SPATIAL GENETIC STRUCTURE OF A TROPICAL UNDERSTORY SHRUB, *PSYCHOTRIA OFFICINALIS* (RUBIACEAE)\(^1\)

**BETTE A. LOISELLE,\(^3\) VICTORIA L. SORK,\(^2,3\) JOHN NASON,\(^4\) AND CATHERINE GRAHAM\(^3\)**

\(^1\)Department of Biology, University of Missouri-St. Louis, St. Louis, Missouri 63121; and
\(^2\)Department of Botany, University of Georgia, Athens, Georgia 30602

Analyses of fine-scale and macrogeographic genetic structure in plant populations provide an initial indication of how gene flow, natural selection, and genetic drift may collectively influence the distribution of genetic variation. The objective of our study is to evaluate the spatial dispersion of alleles within and among subpopulations of a tropical shrub, *Psychotria officinalis* (Rubiaceae), in a lowland wet forest in Costa Rica. This insect-pollinated, self-incompatible understory plant is dispersed primarily by birds, some species of which drop the seeds immediately while others transport seeds away from the parent plant. Thus, pollination should promote gene flow while at least one type of seed dispersal agent might restrict gene flow. Sampling from five subpopulations in undisturbed wet forest at Estación Biológica La Selva, Costa Rica, we used electrophoretically detected isozyme markers to examine the spatial scale of genetic structure. Our goals are: 1) describe genetic diversity of each of the five subpopulations of *Psychotria officinalis* sampled within a contiguous wet tropical forest; 2) evaluate fine-scale genetic structure of adults of *P. officinalis* within a single 2.25-ha mapped plot; and 3) estimate genetic structure of *P. officinalis* using data from five subpopulations located up to 2 km apart. Using estimates of coancestry, statistical analyses reveal significant positive genetic correlations between individuals on a scale of 5 m but no significant genetic relatedness beyond that interplant distance within the studied subpopulation. Multilocus estimates of genetic differentiation among subpopulations were low, but significant (\(F_{st} = 0.095\)). Significant \(F_{st}\) estimates were largely attributable to a single locus (Lap-2). Thus, multilocus estimates of \(F_{st}\) may be influenced by microgeographic selection. If true, then the observed levels of IBD may be overestimates.

The spatial distribution of genetic variation within and among populations is the outcome of gene flow, genetic drift, and natural selection. In plant populations, gene flow through pollination and seed dispersal determines the extent to which genes are locally or more widely dispersed (Bradshaw, 1972; Levin and Kerster, 1974). For tropical understory shrubs whose flowers are pollinated by insects and whose fruits are dispersed by birds, the opportunity for gene flow should be extensive. Yet, the behavior of animal mutualists also could result in restricted pollen or seed movement. This point is particularly true for tropical plants whose seeds are dispersed by an assemblage of vertebrate species, some of which drop seeds during fruit handling while others may move away from a parent plant before dropping seeds (Howe and Vande Kerckhove, 1981; Levey, 1987). Because observations of frugivorous vertebrates do not reveal which animal species are dispersing seeds that become seedlings, their impact on the genetic structure of plant populations must be evaluated by examining the spatial distribution of genotypes.

An important consequence of restricted gene flow is isolation by distance (IBD) (Wright, 1943, 1946, 1969). IBD means that the frequencies of selectively neutral genes in the local population will not reflect those of the global population due to genetic drift resulting from diminishing gene exchange with increasing distance and random loss of alleles (Heywood, 1991). Heywood (1991) reviews methods of evaluating the spatial structure of genetic variation of populations including the use of \(F\) statistics (Wright, 1951; Weir and Cockerham, 1984) and spatial autocorrelation analysis (Sokal and Oden, 1978; Sokal and Wartenberg, 1983; Barbujani, 1987; Ep- person, 1989). Under models of IBD, one would expect that \(F_{st}\), a measure of genetic differentiation among subpopulations, would increase as subpopulations become more widely separated. Thus, a significant value of \(F_{st}\) may indicate IBD.

