

CRYPTIC INTERSPECIFIC INTROGRESSION AND GENETIC DIFFERENTIATION WITHIN *GOSSYPIUM ARIDUM* (MALVACEAE) AND ITS RELATIVES

INÉS ÁLVAREZ¹ AND JONATHAN F. WENDEL²

¹Real Jardín Botánico de Madrid, Consejo Superior de Investigaciones Científicas, Plaza de Murillo 2, E-28014 Madrid, Spain
E-mail: ines@ma-rjb.csic.es

²Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, Iowa 50011
E-mail: jfw@iastate.edu

Abstract.—Interspecific gene flow is increasingly recognized as an important evolutionary phenomenon in plants. A surprising observation is that historical introgression is often inferred between species that presently have geographic and reproductive barriers that would appear to prohibit the inferred sexual exchange. A striking example concerns *Gossypium aridum* (subsection *Erioxylum*); previous analyses have shown that populations from Colima (southwestern Mexico) have a chloroplast genome (cpDNA) similar to that of a different taxonomic subsection (*Integrifolia*) that presently is confined to Baja California and the Galapagos Islands, whereas other *G. aridum* populations share a cpDNA lineage with each other and with other species in subsection *Erioxylum*. To evaluate further the possibility that this cpDNA evidence reflects introgression as opposed to some other evolutionary process, as well as to explore patterns of genetic diversity and similarity in both subsections, we conducted amplified fragment-length polymorphism (AFLP) analysis using 50 populations representing all seven species in the two subsections. Genetic diversity is high in *G. aridum*, and is strongly correlated with geography, as are similarities among the five species in subsection *Erioxylum*. This subsection is genetically distant from the two species in subsection *Integrifolia*, whose populations are highly similar inter se. Populations of *G. aridum* from Colima are genetically distinct from the remainder of the species, and exhibit a comparatively high frequency of AFLP fragments that otherwise are diagnostic of the *Integrifolia* lineage. These data implicate intersubsectional introgression between presently allopatric and genetically isolated clades, giving rise to a morphologically cryptic, introgressant entity. Biogeographic considerations suggest that this history was initiated following migration of one or more seeds from Baja California to the Colima coast, perhaps during the Pleistocene. We suggest that cryptic and seemingly improbable interspecific introgression and molecular differentiation may be more common than appreciated in angiosperm evolution.

Key words.—Phylogeography, genetic diversity, introgression, molecular differentiation, amplified fragment-length polymorphism, *Gossypium aridum*.

Received April 1, 2005. Accepted January 10, 2006.

One important insight from comparative molecular phylogenetic analyses over the last 20 years is that interspecific gene flow and hybrid speciation are common phenomena in plants (Rieseberg and Soltis 1991; Rieseberg and Wendel 1993; Rieseberg et al. 1995; Arnold 1997; Rieseberg 1997; Cronn et al. 2003; Cronn and Wendel 2004; Hegarty and Hiscock 2005). Although only a small fraction of plant genera have been examined using multiple, independent phylogenetic datasets, incongruence among two or more molecular datasets has shown that hybridization has been even more prevalent than indicated by morphological and cytogenetic data. Cryptic historical introgression appears to be common in some well-studied genera, such as *Glycine* (Doyle et al. 2004) and *Gossypium* (Cronn and Wendel 2004), suggesting that interspecific hybridization is an important and creative evolutionary process (Rieseberg and Wendel 1993; Rieseberg et al. 1995, 2003; Arnold 1997; Rieseberg 1997; Hegarty and Hiscock 2005). Adaptive diversification and speciation due to hybridization has been theoretically argued (Barton 2001) and experimentally demonstrated in several genera, including *Helianthus* (Rieseberg et al. 2003), *Iris* (Arnold et al. 2004), and *Ipomopsis* (Campbell 2004), in which hybrid and introgressant genotypes may be more fit than their parents in novel habitats. A remarkable aspect of many recent studies is that interspecific gene flow has been inferred between species that presently show strong geographic or reproductive barriers that would appear to prohibit the inferred sexual exchange. These barriers include extreme geographic allopatry as well

as intrinsic barriers to mating and genetic exchange (Rieseberg and Wendel 1993; Rieseberg et al. 1995; Arnold 1997; Rieseberg 1997, 2003; Cronn et al. 2003; Cronn and Wendel 2004).

A case in point concerns the evolutionary history of subsection *Erioxylum*, a group of American *Gossypium* species centered in southern and western Mexico. Phylogenetic affinities among American diploid cottons (*Gossypium* L., subgenus *Houzingenia* Fryxell) have been studied using multiple datasets, including chloroplast DNA (cpDNA) restriction sites (DeJode 1992; Wendel and Albert 1992), cpDNA spacer and gene sequences (Cronn et al. 2003), nuclear ribosomal DNA (nrDNA) sequences (Cronn et al. 1996; Seelanan et al. 1997), and sequences of low-copy nuclear genes (Small and Wendel 2000; Cronn et al. 2003; Álvarez et al. 2005). All of these analyses support the monophyly of the subgenus (reviewed in Wendel and Cronn 2003) and the taxonomic recognition of six subsections (*Austroamericana*, *Caducibracteolata*, *Erioxylum*, *Houzingenia*, *Integrifolia*, and *Selera*). Notwithstanding these multiple analyses, relationships among subsections remain unresolved in some cases, in part reflecting rapid cladogenesis early in the history of this subgenus (Álvarez et al. 2005), the probable occurrence of multiple episodes of historical interspecific hybridization (Cronn et al. 2003; Cronn and Wendel 2004), and perhaps insufficient population sampling.

Among the 13 species of American diploid cottons, a group of four Mexican species (subsection *Erioxylum*) stand out

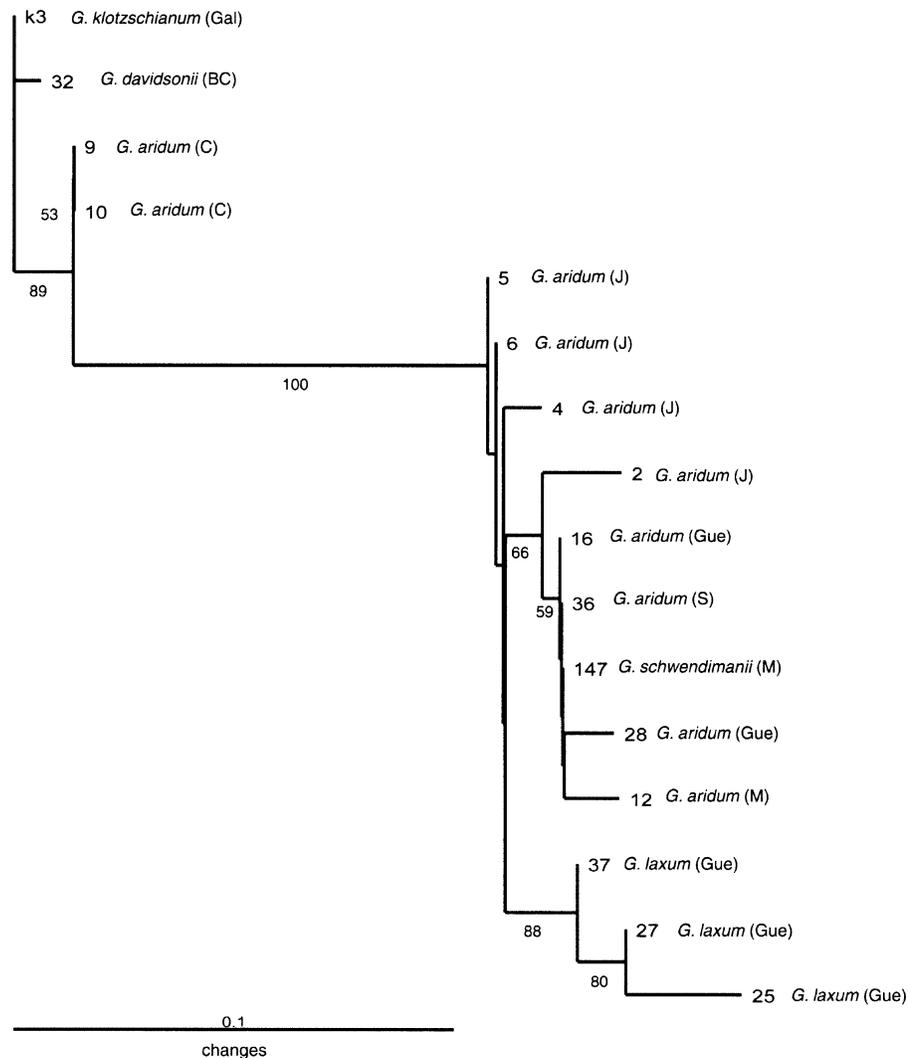


FIG. 1. Midpoint-rooted neighbor-joining tree using Nei-Li distances for chloroplast DNA restriction-site data (DeJode 1992) of individuals from 16 populations (two from Colima) included in the present work. Monomorphic sites and missing data were excluded, and a total of 28 markers was analyzed. Numbers in terminals correspond to population number in this study (see Appendix, available online). Geographic origin of populations is indicated within parentheses as follows: BC, Baja California; C, Colima; Gal, Galapagos Islands; Gue, Guerrero; J, Jalisco; M, Michoacan; S, Sinaloa. Bootstrap values $>50\%$ are shown below branches.

morphologically by virtue of their arborescent habit (trees up to about 20 m) as well as other characters. Previous work (Small and Wendel 2000; Álvarez et al. 2005) reported high levels of intraspecific variation in the four currently recognized species (*G. aridum*, *G. laxum*, *G. lobatum*, and *G. schwendimanii*) of the subsection, as well as an absence of AdhA allele coalescence within species (Small and Wendel 2000). This trans-specific polymorphism could reflect either interspecific gene flow, a lack of coalescence following speciation from a polymorphic ancestor, or both phenomena.

