

## GENETIC ANALYSIS OF A BROAD HYBRID ZONE IN *AESCULUS* (SAPINDACEAE): IS THERE EVIDENCE OF LONG-DISTANCE POLLEN DISPERSAL?

David T. Thomas,\* Adjoa R. Ahedor,† Charles F. Williams,‡ Claude dePamphilis,§  
Daniel J. Crawford,|| and Qiu-Yun (Jenny) Xiang<sup>1,\*</sup>

\*Department of Plant Biology, North Carolina State University, Raleigh, North Carolina 27695-7612, U.S.A.; †Department of Botany and Microbiology, University of Oklahoma, Norman, Oklahoma 73019, U.S.A.; ‡Department of Biological Sciences, Idaho State University, Pocatello, Idaho 83209-8007, U.S.A.; §Department of Biology and Institute of Molecular Evolutionary Genetics, Penn State University, University Park, Pennsylvania 16801, U.S.A.; and ||Department of Ecology and Evolutionary Biology and Natural History Museum and Biodiversity Research Center, University of Kansas, Lawrence, Kansas 66045-2106, U.S.A.

The genetic structure of a broad hybrid zone involving three hybridizing *Aesculus* species, *Aesculus flava*, *Aesculus pavia*, and *Aesculus sylvatica*, was examined. The objectives were to assess genetic variability, to test previously reported hypotheses on patterns of gene flow, and to infer the genetic structure and evolutionary processes in the hybrid zone. Samples from 24 populations within parental ranges and the hybrid zone were analyzed for variation at microsatellite and intersimple sequence repeat loci. The results indicated that genetic variability was similar among parental and hybrid populations, indicating no evident increased diversity in the hybrid zone. Most hybrid individuals were genetically more similar to *A. sylvatica* than to the other two species, and the overall genetic structure of the hybrid zone is asymmetrically biased toward *A. sylvatica*. Our analyses supported occasional recurrent long-distant gene flow from *A. pavia* and frequent gene flow from *A. sylvatica* into the hybrid zone, agreeing with results of a previous allozyme study. Collectively, the data from our study and previous allozyme and chloroplast DNA studies indicate that both historical localized gene flow and recurrent long-distant gene flow have contributed to the existence of the hybrid zone, that is, its origin via historical localized gene flow, while its maintenance involves ongoing long-distance pollen dispersal.

**Keywords:** *Aesculus*, hybrid zone, ISSR, long-distance gene flow, microsatellites.

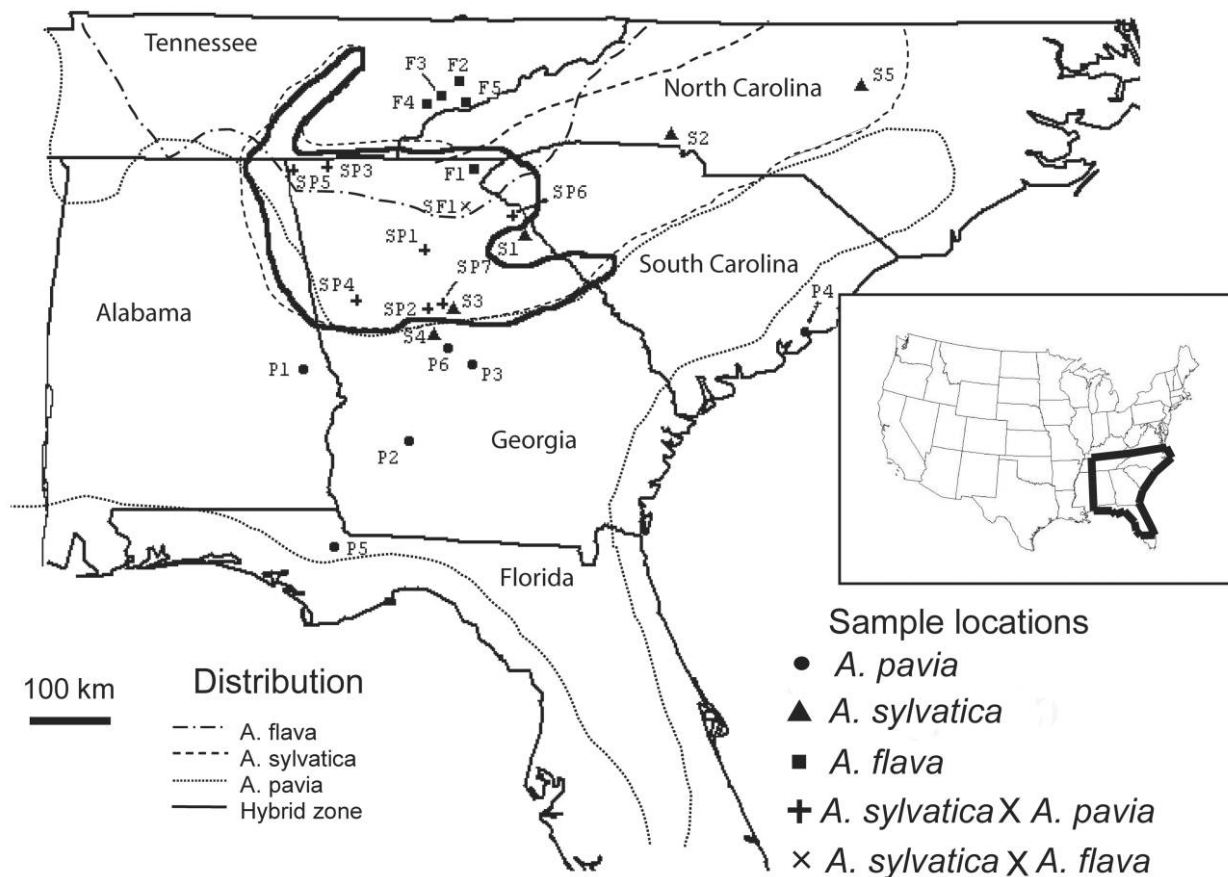
### Introduction

The prevalence and evolutionary significance of natural hybridization has long been debated. In the early to mid-twentieth century, prominent plant evolutionists accepted it as a significant evolutionary process, while those studying animals saw hybridization as functioning to reinforce reproductive isolation (Arnold 1997, chap. 1). The basic difference in these two opposing views of hybridization is that botanists (e.g., Anderson and Stebbins 1954) thought that it could produce novelties that would be more fit than their parents in certain habitats, whereas zoologists (e.g., Mayr 1942) saw hybrids as less fit. Not all prominent plant evolutionists, however, saw hybridization as a creative force, some referring to it as “evolutionary noise” (Wagner 1969, 1970). Spurred in part by the availability of molecular markers for measuring gene flow, there has been recent renewed interest in natural hybridization as a dynamic process (Rieseberg and Wendel 1993; Rieseberg 1995, 1997, Barton 2001; Rieseberg et al. 2003). Investigations of the evolutionary significance of hybridization have focused on hybrid zones as model systems for studying diversification and speciation (Arnold 1997; Campbell and Waser 2001; Campbell 2003). The dispersal ability of parental species and natural selection play significant roles in the origin,

maintenance, and dynamics of hybrid zones (Barton and Gale 1993; Harrison 1993), and the fitness of hybrid derivatives and their parental species in hybrid zones are of particular interest (Arnold and Hodges 1995; Rieseberg 1995; Arnold 1997). Studies have demonstrated that hybrid derivatives may be more fit than their parents in certain environments (e.g., Campbell and Waser 2007), that derivatives from a single hybridization event can produce hybrids more or less fit than their parents (e.g., Johansen-Morris and Latta 2006), and that long-term studies of hybrid zones provide refined estimates of hybrid fitness (e.g., Miglia et al. 2005). These and other studies demonstrate that hybrid zones can generate plants more fit than their parents and could influence the dynamics of hybridizing lineages by facilitating the movement of novel genotypes among hybrids and parents, potentially leading to the replacement of parental species or the formation of new species (Arnold 1993; James and Abbott 2005).