Spatial autocorrelation analysis is a statistical procedure that has been used to identify correlations among the genotypes of mapped individuals (Heywood, 1991). These statistics have the advantage of identifying the scale of genetic structure without prior knowledge of that scale (Heywood, 1991). In making specific assumptions about the origins of alleles identical in state, one can estimate the probability that genes in different individuals within subpopulations are identical by descent by using a measure of coancestry, \(\rho\), a genetic structure statistic. Under IBD, one would expect that estimates of coancestry should decline gradually for pairs of individuals located at increasing geographic interplant distance. Thus,
spatial autocorrelation analysis provides a means of evaluating the genetic consequences of dispersal over fine spatial scales within subpopulations while $F_c$ allows one to evaluate the impact of dispersal over longer distances between subpopulations.

In this paper, we evaluate the spatial dispersion of genes within and among subpopulations of a tropical understory shrub, Psychotria officinalis, in a lowland wet forest in Costa Rica. This understory plant has distyloous flowers, is reportedly self-incompatible, and is pollinated by butterflies and small bees (Bawa and Beach, 1983; Kress and Beach, 1994). Its berries are dispersed primarily by understory birds (Loiselle and Blake, 1990, 1993). Within a population, we would expect the pollination system to result in widespread gene flow. However, depending on the behavior of birds as dispersal agents, substantial fine-scale genetic structure may develop over short-distance intervals within subpopulations if seeds are dropped near the parent plant, while little spatial structuring may result if seeds are transported farther away before deposition. Consequently, an analysis of fine-scale genetic structure should indicate the pattern of effective seed dispersal by birds as well as patterns of relationship among established individuals.

Specific objectives of our study are to: 1) describe genetic diversity of each of the five subpopulations of Psychotria officinalis sampled within a contiguous wet tropical forest; 2) evaluate fine-scale spatial genetic structure of adults of P. officinalis within a single 2.25-ha mapped plot; and 3) estimate larger scale spatial genetic differentiation in P. officinalis using data from five subpopulations distributed at different distances from each other within the study site.

**MATERIALS AND METHODS**

**Study site and species**—This study was conducted in tropical wet forest at Estación Biológica La Selva, a field station operated by the Organization for Tropical Studies (OTS) in northeast Costa Rica. La Selva, which encompasses 1,536 ha, is adjacent to Parque Nacional Braulio Carrillo (44,000 ha). La Selva receives 4 m rain annually with the lowest rainfall on average occurring in February and March and the wettest months in July and August (McDade and Hartshorn, 1994; Sanford et al., 1994). A complete description of this site is available in McDade et al. (1994; and references therein).

The study plant is Psychotria officinalis (Aubl.) Sandw., an understory shrub in the Rubiaceae. P. officinalis is relatively common at La Selva, but is restricted mainly to residual soils derived from old lava flows; few individuals occur on recent or old alluvial soils near the front of the property (Loiselle, personal observation). P. officinalis can reach heights of 5–6 m and 5-cm stem diameter (Loiselle and Sork, unpublished data). Small white flowers are produced in terminal panicles and are likely pollinated by small bees and butterflies (Bawa and Beach, 1983; Kress and Beach, 1994). Two-seeded berries (6–7 mm diameter) turn purple-black at maturity and are eaten primarily by understory birds, especially manakins (Pipridae), but also thrushes (Turdinae) and tanagers (Thraupinae) (Loiselle and Blake, 1990, 1993).

**Field sampling**—During July and August 1993, we located four study populations of P. officinalis within undisturbed forest at La Selva (Fig. 1). Populations located adjacent to Lindero Occidental (populations 2, 3, and 4) occurred on the Jaguar residual solid consociation, whereas population 1 occurred on the Matahuey soil consociation (Sollins et al., 1994). These two residual soil types are derived from different lava flows, but are chemically and physically similar. These soils are considered in the ultisol group and are strongly acidic, with rich organic matter, relatively large amounts of exchangeable acidity, and low base saturation (Sollins et al., 1994).

At population 1, we mapped the position of all adult P. officinalis individuals ($N = 175$) within a 2.25-ha area (Fig. 2). In general, adults were defined as those individuals $>0.75$-cm stem diameter, a size at which reproductive activity had been observed in a long-term study of understory fruit-eating birds and bird-dispersed plants (e.g., Loiselle and Blake, 1990, 1993). At populations 2–4, we sampled adults over an area of $\sim 1$ ha; leaves were only collected from adults that were located $2–5$ m or more apart.