Earlier indications that *G. aridum* had an unconventional evolutionary history trace back more than a decade (Wendel and Albert 1992), when a single accession of *G. aridum* included in a phylogenetic analysis of cpDNA restriction-site variation was shown to share its cpDNA ancestry with members of the otherwise cladistically distant subsection *Integrifolia* (*G. davidsonii* and *G. klotzschianum*), rather than with the other Mexican arborescent species. These two subsections

are taxonomically well defined and morphologically distinct, as well as geographically allopatric. Subsection *Integrifolia*, from Baja California (*G. davidsonii*) and the Galapagos Islands (*G. klotzschianum*), is represented by shrubby plants with rotate yellow corollas, lacinate to dentate involucre bracts, and sparse seed pubescence. In contrast, members of subsection *Erioxylum*, broadly distributed in western Mexico exclusive of Baja California, are trees with funnelliform pink corollas, entire involucre bracts, and dense seed pubescence. Thus, the cpDNA data were surprising, and were hypothesized to reflect cytoplasmic introgression into *G. aridum* from a taxon in the subsection *Integrifolia* lineage. Further experiments involving the same species but deeper population sampling and both cpDNA restriction-site and isozyme analyses (DeJode 1992) confirmed the previous observation of putative cytoplasmic introgression from *Integrifolia* (Fig. 1), but added the twist that this introgression is evident only in accessions of *G. aridum* from the single Mexican state of

Colima, a small region located on the southwestern Mexican coast. All other *G. aridum* accessions sampled, from Guerrero, Jalisco, Michoacan, and Sinaloa, contain “normal” plastid genomes, cladistically monophyletic with those from all other *Erioxylum* species and populations.

Thus, *G. aridum* offers a striking example of interpopulational, intraspecific polymorphism in chloroplast genomes, raising the question as to how and when the polymorphism originated. Additionally, it is of interest to ask whether the putative hybridization event left footprints in the nuclear genome of *G. aridum* from Colima, and, if so, how extensive nuclear introgression might be, as well as its genomic distribution and composition. To date, evidence of nuclear introgression has not been observed, either in 5S rDNA genes (Cronn et al. 1996) or coding sequences from four independent nuclear loci (Small and Wendel 2000; Álvarez et al. 2005). However, these surveys collectively explore only a tiny portion (~7 kb) of the approximately 900 Mb nuclear genome (Hendrix and Stewart 2005) of these *Gossypium* species.

To gain further insight into the possibility of nuclear introgression in Colima populations of *G. aridum* and to explore genetic similarities with *G. aridum* and sympatric species, it is necessary to sample the genome more thoroughly. With these objectives in mind we employed amplified fragment-length polymorphism (AFLP) analysis (Vos et al. 1995). This commonly used DNA fingerprinting technique (reviewed in Mueller and Wolfenbarger 1999) has applicability to a wide array of evolutionary questions, ranging from studies of hybridization and allopolyploidy (Liu et al. 2001; Adams et al. 2004) to paternity analysis (Pertl et al. 2002; Chauhan et al. 2004) and questions of species circumscription and delimitation (e.g., Van Den Berg et al. 2002; Martínez-Ortega et al. 2004). In the present study we analyze AFLP patterns for a broad sampling of *G. aridum* populations, including from Colima and several populations of all species of subsection *Erioxylum* and subsection *Integrifolia*. We examine levels and patterns of genetic diversity as revealed by AFLP analysis, address the taxonomic and phylogeographic implications of these data, and explore the data for evidence of historical, intersubsectional, nuclear introgression.

MATERIALS AND METHODS

Plant Material

We studied 143 individuals from 50 populations of *Gossypium aridum* (Rose & Standley) Skovsted, *G. laxum* Phillips, *G. lobatum* Gentry, and *G. schwendimanii* (Fryxell & Koch) from subsection *Erioxylum* and of *G. davidsonii* Kellogg, *G. klotzschianum* Andersson from subsection *Integrifolia* (see Appendix, available online only at <http://dx.doi.org/10.1554/05-184.1.s1>). Plants were grown in the greenhouse at Iowa State University, starting with seed samples selected from collections made during previous collecting trips and those described elsewhere (DeJooe 1992; Wendel and Albert 1992; Wendel and Percival 1990), except for accessions 56-1 thru 66-5 (see online Appendix), which were provided by M. Ulloa and J. M. Stewart, and accessions 107-1, 118-1, 147-1, 157-1, 185-1, k3-1, and 32s-1, which were maintained as adult plants in the greenhouse. Sampling of *G.*

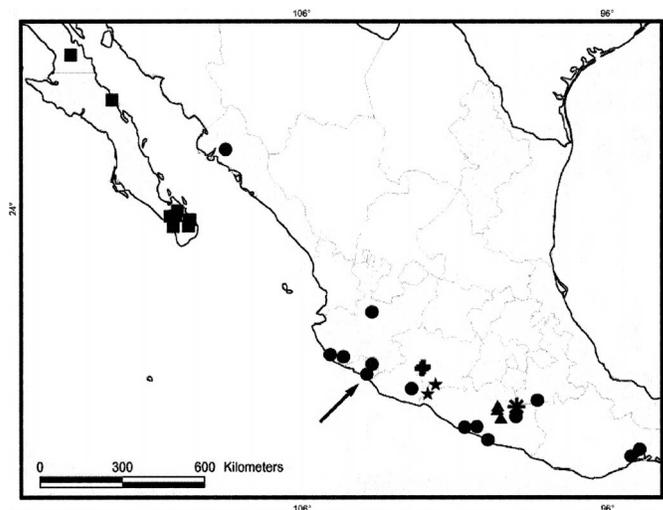


FIG. 2. Location of Mexican populations included in the study: *Gossypium aridum* (circles), *G. davidsonii* (squares), *G. laxum* (triangles), *G. lobatum* (crosses), *G. schwendimanii* (five-tip stars), *Gossypium* sp. (eight-tip stars). Colima populations are indicated by an arrow.

aridum was designed to encompass its range of distribution (see Fig. 2), but in addition all populations from Colima were included. Following seed scarification, only 34% of seeds germinated and survived, and thus the initial 85 populations selected were reduced to 50 as follows: *G. aridum* (24, of which four are from Colima), *G. davidsonii* (10), *G. klotzschianum* (4), *G. laxum* (7), *G. lobatum* (2), *G. schwendimanii* (2), and *Gossypium* sp. (1) (see online Appendix for details). One accession of a taxon suspected to represent an undescribed species (*Gossypium* sp.) also was included (acc. 64-1, US-72) (Álvarez et al. 2005; Ulloa et al. 2006). Vouchers for all individuals were deposited at the Ada Hayden Herbarium at Iowa State University, Ames.

AFLP Analysis

Fresh leaf tissue was used to isolate total DNA with the Plant DNeasy kit (Qiagen, Valencia, CA) following the manufacturer's instructions. Quality of isolated DNA was checked on 1% TAE-agarose gels. Genomic DNA (~200 ng) was digested with 10 units of EcoRI and 5 units of MseI, incubating at 37°C for 3 h. Double-stranded adaptors were prepared from the following complementary single-stranded oligonucleotides: 5' CTC GTA GAC TGC GTA CC 3' and 5' AAT TGG TAC GCA GTC 3' for the EcoRI adapter pair, and 5' GAC GAT GAG TCC TGA G 3' and 5' TAC TCA GGA CTC AT 3' for the MseI adapter pair. Ligation reactions were performed by adding 75 pmoles of the EcoRI adapter, 15 pmoles of the MseI adapter, and 20 units of T4 DNA ligase with its buffer to the digested product, and incubating overnight at 16°C. Ligation products were diluted by adding 160 μ l dH₂O.