The large hybrid zones of the buckeye genus *Aesculus* L. represent a good example of hybridization and introgression in plants (Hardin 1957b, 1957c; dePamphilis and Wyatt 1989, 1990; Modliszewski et al. 2006). A broad hybrid zone (>200 km) is located in the southeastern United States and involves three parapatric species of sect. *Pavia*, *Aesculus flava* Ait. of the Appalachian Mountains, *Aesculus pavia* L. of the southern Coastal Plain, and *Aesculus sylvatica* Bartr. of the Piedmont (fig. 1). The age of the hybrid zone is unknown, although it has been suggested to have been formed during or after the

<sup>1</sup> Author for correspondence; e-mail: jenny\_xiang@ncsu.edu.



**Fig. 1** Current geographical ranges of *Aesculus flava*, *Aesculus sylvatica*, and *Aesculus pavia* in the southeast United States as indicated by Hardin (1957c). The location of the hybrid zone in northern Georgia is encircled with a solid line, and approximate locations of sample populations analyzed in this study are indicated.

Pleistocene (<1.8 Myr ago; Hardin 1957b, 1957c; see also Modliszewski et al. 2006). Species in sect. *Pavia* are clearly delimited by flower and fruit morphological characters (Hardin 1957a). Hardin (1957b, 1957c) studied hybridization among species within this zone using morphological characters. More recently, dePamphilis and Wyatt (1989, 1990) employed allozymes, crossing experiments, pollination biology studies, and pollen viability experiments via in vitro germination of pollen and stainability tests to gain additional insights into the zone's shape and structure; they suggested that this is a zone of secondary contact. Plants from the hybrid zone have more meiotic irregularities during microsporogenesis and pollen fertility is lower than in "pure" taxa, suggesting that these hybrids are less fit than their parents (Hardin 1957b; dePamphilis and Wyatt 1989).

The studies of Hardin (1957b, 1957c) and dePamphilis and Wyatt (1989, 1990) elucidated several interesting aspects of the hybrid zone. The distribution of hybrids between *A. pavia* and *A. sylvatica* is highly asymmetrical. Nearly all hybrids occur beyond the current range of *A. pavia*, which occurs on flat flood plains of rivers and streams or at the edge of swamps and marshes in the Coastal Plain. Hybrids extend up to 145 km into the range of *A. sylvatica*, which occurs on rocky hill-sides and more frequently on well-drained slopes of the Pied-

mont. The border between the distributions of *A. pavia* and *A. sylvatica* populations occurs along the Fall Line, a boundary between the Coastal Plain and the Piedmont physiographic regions.

Hardin (1957c) proposed that the geographic distribution of *A. pavia* once overlapped that of *A. sylvatica*, which allowed for localized gene flow between species and the initial formation of the hybrid zone. In subsequent years, *A. pavia* disappeared from the region and is no longer found within the boundaries of the hybrid zone. This hypothesis is supported by the leading-edge theory (Hewitt 1993), in which climatic changes during the Pleistocene restricted survival of selected species to localized warmer areas or refugia. The surviving populations might have subsequently expanded their range from the refugia during an interglacial warm period, resulting in the range overlap of species, which could have led to the formation of hybrid zones. In some cases, the succeeding environmental conditions might have selectively favored hybrids over parentals, resulting in hybrid zones lacking one or both parental species (Hewitt 1996). The reduction of *A. pavia* from the Piedmont could also have been the result of climatic cooling during the last glacial maximum (>20,000 yr ago), which may have restricted its range to warmer regions to the south and coastal areas (Hewitt 1996).

DePamphilis and Wyatt (1990) documented that allozyme alleles from *A. flava* and *A. pavia* extend into much of the range of the hybrid zone and into the range of *A. sylvatica* in northern Georgia. They confirmed that the genetic structure of the zone is also asymmetrical (biased toward *A. sylvatica*). To explain the presence of *A. pavia* genes in the hybrid zone, they proposed that recurrent long-distance dispersal of pollen occurs from the coastal *A. pavia* into *A. sylvatica* of the Piedmont, potentially mediated by ruby-throated hummingbirds (dePamphilis and Wyatt 1989, 1990).

The two previously proposed hypotheses, historical hybrid events followed by local extinction of *A. pavia* (historical secondary contact [HSC]) and long-distance pollen dispersal of *A. pavia* (long-distance gene flow [LDF]) are not necessarily mutually exclusive and therefore allow for a third hypothesis that incorporates both of these mechanisms in the origin and ongoing maintenance of the hybrid zone. These hypotheses can be further evaluated by analyzing the genetic composition and gene flow among parental and hybrid populations using chloroplast and more variable nuclear markers. Modliszewski et al. (2006) analyzed three chloroplast DNA (cpDNA) regions from populations within and surrounding the *Aesculus* hybrid zone and found little gene flow via seeds among populations, yet they detected the presence of *A. pavia* cpDNA haplotypes (confirmed to be maternally inherited in Modliszewski et al. 2006) in the hybrid zone, supporting the HSC hypothesis. These results, in combination with the allozyme data from dePamphilis and Wyatt (1990), present a scenario that suggests that the hybrid zone was initially formed through historical localized gene flow but is maintained by ongoing long-distance pollen dispersal (Modliszewski et al. 2006).

Although the cpDNA data provide evidence for historical localized seed-mediated gene flow in the hybrid zone, the evidence supporting LDF from the allozyme study (dePamphilis and Wyatt 1990) is limited by the low variability of the markers and the lack of species-specific alleles detected in the parental taxa. Therefore, additional data from more variable nuclear markers may provide additional information to further characterize the genetic structure of the hybrid zone and to test the LDF hypothesis.

If hybridization occurred in situ in the past, followed by local extinction of *A. pavia* sometime during the last glacial maximum, the proportion of nuclear *A. pavia* genes should be extremely low and relatively consistent within hybrid populations compared to *A. sylvatica* genes, assuming that *A. pavia* has been isolated from the hybrid zone since the last deglaciation (~18,000 yr; Prentice et al. 1991). A conservative estimate of the generation time of *Aesculus* based on nursery information is 10 yr; it takes about 3 yr for a seedling to mature, bloom, and produce seeds. Thus, under a "pure" HSC scenario, we would expect to detect a very low proportion of nuclear *A. pavia* genes in hybrid individuals. A significantly higher variance in genetic differentiation between parental and hybrid populations would also be expected if hybridization initially occurred historically and was followed by a long period of restricted gene flow (Latta and Mitton 1999; Cruzan 2005).