At all four study sites, material was collected for electrophoretic analysis by placing freshly clipped leaves (two to four per plant) into individual plastic bags. While in the field, leaves were kept on blue ice in a cooler. (A layer of newspaper was placed between ice packs and leaves to prevent frost damage.) Upon return to the station’s laboratory, leaves were transferred to a liquid nitrogen tank for storage until return to the University of Missouri-St. Louis where they were immediately placed in an ultra-cold freezer ($-75 \,^\circ$C).

**Electrophoretic studies**—Allozyme analyses generally followed methods described by Soltis et al. (1983). We used a phosphate-polyvinylpyrrolidone extraction (“camellia”) buffer (Wendel and Parks, 1982) and liquid nitrogen to crush leaves with a mortar and pestle. This slur-
phoglucomutase (PGM), and triosephosphate isomerase (TPI). These loci are assumed to be selectively neutral and unlinked in the genetic analyses that follow.

Data analyses—Genetic analyses were conducted on five subpopulations. In order to include closely neighbored subpopulations, site 1 was divided into two subpopulations by assigning individuals to east and west subpopulations and excluding any individuals within a 30-m wide area separating the two subpopulations (1E and 1W, respectively). This 30-m gap was a natural swale that contained few Psychotria adults (Fig. 2). In order to emulate sampling techniques used for the other three sites, populations and excluding any individuals within a 30-m gap was a natural swale that contained few Psychotria adults (Fig. 2). In order to examine patterns of fine-scale genetic structure, multilocus allozyme genotypes based on nine polymorphic loci were used to estimate the coancestry, $p_{ij}$, between all possible pairs of individuals within the mapped subpopulations (Cockerham, 1969). In contrast to the spatial autocorrelation measures typically used (e.g., Moran’s I), genetic structure statistics, such as $p_{ij}$, have a sound foundation in population genetics theory and provide a natural means of summarizing data over alleles at a locus and over loci to obtain a more powerful test for spatial genetic structure (Heywood, 1991). $p_{ij}$ measures the correlation in the frequencies of homologous alleles, $p_i$ and $p_j$, at a locus in pairs of mapped individuals $i$ and $j$ and can be estimated as

$$
\hat{p}_{ij} = \frac{\sum_{k=1}^{n} (p_i - \bar{p})(p_j - \bar{p})}{k(n-1)} + \frac{2}{(8k+1)^{0.5} - 1}
$$

where the first term is the expected value of $p_{ij}$, and $k = n(n-1)/2$ is the total number of possible pairwise connections between $n$ individuals located a discrete number of map units away from each other. The second term adjusts for bias attributable to finite sample size and results in $p_{ij}$ having an expected value of zero for a population in Hardy-Weinberg equilibrium. For discrete distance intervals, mean values of $\hat{p}_{ij}$ were obtained by summarizing over all possible pairs of individuals located that distance apart. Results were combined over loci by weighting the result for each locus by its polymorphic index, $\sum p_i (1-p_i)$, to obtain a multilocus measure of spatial genetic structure. Weights for individual loci were also adjusted for differences in sample size due to missing genotypes.

Tests of significance for estimated values of $p_{ij}$ were performed by using randomization procedures to generate populations under the null hypothesis of no spatial genetic structure (see Slatkin and Arter, 1991). Occupied map locations were randomly assigned intact multilocus genotypes drawn at random with replacement from the sample population. For a given distance class, values of $p_{ij}$ from 399 simulation trials were ordered with $\hat{p}_{(2)}$ and $\hat{p}_{(398)}$ used to construct a 99% interval around the null hypothesis of no genetic structure. With $\hat{p}_{(data)}$ (the estimate based on the actual data) forming the 400th statistic, the null hypothesis of $p_{ij} = 0$ was rejected at the $\alpha = 0.01$ level when $\hat{p}_{(data)}$ was found to fall outside of this interval. Assuming that localized adaptation to micrographic conditions is not occurring, significant values of $\hat{p}_{(data)}$ are interpreted to indicate genetic structuring due to IBD resulting from nonrandom gene dispersal at the spatial scale examined. Such nonrandom dispersal presumably is due to localized pollen and/or seed movement.

