: Jones C.E. and Little R.J. (eds), Handbook of Experimental Pollination Biology, Van Nostrand Reinhold Co., New York, NY, pp. 508–514.

Morden C.W., Caraway V.C. and Motley T.J. 1996. Development of a DNA library for native Hawaiian plants. Pacific Sci. 50: 324–335.

Mueller U.G. and Wolfenbarger L.L. 1999. AFLP genotyping and fingerprinting. Trends Ecol. Evol. 14: 389–394.

Nei M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York, NY, USA.

Nei M. and Li W.H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76: 5269–5273.

Randell R.A. and Morden C.W. 1999. Hawaiian plant DNA library II: endemic, indigenous, and introduced species. Pacific Sci. 53: 401–417.

Raybould A.F. 1999. Transgenes and agriculture going with the flow. Trends Plant Sci. 4: 247–248.

Rivera-Ocasio E., Aide T.M. and McMillan W.O. 2002. Patterns of genetic diversity and biogeographical history of the tropical wetland tree, *Pterocarpus officinalis* (Jacq.), in the Caribbean basin. Mol. Ecol. 11: 675–683.

Rohlf F.J. 1998. NTSYSpc Numerical Taxonomy and Multivariate Analysis System. Applied Biostatistics Inc., Setauket, New York.

Sasanuma T., Miyashita N.T. and Tsunewaki K. 1996. Wheat phylogeny determined by RFLP analysis of nuclear DNA. 3. Intra- and interspecific variations of five *Aegilops* Sitopsis species. Theor. Appl. Genet. 92: 928–934.

Schneider S., Roessli D. and Excoffier L. 2000. Arlequin: A software for population genetics data analysis. Ver 2.000. Genetics and Biometry Lab, Department of Anthropology, University of Geneva.

Seelanan T., Schnabel A. and Wendel J.F. 1997. Congruence and consensus in the cotton tribe. Syst. Bot. 22: 259–290.

Seeley T.D. 1985. Honeybee Ecology. Princeton University Press, Princeton, New Jersey, USA.

Senchina D.S., Alvarez I., Cronn R.C., Liu B., Rong J.K., Noyes R.D., Paterson A.H., Wing R.A., Wilkins T.A. and Wendel J.F. 2003. Rate variation among nuclear genes and the age of polyploidy in *Gossypium*. Mol. Biol. Evol. 20: 633–643.

Simioniuc D., Uptmoor R., Friedt W. and Ordon F. 2002. Genetic diversity and relationships among pea cultivars revealed by RAPDs and AFLPs. Plant Breed. 121: 429–435.

Small R.L., Ryburn J.A., Cronn R.C., Seelanan T. and Wendel J.F. 1998. The tortoise and the hare: choosing between noncoding plastome and nuclear *Adh* sequences for phylogeny reconstruction in a recently diverged plant group. Amer. J. Bot. 85: 1301–1315.

Soleimani V.D., Baum B.R. and Johnson D.A. 2002. AFLP and pedigree-based genetic diversity estimates in modern cultivars of durum wheat [*Triticum turgidum* L. subsp. *durum* (Desf.) Husn.]. Theor. Appl. Genet. 104: 350–357.

Stephens S.G. 1963. Polynesian cottons. Ann. Missouri Bot. Garden 50: 1–22.

Stephens S.G. 1964. Native Hawaiian cotton (*Gossypium tomentosum* Nutt.). Pacific Sci. 18: 385–398.

Storfer A. 1999. Gene flow and endangered species translocations: a topic revisited. Bio. Conserv. 87: 173–180.

USDA-AMS,1996. Cotton varieties planted-1996. USDA-AMS, Memphis, TN.

USDA-AMS,2000. Cotton varieties planted-2000. USDA-AMS, Memphis, TN.

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 fingerprinting. Nucleic Acids Res. 23: 4407–4414.

Wendel J.F., Brubaker C.L. and Percival A.E. 1992. Genetic diversity in *Gossypium hirsutum* and the origin of Upland cotton. Amer. J. Bot. 79: 1291–1310.

Wendel J.F. and Cronn R.C. 2003. Polyploidy and the evolutionary history of cotton. Adv. Agron. 78: 139–186.

Wendel J.F., Olson P.D. and Stewart J.M. 1989. Genetic diversity, introgression and independent domestication of Old World cultivated cottons. Amer. J. Bot. 76: 1795–1806.

Wendel J.F., Schnabel A. and Seelanan T. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). Proc. Natl. Acad. Sci. USA 92: 280–284.

Yeh F.C. and Boyle T.J.B. 1997. Population genetic analysis of co-dominant and dominant markers and quantitiative traits. Belgian J. Bot. 129: 157.

Zawko G., Krauss S.L., Dixon K.W. and Sivasithamparam K. 2001. Conservation genetics of the rare and endangered *Leucopogon obtectus* (Ericaceae). Mol. Ecol. 10: 2389–2396.