In contrast, hybridization occurring exclusively by recurrent long-distance pollen dispersal should result in only a few individuals containing high proportions of nuclear *A. pavia*

genes in a population, since long-distance dispersal is expected to be limited and sporadic. Most hybrid individuals should have a low proportion of *A. pavia* genes. Furthermore, high levels of positive disequilibrium in hybrid populations would also be detected if contemporary LDF occurs from parental populations (Cruzan 2005). Combined evidence from previous cpDNA and allozyme studies has suggested that both processes might have contributed to the current genetic structure of the *Aesculus* hybrid zone (dePamphilis and Wyatt 1990; Modliszewski et al. 2006). If both mechanisms were directly involved, we would expect to see outcomes from both processes, i.e., hybrid individuals with a low residual proportion of *A. pavia* alleles yet few hybrid individuals with a high proportion of *A. pavia* alleles. Also, a low variance of genetic differentiation between parental and hybrid populations and high levels of positive disequilibrium in hybrid populations would be expected if long-distance gene flow has not been restricted since the initial formation of the hybrid zone. Therefore, the goals of this study are to estimate the population genetic composition of the hybrid zone with two additional sets of more variable nuclear molecular data, microsatellite and intersimple sequence repeat (ISSR) markers, and to evaluate whether the genetic structure expected under a combination of the HSC and LDF processes is revealed.

## Material and Methods

### Sampling

Leaf samples from 391 individuals were collected from 24 *Aesculus* populations within the hybrid zone, along the border of the zone, and outside the zone (fig. 1). This sampling included six populations of *Aesculus pavia*, five populations of *Aesculus sylvatica*, five populations of *Aesculus flava*, seven hybrid populations between *A. sylvatica* and *A. pavia* ( $S \times P$ ), and one hybrid population between *A. sylvatica* and *A. flava* ( $S \times F$ ). Identification of populations was based on morphological characters of species and hybrid individuals described by Hardin (1957c), in combination with geographic distributions. Hybrid populations are abbreviated using the first letter of the specific epithet for each parental species but do not imply direction for the interspecific cross. The number of individuals per population ranged from 7 to 44, with an average of 16 individuals per population (table 1).

Total genomic DNA was extracted from fresh or silica-dried tissue using the CTAB protocol outlined by Cullings (1992). This protocol was modified by adding 0.5 volumes of 5 M NaCl before the addition of isopropanol to minimize the coprecipitation of polysaccharide contaminants with total DNA. The concentration and purity of DNA samples were measured using a GeneQuant spectrophotometer (Amersham Pharmacia Biotech, Cambridge, UK) before they were used for analysis.

### Microsatellites

Microsatellite data were collected following methods used by Oetting et al. (1995) and Vargo (2000). Eight microsatellite loci were amplified using primers designed for *Aesculus turbinata* Blume by Minami et al. (1998). Each locus was amplified in a 10- $\mu$ L reaction containing 10–20 ng of DNA

**Table 1**

**Population Number, Collection Location, and Number of Individuals Included in Microsatellite (MSAT) and Intersimple Sequence Repeat (ISSR) Analysis from 24 Populations of *Aesculus***

Species or hybrid name, population	Location	Number of individuals	
		MSAT	ISSR <sup>a</sup>
<i>A. pavia</i> :			
P1	Lee County, AL	19	7
P2	Lee County, GA	15	30
P3	Twiggs County, GA	8	11
P4	Charleston County, SC	19	...
P5	Liberty County, FL	22	...
P6	Bibb County, GA	11	...
<i>A. sylvatica</i> :			
S1	Elbert County, GA	12	14
S2	Gaston County, NC	10	12
S3	Jasper County, GA	17	20
S4	Monroe County, GA	15	21
S5	Wake County, NC	26	...
<i>A. flava</i> :			
F1	Rabun County, GA	9	13
F2	Blount County, TN	44	...
F3	Blount County, TN	18	...
F4	Sevier County, TN	7	...
F5	Sevier County, TN	10	...
<i>A. sylvatica</i> × <i>A. pavia</i> :			
SP1	Gwinnett County, GA	12	...
SP2	Butts County, GA	18	20
SP3	Catoosa County, GA	14	18
SP4	Coweta County, GA	19	23
SP5	Dade County, GA	10	22
SP6	Hart County, GA	15	18
SP7	Jasper County, GA	22	...
<i>A. sylvatica</i> × <i>A. flava</i> :			
SF1	Banks County, GA	19	26

<sup>a</sup> Populations missing in the ISSR analysis did not have material available when the ISSR analysis was performed, 2 yr before the MSAT analysis.

template, 2.5 mM MgCl<sub>2</sub>, 200 μM of each dNTP, 0.2 μg/μL BSA, 32 nM fluorescent-labeled M13F primer (LI-COR, Lincoln, NE), and 0.4 U of *Taq* DNA polymerase (Promega, Madison, WI), along with 10× PCR reaction buffer and 0.01 μM of each specific flanking primer. Amplification reactions were temperature cycled in 96-well plates using a Robocycler thermocycler (Stratagene, La Jolla, CA) with the following touchdown protocol: initial denaturing at 95°C (3 min); six cycles of denaturing at 95°C (45 s), annealing at 68°C (45 s, ramping down 2°C per cycle), and elongation at 72°C (1 min); followed by 30 cycles of denaturing at 95°C (45 s), annealing at 52°C (45 s), and elongation at 72°C (1 min), with an additional 10 min extension at 72°C.

After amplification, 9 μL of loading buffer (95% deionized formamide, 20 mM EDTA, 0.8 mg/mL bromophenol blue) was added to each sample. Samples were denatured at 90°C for 4 min, cooled on ice, and loaded onto a 6.5% KB Plus polyacrylamide gel (LI-COR). Samples were electrophoresed for 2 h (46°C, 1500 V) on a LI-COR 4200L automated DNA

sequencer. Six evenly spaced standards, along with a positive and a negative control, were included on the gel to minimize errors. The resulting gel images were scored using the program Saga GT (ver. 2.1.2; LI-COR).

### ISSRs

Four ISSR primers were selected based on successful and reliable amplification that produced the maximum number of bands. Primer names and sequences used included 844: (CT)<sub>8</sub>-RC; 814.1: (CT)<sub>8</sub>-TG; 17899: (CA)<sub>6</sub>-RG; and oligo7: (CT)<sub>8</sub>-RG. Amplification reactions (25 μL) included 30 ng of DNA template, 4.0 mM MgCl<sub>2</sub>, 200 μM of each dNTP, and 1 U of *Taq* DNA polymerase (Promega), along with 10× PCR reaction buffer and 20 μM of a single primer. Temperature cycling was done using a Robocycler thermocycler with the following temperature protocol: 1.5 min at 94°C; 35 cycles of 94°C (45 s), 45°C (45 s), and 72°C (1.5 min); and a final cycle of 94°C (45 s), 44°C (45 s), and 72°C (5 min). Amplification products were resolved using 1.5% agarose/0.7% synergel and stained with ethidium bromide. Bands were scored as present (1) or absent (0). Control lanes were run for each gel to ensure consistency of bands across gels of the same primer, and only repeatable bands were scored.

### Genetic Variability and Parental Species Specific Alleles

To estimate the current genetic composition of parental and hybrid populations of *Aesculus*, we first performed the following analyses for the microsatellite data. The total number of alleles and genotypes per locus and mean number of alleles and genotypes per locus for each species or hybrid were calculated using the program POPGENE (Yeh et al. 1999). Allelic richness ( $R_s$ ), which estimates the total number of alleles standardized for population size variation, was calculated for each population using the program FSTAT (Goudet 2001). FSTAT uses a rarefaction index to correct for differences in sample size, which influences the amount of genetic variability detected (Hurlbert 1971; Petit et al. 1998). To estimate genetic diversity from ISSRs, the total number and frequency of polymorphic bands per population were calculated and a measure of gene diversity based on mean allele frequencies ( $H_T$ ) was estimated using POPGENE. Species-unique alleles from microsatellites and ISSRs, defined as alleles present at a frequency of 0.05 or higher in at least one population of a parental species but absent in all populations of the other parental species, were identified by visual inspection of the raw data.