Inbreeding coefficients, $F_{is}$, $F_{ir}$, and $F_{is}$ (Wright, 1951) were calculated using the methods of Weir and Cockerham (1984) with standard errors based on jackknifing procedures (Weir, 1990). $F_{is}$ is the probability of uniting genes being identical by descent relative to genes within...
that subpopulation. $F_{it}$ is the probability of uniting genes being identical by descent relative to genes within the entire population. Positive values of $F_{is}$ and $F_{it}$ indicate a deficit of heterozygotes and negative values indicate a surplus. For each polymorphic locus, a chi-square test was used to assess deviations from Hardy-Weinberg equilibrium: $\chi^2 = F^2N(a - 1)$, df = $(a - 1)/2$ where $F$ represents either $F_{is}$ or $F_{it}$, $N$ is the total sample size, and $a$ is the number of alleles at a locus (Li and Horvitz, 1953). To test whether mean values of $F_{is}$ and $F_{it}$ were significantly different from zero, we used the approximation of a 95% confidence level (1.96 times the standard error).

$F_{is}$ measures identity by descent for genes in different individuals within one subpopulation compared to the identity between populations. $F_{is}$ is often used as a measure of genetic differentiation among subpopulations. For each locus, statistical significance of $F_{is}$ values was based on the chi-square test, $\chi^2 = 2NF_{is}(a - 1)$ with $(a - 1)(s - 1)$ degrees of freedom, where $s$ is the number of subpopulations (Workman and Niswander, 1970). For the mean $F_{is}$ across loci, we tested whether mean values of $F_{is}$ were significantly different from zero by using the approximation of a 95% confidence interval (1.96 times the standard error). Assuming that genetic differentiation among subpopulations is the result of genetic drift in an island model of population genetic structure, we obtained an indirect estimate of average gene flow among subpopulations by calculating the effective number of migrants per generation ($N_{em}$) from individual and mean values of $F_{is}$ using the formula $F_{is} = 1/(4Nm + 1)$ (Weir and Cockerham, 1984; Slatkin, 1985; Slatkin and Barton, 1989).

### RESULTS

Genetic diversity measures indicate that all five populations of *P. officinalis* possess similar levels of genetic variation but population 4 has the least (Table 1). These levels are relative high compared to other woody plant species (Hamrick and Godt, 1989; Loiselle, Sork, and Graham, in press). Of the 14 putative loci scored, $\approx 50%$ were polymorphic within each subpopulation. Overall, nine of the 14 loci were polymorphic in at least one subpopulation. The mean number of alleles per locus was less than two, but as many as three alleles are observed for some loci (e.g., FES-1, GDH-1, and CE-1). The mean observed heterozygosity over loci ranged from 0.125 in population 4 to 0.198 in population 3 and expected heterozygosity ranged from 0.143 in population 4 to 0.218 in population 2.

Analysis of fine-scale genetic structure in population 1 indicates a significant positive autocorrelation among individuals located up to 5-m apart (Fig. 3). Coancestry values were not significantly different from zero beyond this closest distance class with the exception of a significant positive and negative correlation at 90 m and 105 m, respectively.

Analysis of genetic structure reveal that the estimated mean inbreeding coefficient, $F_{is}$, across loci within *P. officinalis* subpopulations was 0.055 which is significantly greater than 0 with a Type I error rate of 0.05 (Table 2). However, individual locus values show a great deal of variation. For example, 6PGD-1 shows a significant excess of heterozygotes while FES-1, LAP-1, LAP-2, and

---

**Table 1.** Summary of genetic diversity measures for 14 isozyme loci observed in five subpopulations of *Psychotria officinalis* at Estación Biológica La Selva.