For the preselective polymerase chain reaction (PCR), we added 10 μ l of the diluted ligation product to 40 μ l of a preselective PCR mix consisting of: 13 μ l dH₂O, 5 μ l 10 \times PCR buffer, 1.5 μ l MgCl₂ (50 mM), 4 μ l dNTP (2.5 mM), 8 μ l (5 pmol/ μ l) of each preselective primer, and 0.5 μ l of

TABLE 1. Primer combinations used in the selective PCR and total number of scored fragments obtained with each combination.

	Mse + CAA	Mse + CAC
EcoRI + AGC (6-FAM)	60	36
EcoRI + ACG (TET)	49	42
EcoRI + AAC (TET)	102	63
EcoRI + ACA (6-FAM)	77	66

Taq DNA polymerase (5 U/ μ l). Sequences of preselective primers are: EcoRI + A: 5' GAC TGC GTA CCA ATT CA 3', and MseI + C: 5' GAC GAT GAG TCC TGA GTA AC 3'. Preselective PCR conditions were a preliminary 75°C extension for 2 min followed by 20 cycles of 94°C for 30 sec, 56°C for 30 sec, 75°C for 2 min, finishing with 1 cycle of 60°C for 30 min. Ten μ l of this PCR product were electrophoresed through 1.5% TAE-agarose gels and stained with ethidium bromide to verify adequate preselective amplification. The remaining 40 μ l were diluted with 740 μ l dH₂O.

For the selective PCR, we added 5 μ l of diluted preselective PCR product to 20 μ l of the selective PCR mix consisting of: 11.5 μ l dH₂O, 2.5 μ l 10 \times PCR buffer, 0.75 μ l MgCl₂ (50 mM), 3 μ l dNTP (2.5 mM), 0.75 μ l (5 pmol/ μ l) each of two EcoRI labeled (6-FAM and TET) selective primers, 0.5 μ l (50 pmol/ μ l) of one MseI unlabeled primer, and 0.25 μ l of Taq DNA polymerase (5 U/ μ l). We performed selective PCR with eight primer pairs (see Table 1). The PCR profile was 1 cycle of 94°C for 2 min, 1 cycle of 94°C for 30 sec, 65°C for 30 sec, and 72°C for 2 min, followed by 9 cycles of a 1°C decrease in annealing temperature per cycle, followed by 35 cycles of 94°C for 30 min, 56°C for 30 sec, and 72°C for 2 min, and a final extension at 60°C for 30 min. For all samples we performed duplicates of preselective and selective reactions to verify reproducibility. Selective PCR products were electrophoretically separated using automated sequencing gels on an ABI Prism 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA) at the DNA Sequencing and Synthesis Facility at Iowa State University. See www.eeob.iastate.edu/faculty/WendelJ/afpl.htm for more details on our AFLP protocol.

Scoring AFLP Data

Gel images were analyzed with ABI GeneScan Analysis 3.0 and Genotyper 2.5 (Applied Biosystems) software. To

verify reproducibility, samples were run in a manner that permitted duplicates to be visualized side by side on each gel. GeneScan row data was transformed into a Microsoft Excel spreadsheet with the aid of the program Genescan Step 3 (R. J. Dyer, pers. comm.). Peaks between 100 and 500 bp in length were scored, and only those with 100% reproducibility in all replicate samples were included in the data matrix. To avoid ambiguities, we eliminated peaks that overlapped more than half with others, peaks with relative fluorescence units below 100, and other possible gel artifacts. Our guiding philosophy was to use only the most robust, repeatable, and unambiguous bands. A presence/absence matrix was constructed from the final data table.

Data Analysis

Genetic diversity was estimated based on percentage of polymorphic bands within each taxon under consideration. To test for nuclear introgression within populations of *G. aridum* from Colima, we treated the latter as a different entity where appropriate. Thus, 10 entities were defined a priori for all analyses: *G. aridum* from Colima, *G. aridum* outside of Colima, *G. davidsonii*, *G. laxum*, *G. lobatum*, *G. klotzschianum*, *G. schwendimanii*, *Gossypium* sp., subsection *Erioxylum*, and subsection *Integrifolia*. Genetic similarity (Dice coefficients) was calculated among and within entity (see Table 2). Polymorphic bands were classified into two main categories: exclusive (only present in one entity) and shared (present in more than one entity). We arbitrarily defined subcategories within these two main categories. Thus, we classified exclusive bands as diagnostic (present in 100% of individuals of all populations of the entity), potentially diagnostic (present in at least 25% of the populations in the entity), and rare (present in less than 25% of the populations in the entity). Similarly, shared bands were categorized either as diagnostic (present in at least 25% of the populations of one entity and in less than the 25% of populations of any other entity) or not diagnostic. Based on this classification, genetic composition of each entity was estimated (see Table 3). In an attempt to visualize possible nuclear introgression into Colima populations, we represented the AFLP band distribution (see Fig. 3) for all shared bands present in Colima populations, including bands diagnostic for subsections (*Erioxylum* and *Integrifolia*).

Analyses of molecular variance (AMOVA) at different tax-

TABLE 2. Dice similarity coefficients between and within entities. Abbreviations used are: ARI, *G. aridum* outside of Colima; ARIC, *G. aridum* from Colima; DAV, *G. davidsonii*; ERI, subsection *Erioxylum*; INT, subsection *Integrifolia*; KLO, *G. klotzschianum*; LAX, *G. laxum*; LOB, *G. lobatum*; NEW, *Gossypium* sp. (new species); SCH, *G. schwendimanii*.

	ARI	ARIC	LAX	LOB	NEW	SCH	ERI	DAV	KLO	INT
ARI	0.816									
ARIC	0.622	0.900								
LAX	0.665	0.628	0.860							
LOB	0.704	0.666	0.685	0.968						
NEW	0.706	0.651	0.819	0.740	1.000					
SCH	0.693	0.661	0.680	0.750	0.723	0.890				
ERI	0.739	0.673	0.706	0.711	0.728	0.695	0.718			
DAV	0.489	0.486	0.477	0.495	0.515	0.510	0.487	0.926		
KLO	0.478	0.479	0.462	0.487	0.504	0.500	0.476	0.923	0.958	
INT	0.487	0.484	0.474	0.494	0.513	0.508	0.485	0.926	0.930	0.927

TABLE 3. Distribution of diagnostic and potentially diagnostic fragments scored for each entity analyzed. Ni, number of individuals; Nf, number of fragments scored; Np, number of populations; p, percentage of polymorphic fragments within each entity; e, percentage of exclusive fragments; shC, percentage of fragments exclusively shared with Colima populations; shE, percentage of shared fragments diagnostic for subsection *Erioxylum*, shI, percentage of shared fragments diagnostic for subsection *Integrifolia*.

Entity	Ni	Np	Nf	p	e	shE	shC	shI
<i>G. aridum</i> exclusive of Colima	50	20	285	89.8	26.6	41.1	4.6	4.9
<i>G. aridum</i> from Colima	16	4	169	58	13.7	45.6	—	10.7
<i>G. davidsonii</i>	36	10	166	48.2	7.8	18.1	0.6	33.1
<i>G. klotzschianum</i>	9	4	134	27.6	0	9.7	0	41
<i>G. laxum</i>	22	7	201	80.1	20.9	46.5	1.5	3
<i>G. lobatum</i>	6	2	134	23.9	2.2	57.5	1.5	3
<i>G. schwendimanii</i>	3	2	159	50.3	9.5	51.6	1.2	4.5
<i>Gossypium</i> sp.	1	1	118	—	2.5	54.2	0.9	1.6
Subsection <i>Integrifolia</i>	45	14	166	—	39.7	10.8	3	—
Subsection <i>Erioxylum</i>	98	36	429	—	39.4	—	12.8	4.2

onomic and geographic hierarchical levels were conducted using ARLEQUIN 2.000 (Schneider et al. 2000). To estimate genetic similarities among AFLP phenotypes, we calculated a distance matrix based on Nei and Li's (1979) method and generated a neighbor-joining tree (NJ). To more readily visualize the two main groupings, the tree was rooted at its

midpoint. A bootstrap analysis with 1000 replicates was performed to estimate branch support. These analyses were conducted with PAUP* 4.0b10 (Swofford 1999). A principal coordinate analysis (PCoA) was performed to infer relative spatial locations of AFLP patterns based on Dice distances (Sneath and Sokal 1973), computed using NTSYS-pc 2.10t software (Rohlf 1998).