### Genetic Structure and Gene Flow

To test the LDF, HSC, and combined LDF + HSC hypotheses, we estimated the overall genetic structure and patterns of effective gene flow among populations by calculating genetic differentiation ( $\Phi_{ST}$ ), estimated effective gene flow ( $Nm_{est}$ ), and AMOVA from microsatellite and ISSR data. The genetic structure of the *Aesculus* hybrid zone was explored by calculating pairwise comparisons of  $\Phi_{ST}$  values between each population using the program Arlequin (Schneider et al. 2000). A nonparametric permutation scheme that tests significance by permuting individuals between populations as implemented

in Arlequin was used to test for the significance of each pairwise comparison, using 5040 permutations at a significance level of 0.05. Estimates of effective gene flow using Wright's island model approximation,  $Nm = (1/4)[(1/F_{ST}) - 1]$  (Wright 1931), within each parental and hybrid group were calculated based on  $\Phi_{ST}$  or  $G_{ST}$  values using POPGENE for microsatellite or ISSR data, respectively. Inferences into effective gene flow patterns from AMOVA were determined by using estimates of  $\Phi_{ST}$  in Arlequin. The total genetic variance was partitioned among individuals within populations, among populations within groups (within a species), and among groups (between species). The program Migrate was also used to estimate effective gene flow from microsatellite data, which uses a Metropolis-Hastings algorithm to estimate migration rates between multiple populations based on coalescent theory (Beerli and Felsenstein 2001). It allows for unequal population sizes and asymmetric gene flow and estimates gene flow between population pairs. A maximum likelihood analysis was run using default parameters defined for the microsatellite stepwise mutation model in the program.

To determine whether gene flow is restricted by geographic distance, we assessed isolation by distance using pairwise estimates of genetic differentiation ( $\Phi_{ST}/(1 - \Phi_{ST})$ ) calculated using Arlequin and geographical distance. Geographical distances between all populations were calculated using the software ArcView GIS (ver. 3.3). The program IBD (Bohonak 2002) was used to determine the significance of the association between genetic differentiation and geographical distance using a Mantel test with 10,000 randomizations. For microsatellite data, we estimated within population inbreeding by calculating the fixation index ( $F_{IS}$ ) using FSTAT.

#### Genetic Admixture

To gain additional information about the genetic structure of the hybrid zone, the program Structure (ver. 2.2; Pritchard et al. 2000) was used to estimate genetic admixture based on assignment probabilities of parental and hybrid individuals to parental populations. Structure uses a model-based Bayesian method that implements a Monte Carlo Markov chain algorithm to assign individual genotypes to each of  $K$  clusters. In this analysis, we set  $K$  equal to three clusters representing the three parental species. The microsatellite data set was run using the admixture model, which allows for mixed ancestry. The no-admixture model was used for the dominant ISSR

data set, which allows for the approximation of genetic admixture proportions when the heterozygous genotype cannot be determined (Pritchard et al. 2000). Analyses were run for 300,000 iterations after 30,000 iterations allotted for a burn-in period as determined by multiple runs.

Deviations from Hardy-Weinberg equilibrium (HWE) and genotypic linkage disequilibrium between all locus pairs were estimated with Fisher's exact tests using the program Genepop on the Web (M. Raymond and F. Rousset, <http://genepop.curtin.edu.au>). A Monte Carlo Markov chain with 1000 iterations was used to determine significance, in which  $P < 0.05$  indicated a statistically significant departure from HWE. For ISSRs, genotypic linkage disequilibrium between all pairs of polymorphic loci was estimated using Arlequin. Linkage disequilibrium ( $D'$ ), standardized for differences in allele frequencies (Lewontin 1964), was also calculated in Arlequin for all microsatellite locus pairs for the entire hybrid zone and parental species on a population basis.

## Results

### Microsatellite and ISSR Polymorphism

All eight microsatellite loci were highly polymorphic among the 24 populations. The 391 individuals produced 187 alleles, with the number of alleles per locus ranging from 15 (AT6D11) to 35 (AT3D6). The number of genotypes per locus ranged from 40 (AT6D2) to 164 (AT7D8). Measures of genetic variability, including mean number of alleles and genotypes per locus and allelic richness, were not significantly different between parental and hybrid *Aesculus* populations (table 2). Forty-five species-unique alleles were detected among the parental populations (data not shown). The majority of these alleles (24) were detected in populations of *Aesculus pavia*, with nine in *Aesculus sylvatica* and 12 in *Aesculus flava*. Eleven alleles unique to *A. pavia* were also detected in  $S \times P$  hybrid populations located within the hybrid zone. Only three alleles unique to *A. sylvatica* and two alleles unique to *A. flava* were detected in  $S \times P$  hybrid populations.

We scored 124 ISSR bands among 14 populations of *Aesculus*. Thirty-four species-unique bands were detected from populations of *A. pavia* (12) and *A. sylvatica* (22). The number of polymorphic loci and level of gene diversity were also similar among parental and hybrid populations (table 2).

Table 2

Measures of Microsatellite (MSAT) and Intersimple Sequence Repeat (ISSR) Polymorphism Detected within Parental Species and Hybrids of *Aesculus*

Taxon	MSAT			ISSR		$H_T$
	Mean alleles/locus	Mean genotypes/locus	Allelic richness	No. polymorphic bands	% polymorphic bands	
<i>A. pavia</i>	8.375 (1.12)	9.646 (2.43)	5.149 (1.54)	90	72.6	.143
<i>A. sylvatica</i>	7.850 (1.51)	9.825 (2.26)	5.230 (1.39)	95	76.6	.138
<i>A. flava</i>	6.975 (2.05)	9.225 (5.19)	4.822 (1.44)	57	46.0	...
$S \times P$	8.714 (1.18)	9.732 (2.19)	5.618 (1.36)	108	87.1	.152
$S \times F$	9.125	11.750	5.341	69	55.7	...

Note.  $H_T$  = gene diversity. Standard deviations are indicated in parentheses.  $S \times P = A. sylvatica \times A. pavia$ ;  $S \times F = A. sylvatica \times A. flava$ .

### Genetic Structure and Gene Flow

Pairwise  $\Phi_{ST}$  values from microsatellite data ranged from 0.020 to 0.240, and most comparisons were significantly different from 0, indicating weak to moderate genetic differentiation between populations. The few exceptions, where more extensive gene flow is suggested, were not necessarily found between neighboring populations or populations of the same taxa (e.g., between P4 and P6, P4 and SP3, or P4 and SP5; data not shown). This apparent lack of isolation by distance was further shown by nonsignificant correlations between pairwise geographic distance and genetic differentiation among parental and hybrid populations and by nonsignificant slopes of their regression ( $P = 0.24\text{--}0.51$ ). Average pairwise  $\Phi_{ST}$  values were highest between populations of different parental taxa (e.g., between *A. pavia* and *A. sylvatica*), while  $\Phi_{ST}$  values among populations of the same taxon or between parental and hybrid taxa were typically lower and similar to one another (table 3). However, estimates of effective gene flow ( $Nm_{est}$ ) from microsatellites were similar among populations of each parental species ( $Nm_{est} = 2.05$  in *A. pavia*, 1.93 in *A. sylvatica*, and 2.09 in *A. flava*) and among hybrid populations ( $Nm_{est} = 1.72$  in  $S \times P$ ). Overall, these data suggest that gene flow between parental and hybrid taxa occurs at rates similar to that within a taxon and that hybrid individuals may act to bridge the genetic gap between parental species where lower levels of interpopulation gene flow occur.