<table>
<thead>
<tr>
<th>Locus</th>
<th>$F_{is}$</th>
<th>$F_{it}$</th>
<th>$F_{st}$</th>
<th>$N_{em}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDH-1</td>
<td>0.134*</td>
<td>0.143</td>
<td>0.113***</td>
<td>22.48</td>
</tr>
<tr>
<td>GDH-2</td>
<td>0.087</td>
<td>0.088</td>
<td>0.002</td>
<td>62.25</td>
</tr>
<tr>
<td>CE-1</td>
<td>0.090</td>
<td>0.102</td>
<td>0.013</td>
<td>18.98</td>
</tr>
<tr>
<td>PGI-2</td>
<td>0.101</td>
<td>0.121</td>
<td>0.023**</td>
<td>10.62</td>
</tr>
<tr>
<td>6PGD-1</td>
<td>-0.228***</td>
<td>-0.090</td>
<td>0.113***</td>
<td>1.96</td>
</tr>
<tr>
<td>6PGD-2</td>
<td>0.001</td>
<td>-0.004</td>
<td>-0.005</td>
<td>na</td>
</tr>
<tr>
<td>Weighted mean</td>
<td>0.055</td>
<td>0.144</td>
<td>0.095</td>
<td>2.37</td>
</tr>
</tbody>
</table>

*P* < 0.05; **P* < 0.01; ***P* < 0.001.
IDH-1 show a significant excess of homozygotes in a manner consistent with inbreeding (Table 2).

The analysis of macrogeographic genetic structure indicates that a significant amount of genetic differentiation occurs among subpopulations with a mean $F_{st}$ across loci of 0.095 which is significantly greater than 0 with a Type I error rate of 0.05 (Table 2). Six of the nine polymorphic loci had single-locus $F_{pis}$ estimates that were significantly greater than zero (Table 2). The significant estimates varied considerably across loci: values for LAP-1, LAP-2, and 6PGD-1 were 0.191, 0.351, and 0.113, respectively, while estimates for the remaining three significant loci were closer to zero. Values of $N_m$ vary across loci with a low of 0.71 migrants per generation for LAP-2 to a high of 62.25 for GDH. The mean value of $F_{st}$ across loci suggests a gene flow rate of 2.37 migrants per generation.

DISCUSSION

Our analysis of genetic structure within and among adult subpopulations of *Psychotria officinalis* indicates little IBD. Within a mapped subpopulation, we observed significant estimates of coancestry among individuals located within 5 m of each other but no significant estimates beyond that distance. Among subpopulations, we observed low but significant values of genetic differentiation. The pattern of fine-scale genetic structure that we observed within a single subpopulation might occur if seed dispersal were restricted such that either siblings or parent and offspring reside near each other. Similar spatial patterns have been found in a population of northern red oak, *Quercus rubra*, occurring in a temperate oak-hickory forest (Sork, Huang, and Wiener, 1993) and in an alpine population of the herbaceous monocarpic plant, *Ipomopsis aggregata* (Campbell and Dooley, 1992). All of these species are characterized by high outcrossing rates: *P. officinalis* and *I. aggregata* are self-incompatible (Bawa and Dooley, 1983; Campbell and Dooley, 1993) and *Q. rubra* rarely self-fertilizes (Sork, Huang, and Wiener, 1993). Thus, this pattern of genetic correlation among near neighbors compared to the subpopulation as a whole is observed in species with high outcrossing and presumably effective pollen flow. The high levels of gene flow through pollination would result in neighborhood areas much greater than the 5–10 m scale of autocorrelation reported for these studies. For this reason, we speculate that the relatedness among neighboring individuals is due to localized seed dispersal rather than isolation by distance resulting from localized dispersal of both pollen and seed.

It is interesting to examine our results in light of a study of tropical woody plant species in Panama (Hamrick, Murawski, and Nason, 1993). Although that study used a different measure of genetic similarity (mean number of alleles in common per loci, NAC) than a coancestry index, their method generates a comparable indicator of the spatial scale of genetic structure for two wind-dispersed canopy tree species, *Alseis blackiana* and *Platypodium elegans*, and for an understory treelet, *Swarzia simplex*. The analysis of NAC vs. distance class revealed positive genetic structuring was present in *Alseis* and *Platypodium* juveniles out to distances of 30 m and 100 m, respectively, but was lost in adults, presumably, due to random mortality during recruitment. *Swarzia*, on the other hand, was found to show genetic structuring in juvenile and adult size classes out to a distance of 10 m, similar in scale to that observed in *P. officinalis*, which is also a bird-dispersed understory species.