RESULTS

AFLP Patterns and Genetic Diversity

From the eight selective primer pairs (see Table 1) and 143 individuals, 495 fragments ranging from 100 to 500 bp were scored, of which 28 (13.9%) were monomorphic. The number of fragments per individual ranged from 129 to 180, with an average of 151 fragments. The highest averages of number of fragments per individual were found in *G. aridum* from Colima and *G. schwendimanii* (both 162), and in *G. laxum* (158), whereas the lowest averages were found in *G. klotzschianum* (135) and *G. davidsonii* (145). Four identical AFLP patterns were found for the following three populations of *G. davidsonii*: 30 (individuals 30-2, 30-4); 44 (individuals 44-1, 44-5); and 46 (individuals 46-1, 46-2, 46-5, 46-6).

The percentage of polymorphic AFLP bands within an entity serves as an estimator of diversity, although in our case this is only heuristic as it may be biased by differences in sample size, and in the case of the new species (*Gossypium* sp.) it is not applicable since only one individual is included. The highest percentage of polymorphic fragments was observed in *G. aridum* from outside Colima (89.8% over 20 pops.) and *G. laxum* (80.1% over seven pops.). More moderate values were obtained for *G. aridum* from Colima (58% over four pops.), *G. schwendimanii* (50.3% over two pops.), and *G. davidsonii* (48.2% over ten pops.). The lowest levels of polymorphism were found in *G. lobatum* (23.9% over two pops.) and in the Galapagos Island endemic *G. klotzschianum* (27.6% over four pops.), which previously has been shown to harbor limited genetic diversity (Wendel and Percival 1990). For a more realistic approach, sampling was standardized applying a rarefaction technique (Coart et al. 2005) with the aid of the program AFLPDIV available at <http://www.pierroton.inra.fr/genetics/labo/Software>. This technique is only applicable for entities with a minimum sample size of three populations with at least three individuals per population; thus,

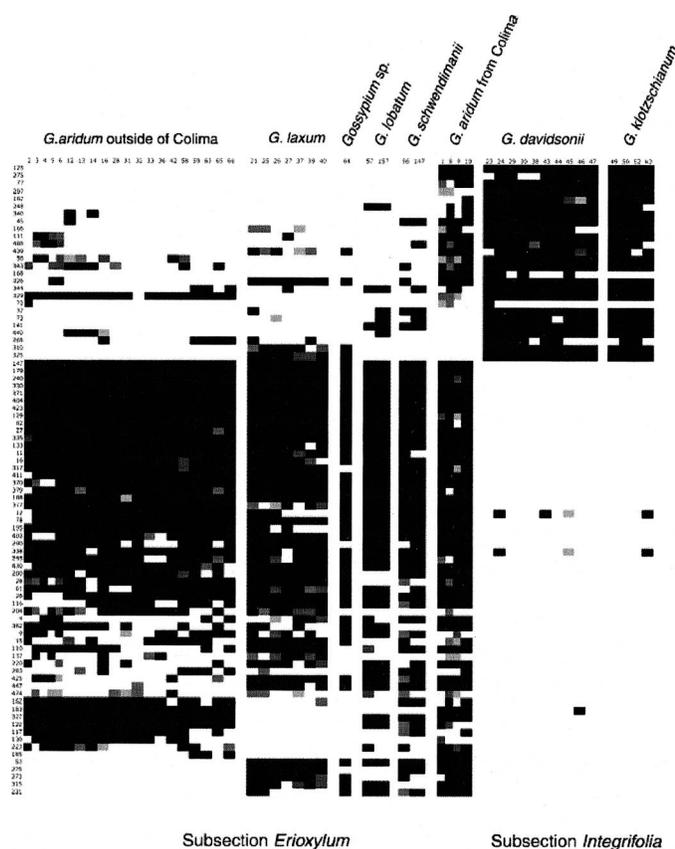


FIG. 3. Distributional representation of all fragments present in Colima populations of *Gossypium aridum* that are diagnostic or potentially diagnostic for one of the two subsections. Columns represent populations and rows correspond to specific amplified fragment-length polymorphism fragments. Black indicates presence in >75% of individuals in a population, dark gray indicates presence in 50–75%, medium gray indicate presence in 25–50%, light gray indicates presence in <25%, and white indicates absence in 100% of individuals of a population.

G. klotzschianum, *G. lobatum*, *G. schwendimanii*, and the new species were excluded from this estimation. In this case, the highest percentage of polymorphic loci was obtained for *G. aridum* from Colima (44.1%) and *G. laxum* (41.5%), while *G. davidsonii* and *G. aridum* from outside Colima present similar lower percentages (35.2% and 35.6%, respectively).

Averages of interspecific and intraspecific pairwise similarities using Dice coefficients are presented in Table 2. As expected, the highest values generally were obtained within entities, ranging from 0.718 to 1.000, and the lowest values were observed when entities from the two different subsections were compared, ranging from 0.474 to 0.515. The species most similar to each other are *G. davidsonii* and *G. klotzschianum*, which exhibit a similarity coefficient (0.923) that is comparable to those obtained among populations within species. Within subsection *Erioxylum*, the species pair with the highest similarity coefficient is *G. laxum*/*Gossypium* sp. (0.819). Notably, the pairwise comparison within species with the lowest similarity is *G. aridum* from Colima and *G. aridum* from elsewhere in the range of the species (0.622). These two entities, however, were equally similar to subsection *Integrifolia* taxa (range 0.478–0.484), a surprising result in light of other evidence for gene flow between the latter clade and *G. aridum* populations from Colima. This might be explained by the large number of exclusive versus shared fragments in Colima populations, which quantitatively overwhelms more subtle similarities evident in the AFLP data.

To explore for evidence of nuclear introgression, we tabulated the distribution of exclusive, diagnostic, and potentially diagnostic fragments and inspected the data for evidence of shared fragments between each entity and between subsections *Erioxylum* and *Integrifolia* and the Colima populations of *G. aridum* (see Table 3). The highest percentages of exclusive fragments are found in *G. aridum* from outside Colima and from Colima itself (26.6% and 13.7% respectively) and in *G. laxum* (20.9%), whereas the lowest percentages were from *G. klotzschianum* (no exclusive bands), *G. lobatum* (2.2%), and *Gossypium* sp. (2.5%). There were only two entities that contained fragments that were positively diagnostic (i.e., present in all individuals of all populations): *G. aridum* from Colima (six fragments; 3.6% of total), and *Gossypium* sp. (three fragments; 2.5% of total; data not shown). In addition, we unexpectedly found that all individuals of *G. aridum* from Oaxaca (five individuals, three pops.) contained nine AFLP fragments not detected elsewhere (5.9% of the total number of fragments in Oaxaca populations, and 3.1% of all *G. aridum* fragments; data not shown). Exclusive fragments for other populations of *G. aridum* were also found, but in each case these were few in number.

To address the question of possible intersubsectional nuclear introgression into *G. aridum* populations from Colima, we inspected for fragments that were shared exclusively between Colima populations and members of subsection *Integrifolia* (data not shown). For comparative purposes we also tabulated the proportion of shared fragments with other taxonomic entities. Not surprisingly, Colima populations share many diagnostic fragments with other *G. aridum* (7.7%), and, also as expected, Colima populations share a much smaller percentage of fragments with other species of Mexican ar-

borescent cottons (*G. laxum* and *G. lobatum*: 1.5% each; *G. schwendimanii*: 1.2%; *Gossypium* sp.: 0.9%). However, Colima populations share a surprisingly high percentage of fragments with the otherwise genetically distant *G. davidsonii* (3.6%) and *G. klotzschianum* (3.7%), which together share 7.1% of AFLP fragments that otherwise are diagnostic for this subsection exclusively with Colima populations of *G. aridum*.

An additional notable result with respect to the distribution of AFLP fragments concerns the new species, *Gossypium* sp. (Ulloa et al. 2006). This taxon shares 4.2% of its fragments exclusively with *G. laxum* and only 0.8% with *G. aridum*, suggesting high similarity to the former species.

Molecular Variance Analysis

Analyses of molecular variance (AMOVA) provide a convenient and powerful means to explore population similarities (see Table 4). All analyses performed were tested using non-parametric permutation procedures (Excoffier et al. 1992). When the data are structured taxonomically, treating each species as a group (structure B in Table 4), 56.4% of the total genetic diversity is attributed to variation among species, 35.3% to variation among populations within a species, and 8.3% to variation within populations. When, however, populations of *G. aridum* from Colima are treated as an entity distinct from the remainder of *G. aridum* (structure C in Table 4), the percentage of variation explained increases, with 65.4% of total variance attributable to interentity variation and 26% arising from among populations within entities. Comparable results are obtained when Colima populations are excluded from the analysis (structure D in Table 4); 65% and 27.4% of variation is explained by among- and within-species differences, respectively. In each of these analyses, there exists relatively less variation within populations (all <9%). To explore whether the increase in variation explained when Colima populations are excluded from *G. aridum* is unique to populations from that region, we performed AMOVA with structure B (see Table 4), sequentially removing or treating as separate entities populations of *G. aridum* from Guerrero, Jalisco, Michoacan, Oaxaca, Puebla, and Sinaloa (data not shown in Table 4). This analysis indicates that only populations from Colima have a notable effect on percentage of variation explained among groups (Fig. 4).

