Conservation Biology

Contributed Paper

Efficacy of Land-Cover Models in Predicting Isolation of Marbled Salamander Populations in a Fragmented Landscape

KATHERINE R. GREENWALD,* H. LISLE GIBBS, AND THOMAS A. WAITE

Department of Evolution, Ecology, and Organismal Biology, Ohio State University, Columbus, OH 43210-1293, U.S.A.

Abstract: *Amphibians worldwide are facing rapid declines due to habitat loss and fragmentation, disease, and other causes. Where habitat alteration is implicated, there is a need for spatially explicit conservation plans. Models built with geographic information systems (GIS) are frequently used to inform such planning. We explored the potential for using GIS models of functional landscape connectivity as a reliable proxy for genetically derived measures of population isolation. We used genetic assignment tests to characterize isolation of marbled salamander populations and evaluated whether the relative amount of modified habitat around breeding ponds was a reliable indicator of population isolation. Using a resampling analysis, we determined whether certain land-cover variables consistently described population isolation. We randomly drew half the data for model building and tested the performance of the best models on the other half 100 times. Deciduous forest was consistently associated with lower levels of population isolation, whereas salamander populations in regions of agriculture and anthropogenic development were more isolated. Models that included these variables and pond size explained 65–70% of variation in genetically inferred isolation across sites. The resampling analysis confirmed that these habitat variables were consistently good predictors of isolation. Used judiciously, simple GIS models with key land-cover variables can be used to estimate population isolation if field sampling and genetic analysis are not possible.*

Keywords: *Ambystoma opacum*, geographic information systems, genetic isolation, habitat fragmentation, land cover, marbled salamander

Eficacia de los Modelos de Cobertura de Suelo para la Predicción del Aislamiento de Poblaciones de *Ambystoma opacum* en un Paisaje Fragmentado

Resumen: *A nivel mundial, los anfibios enfrentan declinaciones rápidas debido a la pérdida y fragmentación del hábitat, enfermedades y otras causas. Donde está implicada la alteración del hábitat, existe una necesidad de planes de conservación espacialmente explícitos. Los modelos construidos con sistemas de información geográfica (SIG) son utilizados frecuentemente para informar dicha planificación. Exploramos el potencial del uso de modelos SIG de conectividad funcional del paisaje como un sustituto confiable de medidas del aislamiento poblacional derivadas genéticamente. Utilizamos pruebas de asignación genética para caracterizar el aislamiento de poblaciones de* Ambystoma opacum *y evaluamos sí la cantidad relativa de hábitat modificado alrededor de las charcas de reproducción era un indicador confiable del aislamiento de la población. Mediante un análisis de remuestreo, determinamos sí ciertas variables de cobertura de suelo describían el aislamiento de la población consistentemente. Aleatoriamente extrajimos la mitad de los datos para construir el modelo y probamos el funcionamiento de los mejores modelos 100 veces en la mitad restante. El bosque deciduo se asoció consistentemente con niveles bajos de aislamiento poblacional, mientras que las poblaciones de salamandras estaban más aisladas en regiones agrícolas y con desarrollo antropogénico. Los modelos que incluyeron estas variables y el tamaño de la charca explicaron 65–70% de la variación en el aislamiento genético inferido. El análisis de remuestreo confirmó que estas variables de*

*email greenwald.35@osu.edu
Paper submitted August 18, 2008; revised manuscript accepted December 23, 2008.

Conservation Biology, Volume 25, No. 5, 1232–1241

DOI: 10.1111/j.1523-1739.2009.01204.x

hábitat consistentemente fueron buenos pronosticadores del aislamiento. Utilizados juiciosamente, los modelos SIG simples con variables de cobertura de suelo claves pueden ser utilizados para estimar el aislamiento de la población si no es posible efectuar muestreos en el campo ni análisis genéticos.

Palabras Clave: aislamiento genético, *Ambystoma opacum*, cobertura de suelo, fragmentación de hábitat, sistemas de información geográfica

Introduction

Because of the global decline of many amphibian taxa, an improved understanding of their dispersal and population structure is critical (Houlahan et al. 2000; Stuart et al. 2004). Amphibian populations can suffer declines and mortality due to habitat loss and degradation, overexploitation, and other enigmatic factors (Stuart et al. 2004). Habitat destruction can be especially pernicious for amphibians because of their generally limited dispersal ability and typical use of both aquatic and terrestrial (upland) habitats. Many amphibians presently exist in subpopulations that have been isolated, in many cases, by habitat destruction (Marsh & Trenham 2001). Habitat fragmentation and the resulting population isolation can have significant demographic consequences. Isolation can negatively influence survival, reproductive success, and dispersal processes (Hovel & Lipcius 2002; Berry et al. 2005; Lloyd et al. 2005). These demographic effects can, in turn, affect genetic structure of populations (Wauters et al. 1994; Andersen et al. 2004).

A number of methods have been developed to measure the extent to which populations are isolated. One of the most frequently used approaches is population genetic analysis, which involves estimates of long-term gene flow (measured as the numbers of migrants, *Nm*) from measures of population differentiation such as *F* statistics (Weir & Cockerham 1984). More recently genetic assignment tests have been used to measure short-term movement between populations (Berry et al. 2005; Zamudio & Wieczorek 2007). Results of many studies of amphibians in fragmented landscapes show substantial genetic differentiation on small spatial scales, which suggests that habitat fragmentation may lead to increased isolation among populations (Storfer 1999; Zamudio & Wieczorek 2007). Furthermore, some researchers have found a relationship between habitat features and genetic differentiation in amphibian populations. For example, there is greater genetic differentiation in populations separated by roads, development, and agriculture than those in undisturbed areas (Gibbs 1998; Noel et al. 2007).

Spatially explicit analyses conducted with geographic information systems (GIS) (Gustafson et al. 2001; Compton et al. 2007) can be used to predict movement and habitat connectivity, and hence degree of isolation, among populations in fragmented landscapes. In many cases data (e.g., land cover) necessary for such analyses are freely available. A few researchers explicitly compared more direct estimates of isolation based on genetic results with more indirect measures based on land-cover data (e.g., Spear et al. 2005; Cushman 2006; Epps et al. 2007).