Lack of strong genetic differentiation and evidence for significant gene flow among parental and hybrid populations are further suggested by AMOVA from both microsatellite and ISSR data. AMOVA indicated that most of the genetic variance is distributed within populations rather than among

populations (within a species) or among groups (between species). These results show weak to moderate genetic differentiation among populations within each parental or hybrid taxon but little or no additional differentiation between taxa (table 4). Estimates of migration inferred from microsatellites using Migrate revealed a moderate level of effective gene flow between most population pairs of *Aesculus*, ranging from 0.07 to 2.82 (data not shown). Long-distance migration of *A. pavia* alleles into the hybrid zone was revealed by Migrate, where  $Nm > 1$ , by gene flow from population P1 into hybrid populations SP1, SP2, and SP5, from P2 into SP4, and from P6 into SP1.

Estimates of  $F_{IS}$  within populations were positive and ranged from 0.246 (S1) to 0.457 (F5), with an overall mean of 0.347. Mean  $F_{IS}$  values varied slightly among populations of each taxon but were not significantly different between parental and hybrid populations. This consistent heterozygote deficit in all taxa suggests that considerable inbreeding occurs within each of the 24 populations.

Estimates of effective gene flow from ISSR data were largely congruent with results from microsatellites. All pairwise comparisons of  $\Phi_{ST}$  were significantly different from 0, ranging from 0.053 to 0.345, and no significant correlation was detected between geographic distance and genetic differentiation. Estimates of  $Nm$  were highest in  $S \times P$  populations (3.762) and  $>1$  in both *A. pavia* (2.467) and *A. sylvatica* (1.963).

### Genetic Admixture

Results from the genetic admixture analysis with microsatellite data indicated that the majority of  $S \times P$  hybrid individuals (53%) were assigned to multiple parental clusters (fig. 2; table 5). A large number of  $S \times P$  hybrid individuals (38%) were assigned with high probability ( $>90\%$ ) to *A. sylvatica*, which is probably the result of repeated backcrossing to *A. sylvatica* of hybrids from historical gene flow and/or current gene flow. A small number of the  $S \times P$  hybrid individuals (10) from different populations were assigned to *A. pavia* with high probability ( $>90\%$ ), which probably resulted from recent hybridization from *A. pavia* and would require long-distance pollen dispersal, because any original hybrid with a high proportion of *A. pavia* alleles would have become extinct in the Piedmont since the end of last glacial period (Prentice et al. 1991).

Estimates of genetic admixture from microsatellites also show that, with a few notable exceptions, most putative parental individuals of *A. pavia* and *A. flava* were correctly assigned to their respective populations (table 5). Individuals from one population of *A. pavia* (P1), which is located near the hybrid zone, were intermediately assigned to *A. pavia* and *A. sylvatica* and represent virtually a hybrid population, although this relationship is not evident from the morphology. Likewise, two populations of *A. flava* (F4, F5) contained individuals that were assigned to *A. flava* and *A. sylvatica*. Putative *A. sylvatica* parental individuals were mostly assigned to the cluster representing *A. sylvatica*, with a certain proportion of individuals assigned to the other parental taxa with moderate probabilities (0.161–0.461), suggesting that the hybrid zone may be expanding. Hybrid individuals from population SP3, which borders the northern boundary of the hybrid zone

**Table 3**

**Average Level of Genetic Differentiation ( $\Phi_{ST}$ ) among Parental and Hybrid Populations of *Aesculus* from Microsatellite Data, with Number of Comparisons ( $N$ ) and SD**

Taxon, parameter	<i>A. pavia</i>	<i>A. sylvatica</i>	<i>A. flava</i>	$S \times P$	$S \times F$
<i>A. pavia</i> :					
$N$	15				
$\Phi_{ST}$	.086				
SD	.026				
<i>A. sylvatica</i> :					
$N$	30	10			
$\Phi_{ST}$	.138	.105			
SD	.050	.044			
<i>A. flava</i> :					
$N$	30	25	10		
$\Phi_{ST}$	.154	.109	.087		
SD	.037	.043	.036		
$S \times P$ :					
$N$		42	35	35	21
$\Phi_{ST}$	.109	.089	.100	.076	
SD	.050	.039	.032	.040	
$S \times F$ :					
$N$	6	5	5	7	...
$\Phi_{ST}$	.149	.075	.079	.073	...
SD	.037	.036	.040	.024	...

Note.  $S \times P = A. sylvatica \times A. pavia$ ;  $S \times F = A. sylvatica \times A. flava$ .

**Table 4****Results from AMOVA from Microsatellite (MSAT) and Intersimple Sequence Repeat (ISSR) Data Using Arlequin**

Variance partition	df	Variance	% of total	<i>P</i>
MSAT:				
Among groups	4	.079	2.51	.002
Among populations	19	.244	7.69	<.0001
Within populations	758	2.851	89.90	<.0001
ISSR:				
Among groups	4	.017	.15	.371
Among populations	9	1.620	13.79	<.0001
Within populations	241	10.101	86.06	<.0001

and represents the collected population most distantly located from *A. pavia*, were mostly assigned to *A. pavia* with moderate probability (0.684), which also suggests long-distance gene flow across the hybrid zone (fig. 2, table 5).

Estimates of genetic admixture from ISSRs revealed results mostly congruent with the microsatellite data. The majority of the  $S \times P$  hybrid individuals showed the highest assignment probabilities to *A. sylvatica* (table 5). However, examples of conflicting estimates were detected between microsatellite and ISSR results. For example, individuals from parental populations P1, P2, and S4 showed considerable difference in assignments to parental clusters between microsatellite and ISSR analyses. The discrepancies between markers in this analysis