Analysis of variance also revealed that the AFLP data are geographically structured, an expected result given the allopatry among some of the genetically distinct species. Treating each region as a group (structure E in Table 4), 59.1% of the total genetic diversity is attributed to variation among regions, whereas 32.1% of variation arises from differences among populations within regions. Similar percentages were obtained when Colima populations were eliminated from the analysis (structure F in Table 4).

Several analyses that included only *G. aridum* were performed to explore population structure within this widespread species and the effect of Colima populations on its molecular variance distribution. When groups are defined by regions, including Colima populations (structure G in Table 4), the highest percentage of variation (51.6%) is accounted for by regions, followed by among populations within regions

TABLE 4. Results of molecular variance analyses (AMOVA) for different entities and geographic regions. Different structures to test are indicated by a capital letter in bold (A–K). df, degrees of freedom; PV, percentage of variation. Each group of the structure to test is included within parentheses: ARI, *G. aridum* from other than Colima; DAV, *G. davidsonii*; KLO, *G. klotzschianum*; LAX, *G. laxum*; LOB, *G. lobatum*; NEW, *Gossypium* sp. (new species); SCH, *G. schwendimanii*; ARICol, *G. aridum* from Colima; ARIGue, *G. aridum* from Guerrero; ARIJal, *G. aridum* from Jalisco; ARIMich, *G. aridum* from Michoacan; ARIOax, *G. aridum* from Oaxaca; ARIPue, *G. aridum* from Puebla; ARISin, *G. aridum* from Sinaloa; Baja, populations from Baja California; Col, populations from Colima; Gue, populations from Guerrero; Jal, populations from Jalisco; Mich, populations from Michoacan; Oax, populations from Oaxaca; Pue, populations from Puebla; Sin, populations from Sinaloa. In all cases $P < 0.001$.

Structure to test	Source of variation	df	PV
Each population as a group	among populations	49	90.1
A	within populations	93	9.9
(ARI + ARICol), (DAV), (KLO), (LAX), (LOB), (NEW), (SCH)	among entities	6	56.4
B	among populations within entities	43	35.3
	within populations	93	8.3
(ARI), (ARICol), (DAV), (KLO), (LAX), (LOB), (NEW), (SCH)	among entities	7	65.4
C	among populations within entities	42	26
	within populations	93	8.6
(ARI), (DAV), (KLO), (LAX), (LOB), (NEW), (SCH)	among entities	6	65
D	among populations within entities	39	27.4
	within populations	81	7.6
(Baja), (Col), (Gal), (Gue), (Jal), (Mich), (Oax), (Pue), (Sin)	among regions	8	59.1
E	among populations within regions	41	32.1
	within populations	93	8.9
(Baja), (Gal), (Gue), (Jal), (Mich), (Oax), (Pue), (Sin)	among regions	7	57.6
F	among populations within regions	38	34.5
	within populations	81	7.9
(ARIGue), (ARIJal), (ARIMich), (ARIOax), (ARIPue), (ARISin), (ARICol)	among regions	6	51.6
G	among populations within regions	18	35.3
	within populations	42	13.1
(ARIGue), (ARIJal), (ARIMich), (ARIOax), (ARIPue), (ARISin)	among regions	5	34.4
H	among populations within regions	14	52.5
	within populations	30	13.1
(ARI), (ARICol)	among entities	1	53.5
I	among populations within entities	22	36.7
	within populations	42	9.9
(ARI + ARICol)	among populations	23	85.4
J	within populations	42	14.6
(ARI)	among populations	19	85.6
K	within populations	30	14.5

(35.3%) and within populations (13.1%). In contrast, when Colima populations are excluded (structure H in Table 4), the results are notably different, with the majority of variation arising from differences among populations within regions (52.5%) as opposed to among regions (34.4%). When Colima populations are treated as a separate entity from the remainder of *G. aridum* (structure I in Table 4), 53.5% and 36.7% of total molecular variance arises from among-entity and within-entity variation, respectively; thus, over half of the variance arises from genetic divergence of Colima populations from the rest of the species. Finally, when each population of *G. aridum* is treated as its own group, the proportion of total variance arising from interpopulation differentiation is similar when Colima populations are either included (85.4%, structure J in Table 4) or excluded (85.6%, structure K in Table 4) from the analysis, with equal variance within populations in each case (14.6% and 14.5%, respectively).

Ordination and Cluster Analyses

To illustrate relative similarities among entities in multi-dimensional genetic space, we performed principal coordinate analysis (see Fig. 5). The first three principal coordinates collectively explain 61.5% of the variance and permit clear delineation of eight clusters. The first axis, which accounts

for 40.5% of the variance, clearly separates two groups corresponding to subsection *Integrifolia* (*G. davidsonii* and *G. klotzschianum*) and subsection *Erioxylum* (*G. aridum*, *G. laxum*, *G. lobatum*, *G. schwendimanii*, and *Gossypium* sp.). The second coordinate (11.9% of total variance) segregates clusters within subsection *Erioxylum*, with the largest group comprising populations of *G. aridum* from Guerrero, Jalisco, Michoacan, and Sinaloa, which collectively are very close to two populations of *G. aridum* from Puebla. *Gossypium aridum* populations from the southernmost part of the species range (from Oaxaca) are located in an intermediate position of the second axis. The most striking result, however, is the remarkable differentiation of populations of *G. aridum* from Colima, which ordinate at the opposite end of PCoA2 from the remainder of the species.

Additional similarities evident in the principal coordinate analysis include the high similarity between *G. lobatum* and *G. schwendimanii* (which are not wholly separated in this analysis), the evident separation of these two species from *G. laxum*, and the relative closeness of this latter species to the new and as yet unnamed species (*Gossypium* sp.). Also, the PCoA analysis fails to separate the closely related but geographically disjunct (Wendel and Percival 1990) species *G. davidsonii* and *G. klotzschianum*.

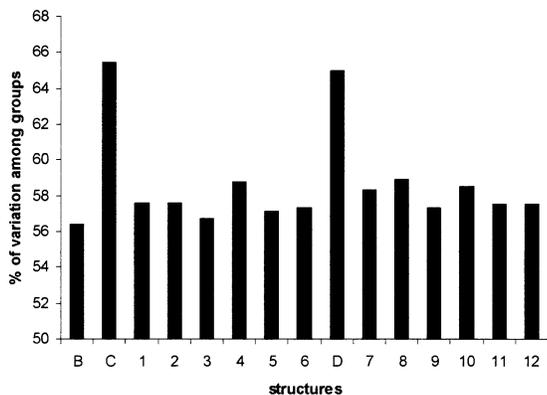


FIG. 4. Graphic representing the percentage of variation explained among groups for different structures tested. Structures are: B, C, and D (Table 4); 1, same as B but treating populations of *Gossypium aridum* from Guerrero as a different entity; 2, same as B but treating populations of *G. aridum* from Jalisco as a different entity; 3, same as B but treating populations of *G. aridum* from Michoacan as a different entity; 4, same as B but treating populations of *G. aridum* from Oaxaca as a different entity; 5, same as B but treating populations of *G. aridum* from Puebla as a different entity; 6, same as B but treating populations of *G. aridum* from Sinaloa as a different entity; 7, same as B but removing populations of *G. aridum* from Guerrero; 8, same as B but removing populations of *G. aridum* from Jalisco; 9, same as B but removing populations of *G. aridum* from Michoacan; 10, same as B but removing populations of *G. aridum* from Oaxaca; 11, same as B but removing populations of *G. aridum* from Puebla; 12, same as B but removing populations of *G. aridum* from Sinaloa.