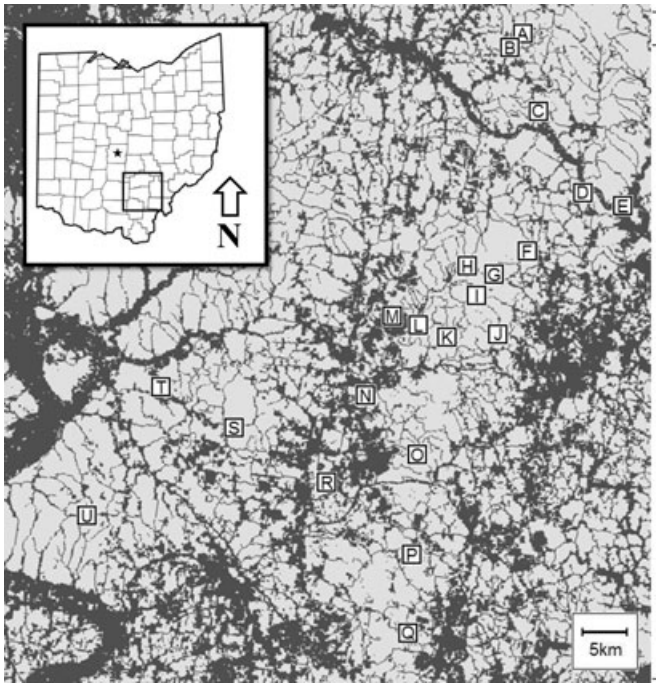
Such studies would be especially valuable if they were to show that land-cover data are an appropriate proxy for measures of isolation based on genetic data. Such predictive ability would be valuable in prioritization of conservation action, whether the goal was to preserve well-connected populations or to connect formerly isolated sites. Land-cover studies might then stand in for more expensive molecular analyses if resources were too limited to conduct both (Scribner et al. 2001; Storfer et al. 2007).

Here we used both GIS methods and genetic data to study population isolation in subpopulations of the marbled salamander (*Ambystoma opacum*). We explored whether simple land-cover models can serve as efficient substitutes for genetically based metrics of population isolation. Broadly speaking we used genetic data to describe regional population isolation and GIS data to evaluate anthropogenic habitat modifications as barriers to movement. Our goals were to (1) quantify, with genetically based measures (i.e., assignment tests), the degree to which marbled salamander populations in Ohio (U.S.A.) are isolated, (2) determine which land-cover types most strongly influenced the levels of isolation, and (3) evaluate the efficacy of the use of GIS models as surrogates for analyses of functional landscape connectivity based on genetic data. The most thorough approach to developing a conservation plan involves studying both landscape and genetic data. Nevertheless, if resources are limited, we suggest that land-cover data may be a cheap, quickly executed substitute for genetically based measures of determining population isolation in conservation planning efforts.

Methods

Sampling and Genetic Analyses

We collected tissue (tail tip) samples (March–June 2005) from marbled salamanders in 21 ponds in southeastern Ohio. A core group of ponds (F–K; Fig. 1) were in a well-forested area consisting of the Wayne National Forest and



Site	Residents	*n*
A	13	25
B	9	15
C	21	24
D	8	12
E	10	10
F	9	25
G	15	25
H	16	25
I	12	25
J	5	25
K	10	21
L	4	25
M	9	14
N	18	22
O	9	25
P	17	25
Q	12	25
R	17	25
S	22	25
T	15	25
U	19	25

*Figure 1. Locations of ponds (letters) in Ohio that were sampled for marbled salamanders (*Ambystoma opacum*) on which population genetic analyses were conducted (light gray, forest; dark gray, agriculture and roads). The number of salamanders assigned as residents and sample sizes (*n*) are shown on right for each pond.*

Zaleski State Forest. The remaining ponds were in areas closer to development and agriculture. We used dip nets and funnel traps to catch larvae in ponds (468 individuals captured). We spaced traps as far from each other as possible around pond perimeters to minimize the probability of sampling siblings. The distance between traps ranged from approximately 5 m in the smallest ponds to approximately 30 m in the largest. To further minimize relatedness of sampled larvae, we haphazardly sampled only one individual per trap when possible. All tissue collection was nondestructive, and the larvae were released at the site of capture. The ponds were located across a wide range of habitat degradation levels, with some in highly forested areas and others in regions that had undergone agricultural and residential development. Our animal use protocol was approved by The Ohio State University Institutional Animal Care and Use Committee (permit numbers 2004A0067 and 2004A0159). We stored tissue samples in a −20 °C freezer in 95% ethanol.

We used Qiagen DNeasy kits (Qiagen, Valencia, California) to extract DNA from tail tissue. The individuals were genotyped at nine microsatellite loci with locus-specific primers and the PCR protocol described by Croshaw et al. (2005) (loci listed in Table 1). We used fluorescent-tagged forward primers to visualize the alleles. We scored the alleles with an Applied Biosystems 3100 automated sequencer and GeneMapper 3.2 software (Applied Biosystems, Foster City, California). We

Table 1. Characteristics of microsatellite loci used to describe isolation in marbled salamander (*Ambystoma opacum*) populations in southeastern Ohio (U.S.A.).[a]

Locus	*Number of alleles*	*Allele size range (bp)*	H_e	H_o	*Repeat type*	F_{ST}[b]
Aop7[c]	25	102–204	0.654	0.204 (0.614)	di	0.033 (0.031)
Aop31	24	130–252	0.757	0.768	tetra	0.083
Aop36	22	101–263	0.738	0.723	tetra	0.048
Aop37[c]	15	197–257	0.616	0.150 (0.611)	di	0.126 (0.101)
AmaD42	23	105–209	0.863	0.872	tetra	0.058
AmaD95	27	140–280	0.864	0.878	tetra	0.052
AmaD328	15	310–366	0.815	0.816	tetra	0.049
AjeD23	18	187–275	0.858	0.872	tetra	0.044
AjeD162	23	120–252	0.862	0.878	tetra	0.057