Results of our fine-scale spatial analysis contrast with those of Dewey and Heywood (1988) for a population of *Psychotria nervosa* in south Florida: they find no evidence of genetic structure on a scale of 1 m to 10 m using individual alleles of two isozyme loci. While their species is also self-incompatible, dispersal agents may have had different effects on the development of genetic structure. In addition, their population shows much less genetic variation than we find in ours. For example, they report only two loci out of 15 tested to be polymorphic while we find polymorphism at nine out of 14 putative loci tested. Thus, their analysis is based on individual single locus tests for five alleles using Moran's I as their spatial statistic while ours pools information from eight loci and 26 alleles to estimate a single population genetic structure statistic. Multilocus estimates provide a more powerful test for genetic structure than do single locus estimates (Heywood, 1991). The contrast in results for congeners with similar mating systems demonstrates the need to incorporate ecological observations and to evaluate genetic structure from a range of populations and species in order to gain more insight about patterns of genetic variation in plant populations.

Our results of spatial genetic structure analyses include the anomalous finding of a significant positive estimate of coancestry, $p_{ij}$, at 90-m distance class and a significant negative $p_{ij}$ at 105 m. One possible explanation for these observations is that they simply are due to chance. Out of 60 data points with 99% confidence limits, we might expect one observation to be significantly different from zero by chance. Unusually similar or dissimilar clusters of individuals located by chance at these distances apart could result in significant estimates of coancestry. Given that these two values are relatively small (both have an absolute value <0.02) and that there is little trend in the estimates at neighboring distance classes, we do not attribute a biological meaning to them.

Estimates of $F_{st}$ indicate low but significant differentiation among subpopulations with a value of $\approx$10% over relatively short distances (i.e., <2 km). This mean value of $F_{st}$ over loci results in an estimate of 2.37 effective migrants per generation. Slatkin (1985) suggests that $N_m$ values of from 1 to 4 should be sufficient to counteract the effects of genetic drift. It should be noted that we observed substantial variation in values of $F_{st}$ across loci. If genetic drift were solely responsible, we would expect all loci to be subject to similar genetic and experimental sampling variation. While the expected variation in $F_{st}$ across loci is unknown (Slatkin and Arter, 1991), some of the higher estimates (e.g., Lap-2) may reflect an effect of microgeographic selection on that locus. If so, then this effect would cause an overestimate of $F_{st}$ as an indicator of IBD. For example, when Lap-2 is removed, $F_{st}$ is not significantly greater than zero, which would suggest no IBD among subpopulations.

In spite of evidence suggesting low levels of IBD, considerable gene flow may be occurring. For example, within the 5-m interplant class, we observe coancestry esti-
mates of \(=0.05\). Assuming random mating, the coancestry estimator is a measure of the inbreeding coefficient between related individuals with an expected value of 0.25 for full-sibs and 0.125 for half-sibs. The observed coancestry value is much less than that expected for half-sibs, suggesting that although some proportion of seed dispersal may be localized, there is substantial mixing of seed shadows. This is consistent with observations of fruit removal by birds: some of the birds that disperse *P. officinalis* consume the fruit and drop seeds in the immediate vicinity of the parent while others take the fruit elsewhere (Loiselle, personal observation). The net result would be a mixing of the seedling pool with some inclusion of localized family clusters. To truly understand whether the observed spatial genetic structure is caused by seed dispersal, pollination mechanism, or patterns of seedling and juvenile survival, additional study is needed. An examination of the spatial structuring of juveniles would, for example, complement and enhance our understanding of the processes generating fine-scale genetic structure in adults. If restricted seed dispersal is important, then we would expect the estimates of coancestry to be even greater and perhaps over a larger spatial scale in juveniles than in adults. In general, nonrandom patterns of seed dispersal and seedling survival contributing to the development of spatial structure can best be identified by a combination of ecological and population genetic approaches.

**LITERATURE CITED**