As with the ordination analysis, cluster analysis (see Fig. 6) clearly reveals the main split between the two subsections *Integrifolia* and *Erioxylum*. Within subsection *Integrifolia*, populations of *G. davidsonii* and *G. klotzschianum*, from Baja California and Galapagos Islands, respectively, are intermingled, although some structure is evident in the figure. Within subsection *Erioxylum*, populations of *G. aridum* from Colima form a group clearly distinct from the remainder of the subsection, with high bootstrap support (100%). High support (100%) also exists for the *G. laxum*/new species grouping and for *G. laxum* by itself (87%), whereas the grouping of *G. lobatum* with *G. schwendimanii* attains modest bootstrap support (57%). Within *G. aridum* exclusive of Colima, populations from Oaxaca are clearly differentiated (100% bootstrap support). The *G. aridum* core group by itself has low bootstrap support, and although populations from different regions form clusters, there is not a clear overall geographic pattern. Within this core group, the main split separates populations from Sinaloa (88% bootstrap support), Puebla (100% bootstrap support) and some from Jalisco (no bootstrap support) from a second group that includes populations from Michoacan (100% bootstrap support), Guerrero (separated in two groups, each one with low bootstrap support), and Jalisco (100% bootstrap support).

DISCUSSION

Genetic Diversity and Similarities of Gossypium aridum and Its Allies

Gossypium aridum is the most morphologically variable American diploid cotton species, as well as the most geo-

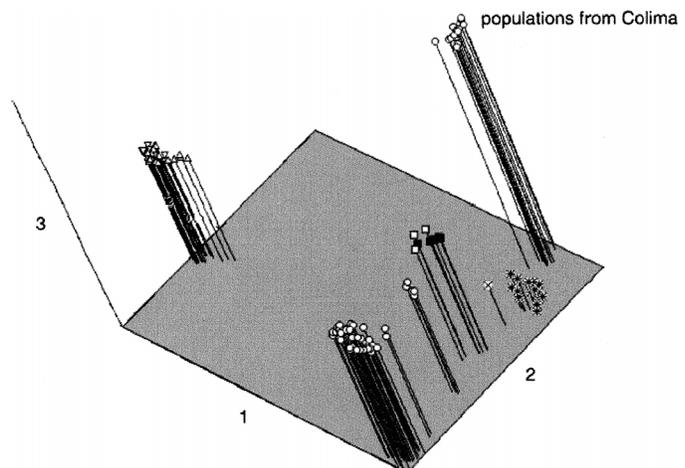


FIG. 5. Principal coordinate analysis based on Dice distances. *Gossypium aridum*, circles; *G. davidsonii*, triangles; *G. klotzschianum*, inverted triangles; *G. laxum*, stars; *G. lobatum*, solid squares; *G. schwendimanii*, open squares; and *Gossypium* sp., cross.

graphically widespread, extending from northwestern to southwestern Mexico with marginal populations reported from as distant as the Mexican Atlantic coast (e.g., from Veracruz). Its aggregate range overlaps the distributions of the remaining species of subsection *Erioxylum*, which are endemic to central Guerrero and Michoacan (*G. lobatum* and *G. laxum*) or are more locally distributed (e.g., *G. schwendimanii* is restricted to northern and eastern Infiernillo in Michoacan and Guerrero, and the new species of *Gossypium* is known only from one locality near the Balsas river in Guerrero; Ulloa et al. 2006). Species of subsection *Erioxylum* have long been thought to be closely related inter se, and molecular phylogenetic analyses focused on this subsection (Small and Wendel 2000) support this contention, as well as a scenario of relatively recent and perhaps rapid radiation. In addition, interspecific hybrids among species of subsection *Erioxylum* are frequently observed in the field (Ulloa et al. 2006), consistent with their occasional sympatry. Amplified fragment-length polymorphism data reported here further document the high level of genetic diversity within *G. aridum*, as revealed by the proportion of polymorphic fragments (see Table 3) and ordination (see Fig. 5) and clustering (see Fig. 6) results. Most notable is the striking divergence of the Colima populations from the remainder of the species (see Table 2) and the otherwise high intraspecific and geographically structured diversity within *G. aridum*. This helps explain why *G. aridum* is so morphologically and genetically diverse as well as why phylogenetic relationships among these closely related and sympatric species have been difficult to discern.

The AFLP data and analyses reported here highlight several features of the structure of genetic diversity in subsection *Erioxylum*. First, there is a clear correspondence between taxonomy and geography, presumably due to the endemic distribution of many taxa as well as some level of geographic partitioning within *G. aridum*. This is evidenced most clearly in the principal coordinate depiction (Fig. 5), although it also is apparent in the NJ tree (Fig. 6). For example, populations

might expect this, given the greater distribution of the former as well as morphological suggestions that this would be the case (see also Wendel and Percival 1990). However, in some more narrowly circumscribed regions (e.g., *G. laxum* and *G. aridum* from Colima) relatively high genetic diversity is observed. Historical patterns of hybridization and/or large long-term effective population sizes may account for high levels of genetic diversity within subsection *Erioxylum* species.

The geographic and taxonomic distribution of AFLP bands (presumed here to be mostly homologous, as opposed to the more unlikely scenario of multiple, spurious, homoplasious characters) lends some insight into evolutionary history and the genesis of present patterns, particularly when considered in light of the principal coordinate and NJ analyses. Perhaps the most noteworthy result is the genetic uniqueness and relative isolation of *G. aridum* from Colima, which displays the highest percentage of exclusive AFLP fragments, is strikingly divergent with respect to genetic similarity (Table 2), and clearly is isolated from the remainder of *G. aridum* in both ordination (Fig. 5) and clustering (Fig. 6) analyses. AMOVA quantifies this isolation of the Colima *G. aridum* populations, in that including Colima populations leads to much higher interpopulational differentiation than when these populations are excluded.

The genetic distinctiveness of Colima populations of *G. aridum* has not been morphologically evident. Recent observations in the field reported by Ulloa et al. (2006) indicate that all coastal, foothill populations of *G. aridum* from Jalisco, Colima, Michoacan, and Guerrero are morphologically similar, and only populations from Oaxaca (showing the largest leaves in the species with dense and fine indumentum) and populations of slopes and canyons from Colima (presenting the smallest leaves and capsules in the species) are macro-morphologically different from the remaining *G. aridum*. With respect to the latter observation, we did not find any characters indicative of morphological differentiation when accessions were grown in the greenhouse. Accordingly, and despite the evidence for some degree of molecular genetic differentiation, we do not consider the Colima populations to be a distinct species.

The AFLP data appear to be quite powerful in characterizing patterns of genetic diversity in subsection *Erioxylum*. In addition to the issue of the isolation of Colima *G. aridum* populations, *G. aridum* from Oaxaca exhibits a relatively high proportion of exclusive AFLP fragments and forms a group distinct from the remaining *G. aridum* populations (see Figs. 3, 4). Similarly, populations of *G. laxum*, which also have a high percentage of exclusive AFLP fragments, appear in clustering and ordination analyses to comprise a relatively isolated evolutionary/genetic unit. In contrast, some taxa are relatively weakly differentiated from congeners. For example, *G. lobatum* and *G. schwendimanii* are suggested to hybridize (Ulloa et al. 2006) based on observations in the field of intermediate morphologies and of sympatry; here, these two entities are revealed for the first time to be close relatives and perhaps sister species. Even more genetically intermingled are populations of *G. klotzschianum* and *G. davidsonii*. The former contains no exclusive AFLP fragments, and its populations interdigitate with those of its close relative *G. davidsonii* in both clustering and ordination analyses. This is

also reflected in the estimate of genetic similarity, which was 0.923 for the species pair (Table 2). These data are consistent with earlier suggestions, based on allozyme data, that these two species are not genetically well-differentiated (Wendel and Percival 1990).

A New Species of Gossypium

Included in the dataset was a single population of a taxon in subsection *Erioxylum* found to be molecularly divergent (Feng et al. 2003) in a preliminary report and considered to be an undescribed species (Ulloa et al. 2006). This putatively new species was first collected by M. Ulloa and J. M. Stewart, who noted that although its morphology in its defoliated condition resembles *G. aridum*, mature leaves become palmately lobed, as in *G. laxum*. Phylogenetic analysis of nuclear gene sequences supported the concept that this entity was a new species (Álvarez et al. 2005). The AFLP data presented here support the distinctive nature of this new species and lend additional support to the possibility that it shares a most recent common ancestor with *G. laxum*, with which it is genetically most similar (see Table 2), closest in multidimensional space (see Fig. 5), and closest in the NJ analysis (see Fig. 6). In addition, both species inhabit the same region (Guerrero), as do some populations of *G. aridum*, suggesting an additional possibility that the new taxon is of hybrid origin. We caution that the foregoing is based on a single collection of what apparently is a rare taxon, for which, hence, additional collections become a priority.