[a] *Key:* H_e, *expected heterozygosity;* H_o, *observed heterozygosity; repeat type refers to whether the microsatellite was a dinucleotide (di) or a tetranucleotide (tetra) repeat.*

[b] *Overall across all populations.*

[c] *Locus had null alleles. Observed heterozygosity values were adjusted in Micro-Checker, and* F_{ST} *values were recalculated (corrected values in parentheses for both).*

used GENEPOP (Raymond & Rousset 1995) to calculate observed and expected heterozygosities and test for a relationship between genetic and geographic distances (isolation by distance [IBD]); FSTAT (Goudet 1995) to calculate overall and pairwise *F*-statistics, and Micro-Checker (Van Oosterhout et al. 2004) to look for evidence of null (nonamplifying) alleles, which result in a deficiency of heterozygotes. The latter program revealed two loci (Aop7 and Aop 37) that showed evidence of null alleles (GENEPOP test for heterozygote deficiency: $p < 0.001$ in both cases). We used the method described by Chakraborty et al. (1992) implemented in Micro-Checker to correct for these null alleles. This procedure provides adjusted genotypes in which a proportion of homozygotes are changed to heterozygotes to restore Hardy–Weinberg equilibrium. We randomly assigned individuals in the data set to represent these putative heterozygotes, creating a dummy allele to substitute for one of the original homozygous alleles. The results were qualitatively the same with this method versus the method in which we eliminated the two loci with null alleles. We present corrected data.

For assignment tests, we used the "detection of first generation migrants" function of GENECLASS2 (Piry et al. 2004) because our sampled individuals were premigratory offspring. Individuals are assigned either as residents or as immigrants. For immigrants, some can be assigned to source populations with high likelihood, whereas others cannot. For our analysis, very few individuals could be assigned as immigrants with high likelihood (see Results). Our response variable was thus the number of individuals assigned as residents with $\geq$95% likelihood at each site. Because this response metric is dichotomous (individuals are assigned to discrete groups, "residents with high likelihood" or "other"), we used a binomial model to analyze the data (see below).

We used the Bayesian analysis of population structure (BAPS 2.2; Corander et al. 2004) to corroborate our assignment test results. Our goal was to determine whether populations identified as highly isolated by the assignment test were similarly identified as isolated by a population-level analysis. We used a burn-in value of 10,000 with 50,000 iterations (Corander et al. 2003, 2004).

GIS and Statistical Analyses

To examine relationships between the surrounding landscape and genetically based measures of isolation, we determined land-cover composition within circular areas centered on each of the 21 sample sites. We performed these analyses in areas with radii of 300 m and 1 km. The 300-m buffer was chosen to encompass frequently utilized upland habitat in this and related species (Semlitsch & Bodie 2003). Dispersal events exceeding 1 km have been documented (Gamble et al. 2007) and may be more common than previously reported (Smith & Green 2005). Analysis at this larger scale allowed us to account for the possibility of such long-distance dispersal. We were also able to evaluate the effect of additional land-cover types because many more habitat categories were present within the 1-km buffer than within the 300-m buffer. As a base layer we used the 30-m resolution National Land Cover Data (NLCD) from 2001 along with a more detailed roadmap overlay, both available from the U.S. Geologic Survey (http://seamless.usgs.gov/). We scored land-cover composition as the proportion of each buffer comprising each land-cover category.

We used two methods to reduce the 15 NLCD land-cover variables for analysis, which allowed us to assess whether a "biologically informed" data-reduction method had better predictive ability than a "biologically blind" method. For the former, we used published resistance values derived from expert surveys (Compton et al. 2007). These resistance values represent the permeability to salamander movement and allowed us to pool land-cover types based on how hospitable biologists expect them to be. Most NLCD land-cover types were easily translated to the categories used by Compton et al. (2007) (Table 2). We used only their low- and high-intensity residential values, rather than urban values, because our study area encompassed only small towns. We also equated our grassland-herbaceous category with their pasture category. Where the grassland-herbaceous category was prominent, it corresponded to open, short-grass (mown) fields and thus appeared more similar to a pasture than to an old-field habitat. Because barren land is folded into open land-old field habitat in Compton et al.'s (2007) data set, we treated it the same way (B. Compton, personal communication).

The expert-assigned resistance values then fell into natural categories named for their dominant land-cover type: FOR, deciduous forest, mixed forest, evergreen forest, and woody wetlands (all with resistance of 1.0); SHRUB, emergent herbaceous wetlands, scrub-shrub, and barren land (all with resistance of 3.0–3.4); LODEV, developed at low intensity and developed at medium intensity (both with resistance of 6.8); AG, grassland-herbaceous, pasture-hay, and cultivated crops (all with resistance of 9.2–10.2); and HIDEV, developed at medium intensity and developed at high intensity (both with resistance of 12.6). We included water as a variable on its own because of its very high resistance value (22.0). We began with all these categories but removed those that were too rare or sparse to include in the analyses (SHRUB and HIDEV at the 300-m scale and HIDEV at the 1-km scale).

Because we wanted to compare this biologically informed method of data reduction with one that did not rely on a priori knowledge, we used factor analysis to combine the original NLCD variables a second time. We retained as variables each factor score (FS) with an eigenvalue >1, which left us with four scores at

Table 2. National Land Cover Data (NLCD) categories used in our study of marbled salamander (*Ambystoma opacum*) population isolation and the comparable categories used in Compton et al. (2007).*

NLCD category	*Compton et al. category*	*Resistance*	*Group*
Deciduous forest	forest	1.0	low
Mixed forest	forest	1.0	low
Evergreen forest	forest	1.0	low
Woody wetlands	forest-vernal pool	1.0	low
Emergent herbaceous wetlands	nonforested wetland	3.0	medium-low
Scrub-shrub	old field	3.4	medium-low
Barren land (rock/sand/clay)	open land-old field	3.4	medium-low
Developed, low intensity	low-intensity, residential	6.8	medium
Developed, open space	low-intensity, residential	6.8	medium
Grassland-herbaceous	pasture	9.2	medium-high
Pasture-hay	pasture	9.2	medium-high
Cultivated crops	row crop	10.2	medium-high
Developed, medium intensity	high-intensity, residential	12.6	high
Developed, high intensity	high-intensity, residential	12.6	high
Open water	pond-lake	22.0	NA

**We matched comparable land-cover categories to "translate" the NLCD categories into the numerical resistance values developed by Compton et al. (2007), shown in the Resistance column. We then grouped variables based on these resistance values (i.e., Group).*

each spatial scale (Kaiser 1960). We used these scores (FS1–FS4) as predictors in the model selection analysis so that we could compare them with the biologically informed land-cover models described above. Loadings for each retained FS on each land-cover variable are shown in Table 3.