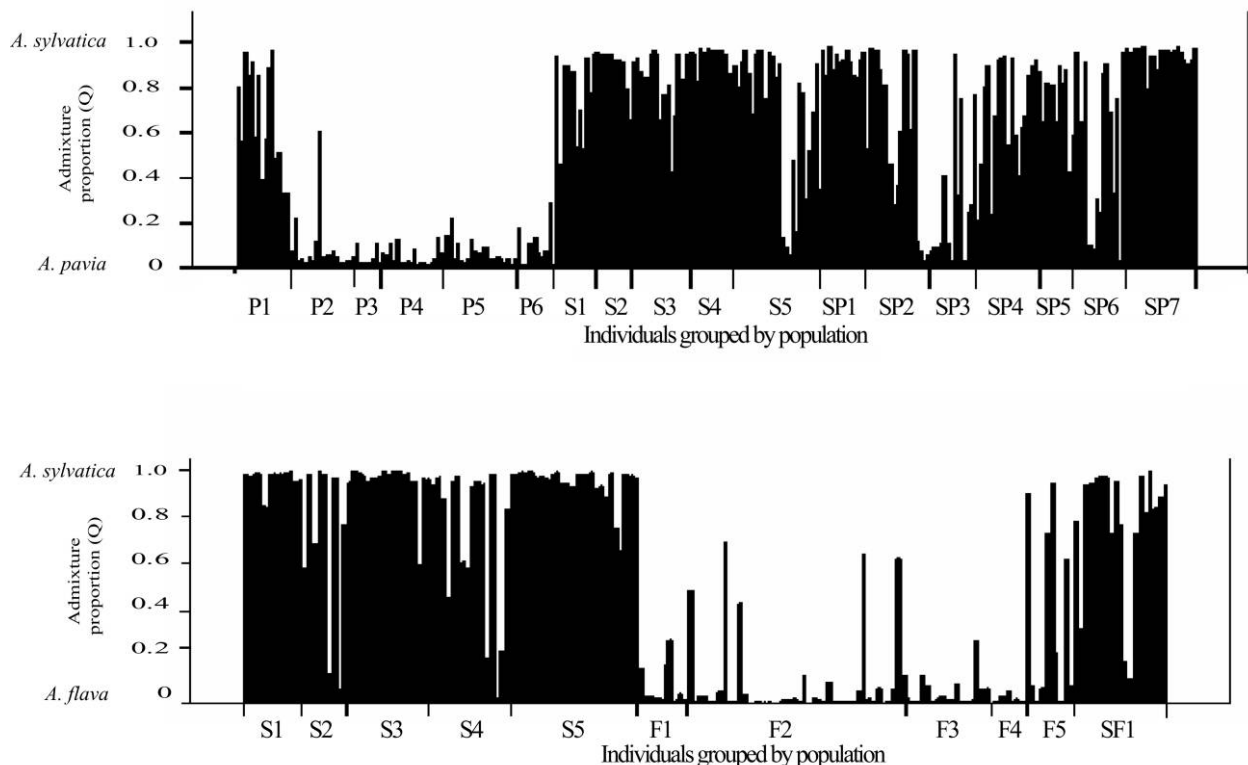
may be due to differences in sampling since the individuals and sampling size for each population were not always the same for both markers (table 1).

Statistically significant deviations from HWE were detected in all populations, although the number of microsatellite loci that showed significant deviations varied among populations. Tests for linkage disequilibrium found that 7.6% of all microsatellite locus pairs for all populations significantly deviated from HWE but were not concentrated in any particular population. Standardized values of  $D'$  were strongly negative and ranged from  $-0.626$  to  $-0.349$ . In ISSRs, the proportion of significant linkage disequilibrium detected among populations (15.2%) was higher than expected at a 0.05 level of significance, although no particular locus pair was in disequilibrium across all populations.

## Discussion

### *Evidence of Long-Distance Gene Flow and a Combination of the HSC and LDF Hypotheses*

In this study, evidence from several analyses supports long-distance dispersal of *Aesculus pavia* alleles into the hybrid zone. Pairwise  $\Phi_{ST}$  data indicate a few cases of extensive gene flow among distantly located populations. For example, microsatellite data revealed nonsignificant  $\Phi_{ST}$  values between a population of *A. pavia* (P6) from Bibb County, Georgia, near the southern border of the hybrid zone, and two hybrid



**Fig. 2** Estimated admixture proportion of individuals assigned to parental taxa from microsatellite data.

**Table 5**  
Average of Individual Assignment Probabilities for 24 Populations of *Aesculus* Based on Estimates from Bayesian Inference from the Program Structure, Using Microsatellite (MSAT) and Intersimple Sequence Repeat (ISSR) Data

Population	MSAT			ISSR		
	<i>A. sylvatica</i>	<i>A. pavia</i>	<i>A. flava</i>	<i>A. sylvatica</i>	<i>A. pavia</i>	<i>A. flava</i>
P1	.518	.410	.072	.696	.013	.291
P2	.022	.907	.071	.611	.245	.144
P3	.020	.954	.026	.007	.725	.268
P4	.020	.959	.021	...	...	...
P5	.026	.951	.023	...	...	...
P6	.060	.921	.019	...	...	...
S1	.604	.296	.100	.681	.013	.307
S2	.487	.051	.461	.721	.008	.271
S3	.776	.161	.063	.923	.013	.064
S4	.525	.039	.437	.089	.715	.196
S5	.679	.232	.089	...	...	...
F1	.036	.039	.925	.274	.246	.479
F2	.042	.026	.932	...	...	...
F3	.095	.019	.886	...	...	...
F4	.225	.035	.740	...	...	...
F5	.386	.046	.568	...	...	...
SP1	.881	.023	.096	.374	.355	.271
SP2	.708	.204	.088	.835	.057	.108
SP3	.095	.684	.221	.074	.707	.220
SP4	.479	.341	.180	.798	.030	.172
SP5	.771	.074	.155	...	...	...
SP6	.420	.369	.211	.474	.280	.246
SP7	.715	.057	.228	...	...	...
SF1	.584	.025	.391	.675	.022	.303

populations of  $S \times P$  (SP3, 0.050; SP5, 0.050) in northern Georgia (fig. 1), separated by a distance of more than 200 km. Both the  $\Phi_{ST}$  data and the AMOVA results (table 4) suggest that gene flow is generally not limited among populations, and the non-significant  $\Phi_{ST}$  values between several populations are consistent with the LDF hypothesis. Estimates of migration using microsatellites from MIGRATE revealed long-distance migration of *A. pavia* alleles into the hybrid zone ( $Nm > 1$ ), e.g., from P1 into hybrid populations SP1, SP2, and SP5, from P2 into SP4, and from P6 into SP1.

The genetic admixture data also revealed the existence of hybrid individuals that were assigned to *A. pavia* with a high probability (>90%) (fig. 2; table 5). Since *A. pavia* is not found in close proximity to hybrid populations, these hybrid individuals with strong genotypic assignment to *A. pavia* would have to be derived from repeated backcrosses to *A. pavia* via long-distance pollen dispersal and thus provide support for the LDF hypothesis. The  $S \times P$  hybrid individuals that were partially assigned to *A. pavia* with low to moderate probabilities could be the result of long-distance gene flow of *A. pavia* alleles followed by repeated backcrossing to *Aesculus sylvatica*, or of historical in situ hybridization followed by a long history of dilution by *A. sylvatica* alleles. Microsatellite data revealed that individuals from four  $S \times P$  hybrid populations (SP2, SP3, SP4, and SP6) were partially assigned to *A. pavia* with moderate probabilities (0.204–0.684). This was also detected from ISSR data in SP1 and SP3 (table 5).

Genetic differentiation was not particularly higher among populations within the *Aesculus* hybrid zone, as might be ex-

pected if the zone had been formed purely as a result of historical contact of parental species followed by isolation between the hybrid zone and parental species and restricted gene flow within the zone (Latta and Mitton 1999). The low level of genetic differentiation and low variance (table 4) between parental and hybrid populations detected suggest that the parental populations have not been separated for a long period of time and/or that local and long-distance gene flow have been sufficient over time to homogenize the gene pool of hybrid populations.

Strong positive linkage disequilibrium in hybrid populations detected in diagnostic marker alleles has been used to reveal recent gene flow from parental species into hybrid populations (Barton and Gale 1993). However, our results show that values of disequilibrium between all pairs of microsatellite loci were strongly negative. There are several possible explanations. One is the strong inbreeding detected within populations ( $F_{IS} = 0.246\text{--}0.457$ ), which results in negative linkage disequilibrium values (Silvela et al. 1999). Second, we were unable to detect fixed species-specific diagnostic microsatellite markers in our data, which limits the power to detect current gene flow from *A. pavia* via linkage disequilibrium (Cruzan 2005). Third, long-distance gene flow from *A. pavia* into the hybrid populations is expected to be rare and therefore may not be revealed in levels of disequilibrium because those levels would be reduced by 50% with each succeeding generation (Cruzan 2005). This suggests that caution should be taken when using linkage disequilibrium alone to detect current gene flow from parental species into hybrid zones.