Intersubsectional Introgression into Colima Populations of Gossypium aridum

Inspection of shared and diagnostic AFLP polymorphisms indicate that the genomic composition of all species within subsection *Erioxylum* includes a mix of species-specific markers and, more commonly, fragments that are shared among two or more of the species. Rarer are instances where members of section *Erioxylum* share AFLP fragments that otherwise are diagnostic of subsection *Integrifolia* (1.6–4.9%; Table 3). This low level of shared AFLP fragments may reflect trans-specific polymorphisms arising from lack of coalescence, homoplasy among bands that are genetically different but of similar sizes, or possibly interspecific gene flow. However, from a comparative standpoint, the level of shared polymorphism between *G. aridum* from Colima and subsection *Integrifolia* becomes more remarkable, in that fully 10.7% of the AFLP fragments from Colima *G. aridum* are diagnostic of the otherwise genetically distant subsection *Integrifolia* (Table 3). Moreover, 3% of the AFLP polymorphisms from Colima are shared exclusively with subsection *Integrifolia* (data not shown), a situation not found for other subsection *Erioxylum* members (see Fig. 3 for details). The observation that Colima populations share more than a two-fold higher percentage of fragments with subsection *Integrifolia* than does any other subsection *Erioxylum* taxon may be taken as evidence of intersectional hybridization. We view lack of coalescence as an unlikely alternative to introgression because we see no reason why this phenomenon should effect the genomic composition of Colima populations more than that of any other *Erioxylum* species.

Additional support for the hypothesis of introgression comes from earlier analyses of chloroplast genomes (DeJode 1992; Wendel and Albert 1992), which demonstrated a shared cytoplasmic ancestry for *Integrifolia* and *G. aridum* populations from Colima (schematically illustrated in Fig. 1). However, previous phylogenetic analyses based on sequences of low-copy nuclear genes (Cronn et al. 1996; Small and Wendel 2000; Álvarez et al. 2005) have failed to confirm the existence of nuclear introgression, perhaps because sampling only three genes is unlikely to reveal low-level nuclear gene flow. In fact, if our estimation for the amount of introgression is true, only about 3% of the AFLP polymorphisms present in Colima populations are exclusively found in subsection *Integrifolia*, meaning a low probability (0.03) of finding introgressed genes randomly drawn genes. The distinctiveness of Colima populations as well as their position in ordination and clustering analyses (Figs. 5 and 6) might be interpreted as a result of introgression between *Integrifolia* lineage and some other unknown and presumably extinct species, but this is unlikely because the morphology of Colima populations lies clearly within the range of variation of *G. aridum* (Ulloa et al. 2006). In addition, gene sequence data for several independent nuclear markers support the inclusion of the Colima populations within *G. aridum* (Cronn et al. 1996; Small and Wendel 2000; Álvarez et al. 2005). A more plausible interpretation of the relative distinctiveness of the Colima populations is that they have experienced some level of local genetic local differentiation, much like other *G. aridum* from other regions, in addition to introgression from *Integrifolia*.

We note that fingerprinting techniques such as AFLP analysis sample both nuclear and chloroplast genomes, and hence the question naturally arises as to whether the introgression we infer here represents bona fide interspecific transfer of nuclear material or merely shared cpDNA fragments. Whereas definitive proof of the genomic origin of shared AFLP fragments would require isolation, cloning, and sequencing of the bands in question, a simple justification for the hypothesis of nuclear introgression emerges from algebraic considerations based on genome sizes (~900 Mbp for the nuclear genome and 0.15 Mbp for the chloroplast genome). With quasi-random sampling of these genomes by AFLP techniques, we expect to observe one chloroplast fragment for every 6000 nuclear fragments on AFLP gels. Our odds, therefore, of having observed any cpDNA fragments from among the 495 bands scored is low, and becomes vanishingly small for multiple, shared cpDNA fragments. A similar algebraic argument applies to mitochondrial DNA. Thus, we conclude that most of the Colima fragments shared exclusively (3%) with *Integrifolia* members, and those diagnostic for *Integrifolia* but present in Colima populations (10.7%) reflect introgression of nuclear material (excluding fragments that cryptically are homoplasious and those that are peculiarly symplesiomorphic to this comparison).

According to previous authors using phenetic and cladistic approaches (Small and Wendel 2000; Álvarez et al. 2005), speciation of American diploid cottons took place following separation of the Gulf of California from what is now mainland Mexico, estimated to have occurred around 6 to 12 million years ago (Larson et al. 1968; Lonsdale 1991; Ferrari et al. 1999; Ingle 2001). If this scenario is correct, introgression

occurred after these geological and cladistic events, and thus the cytoplasmic and nuclear gene introgression from the *Integrifolia* lineage (from Baja California) into what now are Colima cottons required dispersal of *G. davidsonii* from Baja California to the Colima coast in southwestern Mexico (an ocean voyage of ~750 km), followed by interspecific hybridization using the immigrant as the seed parent and *G. aridum* as the pollen parent. This would presumably have been followed by repeated backcrossing of the hybrid, as female, into the paternal lineage, ultimately generating a nearly pure but introgressant *G. aridum* nuclear genome (and morphology) but in an alien cytoplasm. This scenario gains credibility from the predilection for transoceanic dispersal observed in wild *Gossypium* (reviewed in Wendel and Cronn 2003), and from the many other remarkable stories of odds-defying interspecific gene flow in the genus (Cronn and Wendel 2004).

From a population genetic perspective, the foregoing scenario is rendered more plausible by the likelihood that only one or a few individuals of *G. davidsonii* (or its close but extinct relative) crossed the Gulf of California and became established on the Colima coast, and thus pollination would have been mainly or exclusively by *G. aridum*. At maturity, capsules of *G. aridum* dehisce to passively release seed that is dispersed by gravity; thus, long-term retention of a locally introgressed cytoplasm may have been promoted by passive, local seed dispersal, even in the face of continued paternal gene flow from *G. aridum*. The fact that sufficient evidence of this nuclear introgression persists until this day suggests that the process was initiated relatively recently, perhaps during the Pleistocene following formation of the Gulf of California. In addition, the persistence of foreign genomic material may have been promoted by locally small population sizes, long generation times associated with the tree habit and/or, possibly, natural selection favoring introgressed genes or gene interactions. The fact that typical (nonintrogressant) populations of *G. aridum* have not been found in Colima provides correlative support for the possibility of selective maintenance of introgressed genes in these populations. We note that extensive de novo genomic variation has been demonstrated to be provoked by intergeneric hybridization in rice (Wang et al. 2005), as has extensive chromosomal rearrangement in hybrid sunflowers (Lai et al. 2005). Thus, although in many cases mutations and/or changes in chromosomal patterns may be disadvantageous for the hybrid, they might also play an important role for either the transfer or evolution of novel traits related to adaptation and speciation in natural populations.

Among the many unanswered questions raised by this story are those concerned with timing and the dynamics of the introgressive hybridization event, about which at this point we can only speculate. However, it is experimentally feasible to explore the genomic distribution of the introgressed material, a possibility made realistic by the existence of high-density genetic maps (Rong et al. 2004) for diploid and polyploid cotton. It also would be of interest to inventory the introgressant genes and genetic regions, through judicious cloning and sequencing approaches. Such an analysis serves as a necessary prelude to any attempt to appreciate the potential adaptive relevance of this interspecific introgression.

Finally, we note an important implication of the *G. aridum* history, namely, that cryptic and seemingly improbable interspecific introgression may be a more important process in angiosperm evolution than presently realized (Rieseberg and Wendel 1993; Rieseberg et al. 1995, 2003; Rieseberg 1997; Cronn et al. 2003; Cronn and Wendel 2004).

ACKNOWLEDGMENTS

We thank D. DeJoode, E. Percival, J. M. Stewart and M. Ulloa for seed collections; R. Cronn, D. DeJoode, and J. M. Stewart for their comments on the manuscript; C. Aedo for map design; J. Fuertes and G. Nieto for help with the ARLEQUIN and NTSYS programs and for comments; R. Percival for technical assistance; and J. I. Yagüe for his help in the greenhouse. Research support was provided by the U.S. National Science Foundation and by the Spanish Ministry of Science, Education, and Sports.