We used Akaike's information criterion corrected (AIC_c) for small sample size (Burnham & Anderson 2002) to select between candidate models that best explained levels of isolation across the 21 sites. The response variable was the number of individuals assigned as residents at each site with ≥95% likelihood. We used a binomial model with a logit link function to account for the different sample sizes across sites and the dichotomous nature of the response variable. We included three models to account for the inherent characteristics of the ponds and their locations: the perimeter of the pond at maximum pond volume (PP), average pond distance (APD, mean distance from focal pond to all other ponds in the study area), and nearest pond distance (NPD, distance to the pond closest to the focal pond). The latter two models were included to avoid confounding geographic distance between ponds with habitat variation. The average pond distance was lowest for the "core" group of ponds located in the central part of the study area (Fig. 1). The remaining models paired each land-cover variable (both expert- and FS-derived) with pond perimeter because we anticipated that pond size would have a strong influence on our isolation metric. Therefore, at each spatial scale, we tested the following models: PP, APD, NPD, FOR + PP, SHRUB

Table 3. Factor analysis for all National Land Cover Data (NLCD) categories within 300 m and 1 km of marbled salamander (*Ambystoma opacum*) breeding ponds in a study of population isolation.*

	300-m scale				*1-km scale*			
NLCD category	*FS1*	*FS2*	*FS3*	*FS4*	*FS1*	*FS2*	*FS3*	*FS4*
Open water	0.560	0.025	0.190	0.517	0.302	0.606	0.696	0.028
Developed, open space	−0.139	0.767	0.519	−0.026	0.692	0.353	−0.294	0.098
Developed, low intensity	0.050	0.848	−0.223	0.245	0.949	0.065	−0.112	0.135
Developed, medium intensity	NA	NA	NA	NA	0.942	0.046	−0.058	0.149
Barren land (rock/sand/clay)	NA	NA	NA	NA	0.000	0.695	0.656	−0.220
Deciduous forest	−0.900	−0.269	0.071	−0.206	−0.912	0.224	−0.203	−0.115
Evergreen forest	−0.064	0.042	−0.836	0.127	−0.272	−0.026	0.110	0.884
Scrub-shrub	−0.157	−0.053	0.512	0.025	0.356	−0.024	−0.208	−0.041
Grassland-herbaceous	0.880	−0.169	0.139	−0.129	0.928	0.116	−0.094	0.116
Pasture-hay	0.899	−0.112	0.053	−0.380	0.629	−0.522	0.204	−0.283
Cultivated crops	0.105	−0.286	−0.030	0.740	0.242	−0.725	0.416	−0.290
Woody wetlands	0.270	0.302	−0.369	−0.374	−0.126	−0.549	0.549	0.433
Total variance explained (%)	0.284	0.160	0.148	0.124	0.393	0.179	0.137	0.105

**There were four retained factor scores (FS) (eigenvalue >1) at each scale. Variance explained by each FS is in the bottom row; they sum to 71.5% and 81.4% for the 300-m and 1-km scale, respectively (NA, land-cover variables not present at the smaller spatial scale).*

+ PP (1-km scale only), LODEV + PP, AG + PP, FS1 + PP, FS2 + PP, FS3 + PP, and FS4 + PP. We also included the model WATER + PP because of water's very high expert-assigned resistance value. Finally we included a ROADS + PP model composed of the total length of roads within the buffer. Following Burnham and Anderson (2002), we consider models with $\Delta AIC_c < 2$ as good models (strongly supported even though another model was selected).

To evaluate the predictive value of land-cover models, we randomly chose 11 sites for a second AIC_c-based validation analysis. We refer to these 11 as model sites because we used them in the model-building step. Once we determined the best model based on these sites, we tested it on the remaining (unselected) 10 sites. We refer to these 10 sites as test sites because we used them to test the model. This process was repeated 100 times, with a different random drawing of 11 model sites each time. Thus there were 50 runs for the 300-m scale and 50 runs for the 1-km scale. We evaluated model performance by summing the number of iterations in which a given model was selected and the number of times it was a strong candidate ($\Delta AIC_c < 2$). We also noted the sign of the regression coefficient for each land-cover variable in each run. We compared the number of times the regression coefficient was positive to the null expectation (25/50 positive runs) with a chi-square test. Finally we used r^2 as the basis for evaluating predictive value (Spear et al. 2005). To do this we used the regression coefficients generated using the model sites to produce predicted resident assignment values for the test sites. We looked at these predicted values versus the observed assignment data to generate an r^2 for each model with $\Delta AIC_c < 2$. All statistical routines were run with SPSS 16.0 (SPSS, Chicago, Illinois).

Results

Genetic Analyses

The microsatellites were highly polymorphic, with 15–27 alleles per locus (Table 1). Expected and observed heterozygosities were high and within the range of Hardy–Weinberg equilibrium for seven of nine loci. At the other two loci, the frequency of null alleles (r, with the methods of Chakraborty et al. 1992) was 0.524 and 0.608 for Aop7 and Aop37, respectively. Correcting for null alleles resulted in only minor changes in F_{ST} (e.g., original overall $F_{ST} = 0.058$, corrected overall $F_{ST} = 0.056$). We compared assignment test results generated with corrected data to those generated with only the seven loci without null alleles and the results were qualitatively identical. The results presented here are from corrected allele frequencies.