LDF detected in *Aesculus* would be limited almost exclusively to the transfer of viable pollen, given the large size and weight of *Aesculus* seeds (Hardin 1957a; Schopmeyer 1974), as also proposed by dePamphilis and Wyatt (1990). This is consistent with the finding of a cpDNA study (Modliszewski et al. 2006) that gene flow via seeds is highly restricted among populations. According to dePamphilis and Wyatt (1989), frequent visitors to flowers of *Aesculus* species include bees, a butterfly, and hummingbirds. Most of these potential pollinators lack the capability to transfer pollen over long distances. However, the ruby-throated hummingbird is capable of flying long distances during migratory periods, which may allow it to transfer pollen between flowers of distant populations.

Results of genetic analyses of microsatellite and ISSR loci from this study agree with the expectation of a combination of historical hybridization and current gene flow in the *Aesculus* hybrid zone. The total evidence from cpDNA (Modliszewski et al. 2006) and nuclear microsatellite and ISSR analyses together support that both LDF and HSC have contributed to the origin, maintenance, and genetic structure of the hybrid zone. The zone probably originated from HSC and has been maintained by current gene flow, including occasional LDF.

It must be noted that our evidence for LDF relies on the homology of alleles in the hybrid zone and *A. pavia*. It is possible that the sharing of alleles between *A. pavia* and hybrid individuals is the result not of gene flow or common ancestry but of convergent evolution of alleles. However, we consider the impact of convergent evolution of alleles not likely to be significant here. Convergent evolution would have occurred at random, but we see distinct patterns where hybrid populations contain low proportions of *A. pavia* alleles along with a few individuals that contain a considerable proportion of *A. pavia* alleles. Moreover, comparisons between data from two randomly distributed nuclear markers (microsatellites and ISSRs) were mostly congruent, which would not result if there were a high degree of homoplasy.

In addition to homoplasy, incomplete lineage sorting in recently diverged taxa may also give results that can be misinterpreted as gene flow. Parsimony analysis of chloroplast *matK* DNA sequences did not group multiple samples of the same species within sect. *Pavia* together, which may have resulted from incomplete lineage sorting or ancestral polymorphism of cpDNA (Xiang et al. 1998; Modliszewski et al. 2006). However, the effects of incomplete lineage sorting are dependent on the time since divergence of the species and the rate of molecular evolution for a specific marker. A calibrated molecular clock based on ITS data estimated that species of sect. *Pavia*, which includes the hybridizing species, diverged 5–10 Myr ago (Xiang et al. 1998). The effects of incomplete lineage sorting are likely to be negligible in our data, given the estimated time since divergence and the relatively high rate of mutations observed in microsatellite markers ( $5 \times 10^{-3}$ – $5 \times 10^{-5}$  events per locus per generation; Estoup et al. 2002).

#### *Gene Flow Undetected by Morphology and a Stabilized Hybrid Population*

In the allozyme work by dePamphilis and Wyatt (1990), a small proportion of *Aesculus flava* alleles was detected in

some *A. pavia*, *A. sylvatica*, and  $S \times P$  populations that displayed no detectable morphological features characteristic of *A. flava*. Such cases were also detected in our study. Genetic admixture analysis from both microsatellite and ISSR data revealed that  $S \times P$  hybrid individuals from several populations (SP3, SP4, SP5, SP6, and SP7) were partially assigned to *A. flava* with a moderate probability level (table 5), suggesting introgression of *A. flava* genes into the hybrid zone, something not evident from morphology. Another example where morphological characters did not reveal hybrid individuals was found in a population of *A. pavia* (P1) that borders the hybrid zone, in which individuals were intermediately assigned to both *A. pavia* and *A. sylvatica* (fig. 2, table 5).

DePamphilis and Wyatt's study (1990) found three hybrid populations that lacked plants of both parental species, and they proposed that these might have been stabilized hybrid populations. We included two of the three populations (SP3 and SP4) in our study. Microsatellite and ISSR data revealed that the majority of individuals in population SP3 were assigned to *A. pavia* (table 5). Microsatellite data also showed a considerable assignment of individuals from SP4 to *A. pavia*, something not seen with ISSR markers. Given that these populations are located in the central to northern range of the hybrid zone, outside of the current range of *A. pavia*, it is possible that there has been long-distance dispersal of *A. pavia* into these hybrid populations and that selection favors backcrosses to *A. pavia*. With the occurrence of ongoing gene flow, the low genetic divergence of these hybrid populations and that among other populations may have been largely attributed to inbreeding, given that moderate levels of inbreeding are suggested by relatively high  $F_{IS}$  values at all loci. Thus, individuals may have a greater chance to cross with a closely related individual in close proximity than to outbreed over longer distances.

#### *Conclusions*

In conclusion, we detected high genetic diversity at microsatellite and ISSR loci among the species of *Aesculus* sect. *Pavia* in the hybrid zone. This variation allowed us to explore the genetic structure of the hybrid zone and to assess patterns of gene flow both within the hybrid zone and between parental and hybrid populations. Our results agree with those of the allozyme study of dePamphilis and Wyatt (1990) in confirming the existence of the hybrid zone, gene introgression between species that is not evident from morphology, and an asymmetrical genetic structure of the hybrid zone biased toward *A. sylvatica*. Our data revealed almost equivalent levels of genetic variability between parental and hybrid populations, genetic variation mostly partitioned within populations, and moderate levels of inbreeding based on  $F_{IS}$  estimates. The lack of isolation by distance, genetic admixture by *A. pavia* detected in some hybrid individuals, and lack of population structure from AMOVA are all in agreement with the  $\Phi_{ST}$  inference that geographic distance is not a limiting factor for gene flow in these taxa, which is congruent with the LDF scenario. A combination of frequent inbreeding and genetic drift is probably largely responsible for the level of genetic divergence among populations detected by the low, but significant,  $\Phi_{ST}$  values. It is plausible that substantial inbreeding balanced

by gene flow and selection in some hybrid populations could eventually lead to a stabilized hybrid or new species in this region (Arnold 1997). Most important, our data support the idea that LDF occurs between some populations of *A. pavia* in the Coastal Plain and populations in the Piedmont (the LDF hypothesis) and that overall weak genetic differentiation exists among most populations within and surrounding the hybrid zone. Our study on the *Aesculus* hybrid zone using hypervariable nuclear markers in conjunction with cpDNA results from Modliszewski et al. (2006) provides an example of a hybrid zone that involves HSC for the formation of the zone and recurrent long-distance gene flow for its ongoing maintenance.

## Acknowledgments

The authors thank Margaret Ptacek for insightful discussions and technical help on the project and the Core Molecular Facility at Idaho State University for help in collecting preliminary data. We also thank E. Minami and colleagues for development of microsatellite primers used in this study. At North Carolina State University, we thank the labs of Ed Vargo and Fred Gould for use of the LI-COR DNA sequencers. This research was supported by a Faculty Research grant from Idaho State University and a Faculty Research and Development grant from North Carolina State University, both awarded to Q.-Y. Xiang.