LITERATURE CITED

- Adams, K. L., R. Percival, and J. F. Wendel. 2004. Organ-specific silencing of duplicated genes in a newly synthesized cotton allotetraploid. *Genetics* 168:2217–2226.
- Álvarez, I., R. Cronn, and J. F. Wendel. 2005. Phylogeny of the New World diploid cottons (*Gossypium* L., Malvaceae) based on sequences of three low-copy nuclear genes. *Plant Syst. Evol.* 252:199–214.
- Arnold, M. L. 1997. Natural hybridization and evolution. Oxford Univ. Press, New York.
- Arnold, M. L., A. C. Bouck, and R. S. Cornman. 2004. Verne Grant and Louisiana irises: Is there anything new under the sun? *New Phytol.* 161:143–149.
- Barton, N. H. 2001. The role of hybridization in evolution. *Mol. Ecol.* 10:551–568.
- Campbell, D. R. 2004. Natural selection in *Ipomopsis* hybrid zones: implications for ecological speciation. *New Phytol.* 161:83–90.
- Chauhan, N., M. S. Negi, V. Sabharwal, D. K. Khurana, and M. Lakshmikumaran. 2004. Screening interspecific hybrids of *Populus* (*P. ciliata* × *maximowiczii*) using AFLP markers. *Theor. Appl. Genet.* 108:951–957.
- Coart, E., S. Van Glabeke, R. J. Petit, E. Van Bockstaele, I. Roldán-Ruiz. 2005. Range wide versus local patterns of genetic diversity in hornbeam (*Carpinus betulus* L.). *Conserv. Genet.* 6:259–273.
- Cronn, R., and J. F. Wendel. 2004. Cryptic trysts, genomic mergers, and plant speciation. *New Phytol.* 161:133–142.
- Cronn, R. C., X. Zhao, A. H. Paterson, and J. F. Wendel. 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.
- Cronn, R., R. L. Small, T. Haselkorn, and J. F. Wendel. 2003. Cryptic repeated genomic recombination during speciation in *Gossypium gossypoides*. *Evolution* 57:2475–2489.
- DeJoode, D. R. 1992. Molecular insights into speciation in the genus *Gossypium* L. (Malvaceae). M.S. thesis, Iowa State University, Ames.
- Doyle, J. J., J. L. Doyle, J. T. Rauscher, and A. H. D. Brown. 2004. Diploid and polyploid reticulate evolution throughout the history of the perennial soybeans (*Glycine* subgenus *Glycine*). *New Phytol.* 161:121–132.
- Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial-DNA restriction data. *Genetics* 131:479–491.
- Feng, C.-D., J. M. Stewart, M. Ulloa, E. A. García, and S. Godoy. 2003. Genetic diversity among arborescent *Gossypium* species revealed by RAPD and AFLP. Proceedings of the Beltwide Cotton Conferences. National Cotton Council of America. Memphis, TN.
- Ferrari, L., M. López-Martínez, G. Aguirre-Díaz, and G. Carrasco-Núñez. 1999. Space-time patterns of Cenozoic arc volcanism in central Mexico: from the Sierra Madre Occidental to the Mexican Volcanic Belt. *Geology* 27:303–306.
- Hegarty, M. J., and S. J. Hiscock. 2005. Hybrid speciation in plants: new insights from molecular studies. *New Phytol.* 165:411–423.
- Hendrix, B. L., and J. M. Stewart. 2005. Estimation of the nuclear DNA content of *Gossypium* species. *Ann. Bot.* 95:789–797.
- Ingle, J. 2001. History of the Gulf of California: evidence from sediments, subsidence, and Steinbeck. 340th Peninsula Geological Society Meeting, Stanford, CA, 13 November 2001. Abstract available at <http://www.diggles.com/pgs/2001/PGS01-11.htm>.
- Lai, Z., T. Nakazato, M. Salmaso, J. M. Burke, S. Tang, S. J. Knapp, and L. H. Rieseberg. 2005. Extensive chromosomal repatterning and the evolution of sterility barriers in hybrid sunflower species. *Genetics* 171:291–303.
- Larson, R. L., H. W. Menard, and S. M. Smith. 1968. Gulf of California: a result of ocean-floor spreading and transform faulting. *Science* 161:781.
- Liu, B., C. L. Brubaker, G. Mergeai, R. C. Cronn, and J. F. Wendel. 2001. Polyploid formation in cotton is not accompanied by rapid genomic changes. *Genome* 44:321–330.
- Lonsdale, P. 1991. Structural pattern of the Pacific floor offshore peninsula California. Pp. 87–125 in J. P. Dauphin and B. R. T. Simoneit, eds. The Gulf and the peninsular province of the Californias. Memoir. American Association of Petroleum Geologists, Alexandria, VA.
- Martínez-Ortega, M. M., L. Delgado, D. C. Albach, J. A. Elena-Rosselló, and E. Rico. 2004. Species boundaries and phylogeographic patterns in cryptic taxa inferred from AFLP markers: *Veronica* subgen. *Pentasepalae* (Scrophulariaceae) in the Western Mediterranean. *Syst. Bot.* 29:965–986.
- Mueller, U. G., and L. L. Wolfenbarger. 1999. AFLP genotyping and fingerprinting. *Trends Ecol. Evol.* 14:389–394.
- Nei, M., and W. H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* 76:5269–5273.
- Pertl, M., T. P. Hauser, C. Damgaard, and R. B. Jorgensen. 2002. Male fitness of oilseed rape (*Brassica napus*), weedy *B. rapa* and their F(1) hybrids when pollinating *B. rapa* seeds. *Heredity* 89:212–218.
- Rieseberg, L. H. 1997. Hybrid origins of plant species. *Annu. Rev. Ecol. Syst.* 28:359–389.
- Rieseberg, L. H., and D. E. Soltis. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evol. Trend Plant* 5:65–84.
- Rieseberg, L. H., and J. F. Wendel. 1993. Introgression and its consequences in plants. Pp. 70–109 in R. Harrison, ed. Hybrid zones and the evolutionary process. Oxford Univ. Press, New York.
- Rieseberg, L. H., C. Van Fossen, and A. M. Desrochers. 1995. Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature* 375:313–316.
- Rieseberg, L. H., O. Raymond, D. M. Rosenthal, Z. Lai, K. Livingstone, T. Nakazato, J. L. Durphy, A. E. Schwarzbach, L. A. Donovan, and C. Lexer. 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* 301:1211–1216.
- Rohlf, F. J. 1998. NTSYS-PC numerical taxonomy and multivariate analysis system. Ver. 2.0. Exeter Publications, Setauket, N.Y.
- Rong, J., C. Abbey, J. E. Bowers, C. L. Brubaker, C. Chang, P. W. Chee, T. A. Delmonte, X. Ding, J. J. Garza, B. S. Marler, and many others. 2004. A 3347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (*Gossypium*). *Genetics* 166:389–417.
- Schneider, S., D. Roessli, and L. Excoffier. 2000. ARLEQUIN. A software for population genetics data analysis. Ver. 2.000. Genetics and Biometry Laboratory, University of Geneva, Geneva. Available via <http://anthro.unige.ch/arlequin>.
- Seelanan, T., A. Schnabel, and J. F. Wendel. 1997. Congruence and consensus in the cotton tribe (Malvaceae). *Syst. Bot.* 22:259–290.
- Small, R. L., and J. F. Wendel. 2000. Phylogeny, duplication, and intraspecific variation of Adh sequences in New World diploid

- cottons (*Gossypium* L., Malvaceae). *Mol. Phylogenet. Evol.* 16: 73–84.
- Sneath, P. H. A., and R. R. Sokal. 1973. *The principles and practice of numerical classification*. W. H. Freeman, San Francisco, CA.
- Swofford, D. L. 1999. PAUP*. *Phylogenetic analysis using parsimony (*and other methods)*. Ver. 4.02b. Sinauer, Sunderland, MA.
- Ulloa, M., J. M. Stewart, E. A. García-C., S. Godoy-A., A. Gaytan-M., and S. Acosta-N. 2006. Cotton genetic resources in the western states of Mexico: in situ conservation status and germplasm collection for ex situ preservation. *Genet. Res. Crop Evol.* doi: 10.1007/s10722-004-2988-0.
- Van Den Berg, G., J. Bryan, A. Del Rio, and M. Spooner. 2002. Reduction of species in the wild potato *Solanum* section *Petota* series *Longipedicellata*: AFLP, RAPD and chloroplast SSR data. *Theor. Appl. Genet.* 105:1109–1114.
- Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23:4407–4414.
- Wang, Y. M., Z. Y. Dong, Z. J. Zhang, X. Y. Lin, Y. Shen, D. Zhou, and B. Liu. 2005. Extensive de novo genomic variation in rice induced by introgression from wild rice (*Zizania latifolia* Griseb.). *Genetics* 170:1945–1956.
- Wendel, J. F., and V. A. Albert. 1992. Phylogenetics of the cotton genus (*Gossypium*): character-state weighted parsimony analysis of chloroplast-DNA restriction site data and its systematic and biogeographic implications. *Syst. Bot.* 17:115–143.
- Wendel, J. F., and R. C. Cronn. 2003. Polyploidy and the evolutionary history of cotton. *Adv. Agron.* 78:139–186.
- Wendel, J. F., and A. E. Percival. 1990. Molecular divergence in the Galapagos Island-Baja California species pair, *Gossypium klotzschianum* and *G. davidsonii* (Malvaceae). *Plant Syst. Evol.* 171:99–115.

Corresponding Editor: P. Soltis