We observed high levels of variation in pond isolation. Pairwise F_{ST} values ranged from 0.002 (sites K–L) to 0.217 (sites C–E; Fig. 1), although there was no significant IBD effect ($p = 0.12$). Few individuals (6% of putative migrants) were assigned as immigrants with known source ponds, likely because we did not collect samples from the actual source pond in these cases or because our sampled individuals were offspring (offspring of one migrant and one resident may not be definitively assigned to a source population). In contrast, individuals assigned as residents tended to be assigned with much higher certainty (72% [229/319] were assigned with ≥95% likelihood). We used these resident assignments as our genetically based measure of population isolation. Assignment as residents ranged from 16% (4 of 25 individuals assigned as residents with ≥95% likelihood; site L) to 100% (10 of 10 individuals assigned as residents with ≥95% likelihood; site E; Fig. 1).

The results from the BAPS corroborated our assignment test results, although they provided much less resolution in regard to isolation of local populations. Three sites (C, E, and S in Fig. 1) clustered individually as their own populations, whereas all other sites clustered together in one large population. This was the only likely partition and had an associated probability of 1. The three separated sites had the highest observed proportion of individuals assigned as residents in the assignment test analysis (0.88, 1.00, and 0.88, respectively). The BAPS results thus complimented our assignment test results; however, BAPS could not resolve the structure for sites with low-to-moderate levels of isolation as determined by assignment tests. We are unsure as to the reason for this difference in resolution between the methods but speculate that it may be due to the fact that the BAPS method is attempting a more statistically demanding task (assigning each individual to one of n possible clusters), whereas the assignment tests are simply estimating the probability that a given individual is a resident of a single population. We discuss the remainder of our land-cover analyses in reference to the assignment test results only.

Model Selection

At the 300-m scale for all 21 sites, the model with the most support was AG + PP, with a model weight (w_i) of 0.297 (Table 4). Over 70% of the variation in the observed data were explained by these 2 predictors ($r^2 = 0.704$; Fig. 2). Two other models were also strongly supported ($\Delta AIC_c < 2$): FS1 + PP ($w_i = 0.285$) and FOR + PP ($w_i = 0.164$). FS1 at the 300-m scale summarizes both agriculture and forest variables; their inverse relationship is reflected in the sign of the loading values. This FS was driven by high loadings for grassland-herbaceous (0.880) and pasture-hay (0.899) and by a strong negative loading for deciduous forest (−0.900). Loadings for all other land-cover variables were much lower (Table 3).

At the 1-km scale for all 21 sites, FS1 + PP was the best model ($w_i = 0.272$; Table 4). FS1 at this larger scale

Table 4. Akaike's information criterion (AIC) model selection results for 300-m and 1-km land-cover buffers around marbled salamander (*Ambystoma opacum*) breeding ponds.[a]

Model	K	AIC_c	ΔAIC_c[b]	w_i	r^2
300-m scale					
AG + PP	4	104.36	0.00	0.297	0.704
FS1 + PP	4	104.44	0.08	0.285	0.702
FOR + PP	4	105.55	1.19	0.164	0.688
WATER + PP	4	114.98	10.62	0.001	
PP	3	115.04	10.68	0.001	
ROADS + PP	4	115.09	10.73	0.001	
FS3 + PP	4	116.53	12.17	0.001	
LODEV + PP	4	116.71	12.35	0.001	
FS2 + PP	4	117.04	12.68	0.001	
FS4 + PP	4	129.42	25.06	0.000	
NPD	3	148.12	43.76	0.000	
APD	3	148.94	44.58	0.000	
1-km scale					
FS1 + PP	4	108.40	0.00	0.272	0.653
FOR + PP	4	108.93	0.53	0.208	0.664
AG + PP	4	110.94	2.54	0.077	
PP	3	115.04	6.64	0.010	
LODEV + PP	4	115.23	6.83	0.009	
ROADS + PP	4	115.47	7.07	0.008	
WATER + PP	4	115.49	7.09	0.008	
SHRUB + PP	4	115.56	7.16	0.008	
FS3 + PP	4	115.78	7.38	0.007	
FS2 + PP	4	116.68	8.28	0.004	
FS4 + PP	4	117.06	8.66	0.004	
NPD	3	148.12	39.72	0.000	
APD	3	148.94	40.54	0.000	

[a]*The number of individual salamanders assigned as pond residents is the response variable* (n = *21 sites). Key: FOR, expert grouping, low resistance; SHRUB, expert grouping, medium-low resistance; LODEV, expert grouping, medium resistance; AG, expert grouping, medium-high resistance; FS, factor score; PP, pond perimeter;* w_i, *model weight.*

[b]*Models with* $\Delta AIC_c < 2$ *are considered good models;* r^2 *values are shown for these models only.*

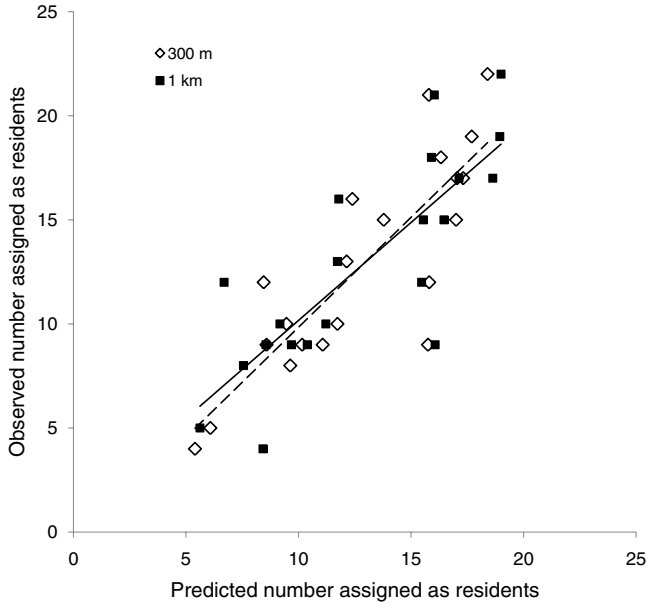


Figure 2. Prediction of the best model of marbled salamander population isolation (as chosen by Akaike's information criterion) versus observed genetic data for the number of salamanders at each site assigned as residents. The best model for all 21 sites at the 300-m scale includes agriculture and pond perimeter and describes 70.4% of the variation in the observed data (open diamonds, dashed line). The best model for all 21 sites at the 1-km scale includes factor score 1 and pond perimeter and describes 65.3% of the variation in the observed data (filled squares, solid line).