## Literature Cited

- Anderson E, GL Stebbins Jr 1954 Hybridization as an evolutionary stimulus. *Evolution* 8:378–388.
- Arnold ML 1993 *Iris nelsonii*: origin and genetic composition of a homoploid hybrid species. *Am J Bot* 80:577–583.
- 1997 Natural hybridization and evolution. Oxford University Press, New York.
- Arnold ML, SA Hodges 1995 Are natural hybrids fit or unfit relative to their parents? *Trends Ecol Evol* 10:67–71.
- Barton NH 2001 The role of hybridization in evolution. *Mol Ecol* 10:551–568.
- Barton NH, KS Gale 1993 Genetic analysis of hybrid zones. Pages 13–45 in RG Harrison, ed. Hybrid zones and the evolutionary process. Oxford University Press, New York.
- Beerli P, J Felsenstein 2001 Maximum likelihood estimation of a migration matrix and effective population sizes in *n* subpopulations by using a coalescent approach. *Proc Natl Acad Sci USA* 98:4563–4568.
- Bohonak AJ 2002 IBD (isolation by distance): a program for analyses of isolation by distance. *J Hered* 93:153–154.
- Campbell DR 2003 Natural selection in *Ipomopsis* hybrid zones: implications for ecological speciation. *New Phytol* 161:83–90.
- Campbell DR, NM Waser 2001 Genotype-by-environment interaction and the fitness of plant hybrids in the wild. *Evolution* 55:669–676.
- 2007 Evolutionary dynamics of an *Ipomopsis* hybrid zone: confronting models with lifetime fitness data. *Am Nat* 169:298–310.
- Cruzan MB 2005 Patterns of introgression across an expanding hybrid zone: analyzing historical patterns of gene flow using nonequilibrium approaches. *New Phytol* 167:267–278.
- Cullings KW 1992 Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. *Mol Ecol* 1:233–240.
- dePamphilis CW, R Wyatt 1989 Hybridization and introgression in buckeyes (*Aesculus*: Hippocastanaceae): a review of the evidence and a hypothesis to explain long-distance gene flow. *Syst Bot* 14:593–611.
- 1990 Electrophoretic confirmation of interspecific hybridization in *Aesculus* (Hippocastanaceae) and the genetic structure of a broad hybrid zone. *Evolution* 44:1295–1317.
- Estoup A, P Jarne, JM Cornuet 2002 Homoplasmy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Mol Ecol* 11:1591–1604.
- Goudet J 2001 FSTAT, a program to estimate and test gene diversities and fixation indices, version 2.9.3. <http://www.unil.ch/izea/software/fstat.html>.
- Hardin JW 1957a A revision of the American Hippocastanaceae. *Brittonia* 9:145–171.
- 1957b Studies in the Hippocastanaceae. III. A hybrid swarm in the buckeyes. *Rhodora* 59:45–51.
- 1957c Studies in the Hippocastanaceae. IV. Hybridization in *Aesculus*. *Rhodora* 59:185–203.
- Harrison RG 1993 Hybrids and hybrid zones: historical perspective. Pages 3–12 in RG Harrison, ed. Hybrid zones and the evolutionary process. Oxford University Press, New York.
- Hewitt GM 1993 After the ice: *Parallelus* meets *Erythropus* in the Pyrenees. Pages 140–164 in RG Harrison, ed. Hybrid zones and the evolutionary process. Oxford University Press, New York.
- 1996 Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol J Linn Soc* 58:247–276.
- Hurlbert SH 1971 The nonconcept of species diversity: a critique and alternative parameters. *Ecology* 52:577–586.
- James JK, RJ Abbott 2005 Recent allopatric, homoploid hybrid speciation: the origin of *Senecio squalidus* (Asteraceae) in the British Isles from a hybrid zone on Mount Etna, Sicily. *Evolution* 59:2533–2547.
- Johansen-Morris AD, RG Latta 2006 Fitness consequences of hybridization between ecotypes of *Avena barbata*: hybrid breakdown, hybrid vigor, and transgressive segregation. *Evolution* 60:1585–1595.
- Latta RG, JB Mitton 1999 Historical separation and present gene flow through a zone of secondary contact in ponderosa pine. *Evolution* 53:769–776.
- Lewontin RC 1964 The interaction of selection and linkage. 1. General considerations; heterotic models. *Genetics* 49:49–67.
- Mayr E 1942 Systematics and the origin of species. Columbia University Press, New York.
- Miglia KJ, ED McArthur, WS Moore, H Wang, JH Graham, DC Freeman 2005 Nine-year reciprocal transplant experiments in the gardens of the basin and mountain big sagebrush (*Artemisia tridentata*: Asteraceae) hybrid zone of Salt Creek Canyon: the importance of multiple-year tracking of fitness. *Biol J Linn Soc* 86:213–225.
- Minami E, Y Isagi, Y Kaneko, H Kawaguchi 1998 Polymorphic microsatellite markers in Japanese horse chestnut *Aesculus turbinata* Blume. *Mol Ecol* 7:1613–1621.
- Modliszewski JL, DT Thomas, C Fan, DJ Crawford, CW dePamphilis, Q-Y Xiang 2006 Ancestral chloroplast polymorphism and historical secondary contact in a broad hybrid zone of *Aesculus* (Sapindaceae). *Am J Bot* 93:377–388.
- Oetting WS, HK Lee, DJ Flanders, GL Wiesner, TA Sellers, RA King 1995 Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. *Genomics* 30:450–458.
- Petit RJ, A El Mousadik, O Pons 1998 Identifying populations for conservation on the basis of genetic markers. *Conserv Biol* 12:844–855.

- Prentice IC, PJ Bartlein, T Webb 1991 Vegetation and climate change in eastern North America since the last glacial maximum. *Ecology* 72:2038–2056.
- Pritchard JK, M Stephens, P Donnelly 2000 Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Rieseberg LH 1995 The role of hybridization in evolution: old wine in new skins. *Am J Bot* 82:944–953.
- 1997 Hybrid origins of plant species. *Annu Rev Ecol Syst* 28:359–389.
- Rieseberg LH, O Raymond, DM Rosenthal, Z Lai, K Livingstone, T Nakazato, JL Durphy, AE Schwarzbach, LA Donovan, C Lexer 2003 Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* 301:1211–1216.
- Rieseberg LH, J Wendel 1993 Introgression and its consequences in plants. Pages 70–109 in RG Harrison, ed. *Hybrid zones and the evolutionary process*. Oxford University Press, New York.
- Schneider S, J-M Kueffer, D Roessli, L Excoffier 2000 Arlequin: a software for population genetics data analysis, version 2.000. Genetics and Biometry Lab, Department of Anthropology, University of Geneva, Geneva.
- Schopmeyer CS 1974 Seeds of woody plants in the United States. USDA Handbook 450. US Government Printing Office, Washington, DC.
- Silvela L, I de la Peña, R Gomez-Ruano 1999 Selection under negative linkage disequilibrium. Random mating versus inbreeding. *Heredity* 82:598–604.
- Vargo EL 2000 Polymorphism at trinucleotide microsatellite loci in the subterranean termite *Reticulitermes flavipes*. *Mol Ecol* 9:817–829.
- Wagner WH 1969 The role and taxonomic treatment of hybrids. *BioScience* 19:785–789.
- 1970 Biosystematics and evolutionary noise. *Taxon* 19:146–151.
- Wright S 1931 Evolution in Mendelian populations. *Genetics* 16:97–159.
- Xiang QY, DJ Crawford, AD Wolfe, YC Tang 1998 Origin and biogeography of *Aesculus* L. (Hippocastanaceae): a molecular phylogenetic perspective. *Evolution* 52:988–997.
- Yeh FC, R-C Yang, T Boyle 1999 POPGENE, Microsoft Windows-based freeware for population genetic analysis, version 1.31. <http://www.ualberta.ca/~fyeh/>.