was driven by development and agriculture and again reflected the inverse relationship that these land-cover types had with forest area. This FS had high loadings for low- and medium-intensity development (0.949 and 0.942, respectively) and grassland-herbaceous (0.928) and intermediate loadings for open-space development and pasture-hay (0.692 and 0.629, respectively). Additionally, it had a strong negative loading for deciduous forest (−0.912; Table 3). FS1 and pond size explained over 65% of the variance in observed resident assignment (Fig. 2). There was only one other good model at the 1-km scale: FOR + PP ($w_i = 0.208$).

Model Validation

For the 300-m scale, AG + PP was again the most strongly supported model. It was the best model 21 times (42% of the time) and had $\Delta AIC_c < 2$ an additional 16 times (a good model 37 times [74%]). It also had the highest mean weight ($w_i = 0.229$) and the second-highest mean r^2 for predicted versus observed resident assignments ($r^2 = 0.654$). In addition, the regression coefficient for AG was positive in 100% of the runs, meaning that increased AG was consistently associated with higher isolation. The model with the second-greatest support at the 300-m scale was FS1 + PP. This was the best model 13 times (26%), was a good model 30 times (60%), and had a positive regression coefficient 100% of the time. The third-ranked model, FOR + PP, was the best model 8 times (16%) and was a good model 28 times (56%). The regression coefficient for FOR was negative 50 out of 50 times, meaning that increased FOR was consistently associated with lower isolation. Beyond that, there was a sharp drop in model success; the remaining models were the best 0–2 times each (Table 5). Nevertheless, four other models had regression coefficients more frequently either positive or negative than expected by chance: PP (100% negative), WATER (88% positive), APD (98% positive), and NPD (98% positive).

At the 1-km scale, FS1 + PP was the most strongly supported model (Table 5). It was the best model 17 times (34%) and a good model 25 times (50%). It also had the

Table 5. Akaike's information criterion (AIC) model selection results for validation analysis of the effect of land-cover within 300 m and 1 km of breeding ponds on marbled salamander (*Ambystoma opacum*) population isolation.[a]

Model	*Sign of β[b] (pos/neg)*	*Mean* w_i	*Best[c] (pos/neg)*	*Good[c] (pos/neg)*	*Total[c]*	*Pseudo* r^2 *(range)[d]*
300-m scale						
AG + PP	50/0*	0.229	21 (21/0)	16 (16/0)	37	0.654 (0.275–0.945)
FS1 + PP	50/0*	0.177	13 (13/0)	17 (17/0)	30	0.689 (0.405–0.874)
FOR + PP	0/50*	0.147	8 (0/8)	20 (0/20)	28	0.576 (0.264–0.861)
PP	0/50*	0.048	2 (0/2)	13 (0/13)	15	0.617 (0.381–0.789)
ROADS + PP	39/11	0.055	2 (2/0)	10 (7/3)	12	0.454 (0.002–0.800)
LODEV + PP	31/19	0.050	1 (1/0)	11 (6/5)	12	0.426 (0.113–0.830)
WATER + PP	44/6*	0.033	2 (2/0)	5 (5/0)	7	0.616 (0.426–0.772)
FS2 + PP	21/28	0.033	1 (0/1)	6 (3/3)	7	0.520 (0.126–0.824)
FS3 + PP	36/14	0.025	0	5 (3/2)	5	0.513 (0.304–0.680)
FS4 + PP	31/19	0.020	0	2 (2/0)	2	0.537 (0.419–0.655)
APD	49/0*	0.002	0	1 (1/0)	1	0.125
NPD	49/1*	0.002	0	0	0	NA
1-km scale						
FS1 + PP	48/2*	0.192	17 (17/0)	8 (8/0)	25	0.422 (0.103–0.750)
FOR + PP	0/50*	0.134	4 (0/4)	19 (0/19)	23	0.524 (0.177–0.794)
LODEV + PP	39/11	0.163	11 (11/0)	10 (7/3)	21	0.134 (0.001–0.604)
ROADS + PP	39/11	0.088	3 (3/0)	10 (6/4)	13	0.225 (0.004–0.574)
AG + PP	50/0*	0.071	3 (3/0)	9 (9/0)	12	0.606 (0.338–0.795)
FS3 + PP	45/4*	0.043	5 (5/0)	4 (4/0)	9	0.488 (0.233–0.690)
PP	0/50*	0.033	1 (0/1)	7 (0/7)	8	0.605 (0.360–0.761)
SHRUB + PP	42/8*	0.031	2 (2/0)	3 (3/0)	5	0.580 (0.376–0.726)
FS2 + PP	7/42*	0.027	2 (0/2)	3 (0/3)	5	0.371 (0.089–0.611)
WATER + PP	44/6*	0.034	1 (1/0)	3 (3/0)	4	0.442 (0.341–0.596)
NPD	50/0*	0.009	1 (1/0)	0	1	0.001
FS4 + PP	22/28	0.015	0	1 (0/1)	1	0.554
APD	49/1*	0.002	0	0	0	NA

[a] *The number of individual salamanders assigned as pond residents is the response variable. Key: FOR, expert grouping, low resistance; SHRUB, expert grouping, medium-low resistance; LODEV, expert grouping, medium resistance; AG, expert grouping, medium-high resistance; FS, factor score; PP, pond perimeter;* w_i, *model weight, averaged for each model across all 50 runs.*

[b] *The number of times each land-cover variable was positive versus negative in the regression equation across all 50 runs. Instances that do not sum to 50 are due to models with* $\beta = 0$. *Asterisks denote ratios significantly different from random expectation (25/50) following the Bonferroni correction.*

[c] *Best is the number of times each model was selected as the best model; good is the number of times it was a good model* ($\Delta AIC_c < 2$). *Direction of the relationship is in parentheses (pos/neg). Total is the sum of best runs and good runs for each model.*

[d] *The relationship between predicted and observed data. Predicted values were generated for the 10 nonselected (test) sites, with coefficients based on the 11 model-building (model) sites. Pseudo* r^2 *therefore gives a measure of the predictive value of the model.*

highest mean model weight ($w_i = 0.192$), but not the highest mean r^2 ($r^2 = 0.422$). It had a positive regression coefficient 96% of the time. The model with the next greatest support was FOR + PP. It was best four times, with 19 instances of $\Delta AIC_c < 2$, negative regression coefficients 100% of the time, and high predictive ability ($r^2 = 0.524$). Although LODEV + PP was selected as the best model the second-highest number of times (11), it was a poor predictor ($r^2 = 0.134$) and the split on regression coefficient sign frequency was not significant (78% positive). Other models were selected as the best model five or fewer times (Table 5), although many models had regression coefficients more frequently either positive or negative than expected by chance. Significantly positive regression coefficients were obtained for AG (100% of the time), FS3 (90%), SHRUB (84%), WATER (88%), NPD (100%), and APD (98%). Significantly negative regression coefficients were obtained for PP (100%) and FS2 (84%).

Discussion

Our most important finding was that limited, easily obtained geographic information can explain much of the variation in a genetically inferred measure of population isolation in marbled salamanders. Agriculture was the strongest individual driver of isolation at both scales. This category was composed of grassland-herbaceous, pasture-hay, and cultivated-crop land cover. Cultivated crops were much less common in this area than the other two categories, so the key contributor to isolation at this scale was open-field habitat. Agriculture had heavy loadings on the very successful predictor FS1 at both scales, was consistently associated with increased isolation, and had good predictive ability in the validation analysis. Our results complement recent work showing that open-field habitat is a strong deterrent to movement in the related spotted salamander (*Ambystoma maculatum*)

(Rothermel & Semlitsch 2002). The importance of developed areas varied; LODEV (i.e., developed categories) and road density had a much stronger influence at the large spatial scale. This is logical because these categories were relatively infrequent within 300 m of the ponds. Development was not as good a predictor as agriculture, with variations in the sign of the regression coefficient and low r^2 values. Nevertheless, development categories were also an integral part of FS1 at the larger scale, which was an excellent predictor of isolation.

Our results strongly indicate that the marbled salamander is susceptible to negative repercussions from habitat modification. The assignment tests showed that the least-isolated populations were located in the continuously forested regions within the Zaleski State Forest. The low levels of isolation within the forest suggest that in a pristine habitat, our sites would constitute a patchily distributed population, rather than a true metapopulation (Hanski & Gilpin 1997; Trenham et al. 2001). Connectivity among breeding ponds may minimize local extinction rates. Nevertheless, at sites that are more affected by anthropogenic factors (especially agriculture), local breeding populations would be much more isolated and might more closely approximate a metapopulation. In these cases the probability of recolonization following extinction is presumably reduced, and in some cases (*e.g.*, sites C and E), such a rescue effect may be virtually impossible. Small, isolated populations also may be subject to negative effects due to inbreeding. Such effects have been documented and associated with certain landscape features in amphibian taxa (e.g., proximity to roads; Lesbarrères et al. 2003). Inbreeding can have negative effects on survival (*Rana sylvatica*; Halverson et al. 2006) along with other components of fitness, such as growth rate and metamorph production (*Bufo calamita*; Rowe & Beebee 2002). Some of our sampled populations were quite small and isolated and thus may be at risk of negative effects of inbreeding.

Our data also allowed us to evaluate the effect of pond size and isolation on our genetically inferred measure of population isolation. Larger ponds had fewer individuals assigned as residents and hence more individuals assigned as immigrants. Migrants may select larger ponds preferentially, or the pattern may arise haphazardly simply because larger ponds make larger targets. In either case management action should focus on maintaining the quality of larger breeding ponds because they have the potential to become dangerous sinks in the event that habitat degradation makes successful reproduction impossible (Battin 2004). Nevertheless, the ponds that we considered large were still relatively small; no pond in the data set was >300 m in circumference. Neither measure of pond isolation (APD and NPD) fared well in the AIC analysis (selected as the best model zero times and as a good model two times). Although these variables lacked power in describing isolation, they were highly consistent. In both cases larger distances between the ponds were nearly always associated with higher levels of isolation. The inclusion of these models allowed us to rule out the possibility that land-cover variables and pond isolation were confounded (i.e., the possibility that deciduous forest is not really an important predictor but rather that ponds located in forested areas are simply closer together). The only variables significantly correlated with either APD or NPD were WATER and FS3, neither of which was a meaningful predictor of isolation.

Grouping land-cover variables on the basis of expert knowledge was slightly preferable to blind FS-based groupings. Although FS1 did quite well at both spatial scales, the expert groupings were more frequent in the best and good models. At the 300-m scale the expert groupings were in the best model 30 times and were in the good models 77 times, versus 14 and 44, respectively, for all FSs. At the 1-km scale the expert groupings were in the best model 20 times and were in the good models 61 times, versus 24 and 40, respectively, for all FSs. For future predictive work we recommend the inclusion of variables equivalent to pond perimeter, FOR (driven by deciduous forest), AG (driven by agriculture), and either APD or NPD to rule out confounding of the importance of land cover with that of pond isolation.

In our study system landscape and genetically based measures of population isolation were strongly connected, which suggests that land-cover data could substitute for genetic data in assessments of functional landscape connectivity. Our results suggest that movement by potential migrants is impeded by open fields and impervious anthropogenic features (development and roads) and that it is aided by forest. These variables along with pond size explained nearly three-quarters of the variation in isolation based on genetic data. We propose that conservation planners may sometimes be justified in using land-cover data to estimate population isolation in the absence of genetic data. Conservation resources are generally limited, and genetic analyses are time-consuming and expensive. Predictive models including key land-cover variables may allow managers to prioritize actions, whether the goal is to protect populations with currently high levels of connectivity or to rescue those at risk of isolation.

Acknowledgments

We thank E. Greenwald, B. Horne, A. Vitz, and M. Bakermans for help and logistics in the field. J. Diaz gave essential assistance with laboratory work. We thank the Gibbs laboratory group, B. Groff, and three anonymous reviewers for valuable feedback on earlier drafts of the manuscript. We thank the Olentangy River Wetland Research Park, The Ohio State University (OSU) Graduate School's Alumni Grants for Graduate Research and

Scholarship Program, the Muelbach Memorial Fund, and especially the Columbus Zoo/OSU Cooperative Research Grant program for funding.

Literature Cited

Andersen, L. W., K. Fog, and C. Damgaard. 2004. Habitat fragmentation causes bottlenecks and inbreeding in the European tree frog (*Hyla arborea*). Proceedings of the Royal Society of London Series B-Biological Sciences **271:**1293–1302.

Battin, J. 2004. When good animals love bad habitats: ecological traps and the conservation of animal populations. Conservation Biology **18:**1482–1491.

Berry, O., M. D. Tocher, D. M. Gleeson, and S. D. Sarre. 2005. Effect of vegetation matrix on animal dispersal: genetic evidence from a study of endangered skinks. Conservation Biology **19:**855–864.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoritic approach. 2nd edition. Springer-Verlag, New York.

Chakraborty, R., M. De Andrade, S. P. Daiger, and B. Budowle. 1992. Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. Annals of Human Genetics **56:**45–57.

Compton, B. W., K. McGarigal, S. A. Cushman, and L. R. Gamble. 2007. A resistant-kernel model of connectivity for amphibians that breed in vernal pools. Conservation Biology **21:**788–799.

Corander, J., P. Waldmann, and M. J. Sillanpää. 2003. Bayesian analysis of genetic differentiation between populations. Genetics **163:**367–374.

Corander, J., P. Waldmann, P. Marttinen, and M. J. Sillanpää. 2004. BAPS 2: enhanced possibilities for the analysis of genetic population structure. Bioinformatics **20:**2363–2369.

Croshaw, D. A., N. A. Schable, M. B. Peters, and T. C. Glenn. 2005. Isolation and characterization of microsatellite DNA loci from *Ambystoma* salamanders. Conservation Genetics **6:**473–479.

Cushman, S. A. 2006. Effects of habitat loss and fragmentation on amphibians: a review and prospectus. Biological Conservation **128:**231–240.

Epps, C. W., J. D. Wehausen, V. C. Bleich, S. G. Torres, and J. S. Brashares. 2007. Optimizing dispersal and corridor models using landscape genetics. Journal of Applied Ecology **44**:714–724.

Gamble, L. R., K. McGarigal, and B. W. Compton. 2007. Fidelity and dispersal in the pond-breeding amphibian, *Ambystoma opacum*: implications for spatio-temporal population dynamics and conservation. Biological Conservation **139:**247–257.

Gibbs, J. P. 1998. Genetic structure of redback salamander *Plethodon cinereus* populations in continuous and fragmented forests. Biological Conservation **86:**77–81.

Goudet, J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. Journal of Heredity **86:**485–486.

Gustafson, E. J., N. L. Murphy, and T. R. Crow. 2001. Using a GIS model to assess terrestrial salamander response to alternative forest management plans. Journal of Environmental Management **63:**281–292.

Halverson, M. A., D. K. Skelly, and A. Caccone. 2006. Inbreeding linked to amphibian survival in the wild but not in the laboratory. Journal of Heredity **97:**499–507.

Hanski, I., and M. E. Gilpin. 1997. Metapopulation biology: ecology, genetics and evolution. Academic Press, San Diego, California.

Houlahan, J. E., C. S. Findlay, B. R. Schmidt, A. H. Meyer, and S. L. Kuzmin. 2000. Quantitative evidence for global amphibian population declines. Nature **404:**752–755.

Hovel, K. A., and R. N. Lipcius. 2002. Effects of seagrass habitat fragmentation on juvenile blue crab survival and abundance. Journal of Experimental Marine Biology and Ecology **271:**75–98.

Kaiser, H. F. 1960. The application of electronic computers to factor analysis. Educational and Psychological Measurement **20:**141–151.

Lesbarrères, D., A. Pagano, and T. Lodé. 2003. Inbreeding and road effect zone in a Ranidae: the case of Agile frog, *Rana dalmatina* Bonaparte, 1840. Comptes Rendus Biologies **326:**68–72.

Lloyd, P., T. E. Martin, R. L. Redmond, U. Langner, and M. M. Hart. 2005. Linking demographic effects of habitat fragmentation across landscapes to continental source-sink dynamics. Ecological Applications **15:**1504–1514.

Marsh, D. M., and P. C. Trenham. 2001. Metapopulation dynamics and amphibian conservation. Conservation Biology **15:**40–49.

Noel, S., M. Ouellet, P. Galois, and F. J. Lapointe. 2007. Impact of urban fragmentation on the genetic structure of the eastern red-backed salamander. Conservation Genetics **8:**599–606.

Piry, S., A. Alapetite, J. M. Cornuet, D. Paetkau, L. Baudouin, and A. Estoup. 2004. GENECLASS2: a software for genetic assignment and first-generation migrant detection. Journal of Heredity **95:**536–539.

Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Journal of Heredity **86:**248–249.

Rothermel, B. B., and R. D. Semlitsch. 2002. An experimental investigation of landscape resistance of forest versus old-field habitats to emigrating juvenile amphibians. Conservation Biology **16:**1324–1332.

Rowe, G., and T. J. C. Beebee. 2002. Population on the verge of a mutational meltdown? Fitness costs of genetic load for an amphibian in the wild. Evolution **57:**177–181.

Scribner, K. T., J. W. Arntzen, N. Cruddace, R. S. Oldham, and T. Burke. 2001. Environmental correlates of toad abundance and population genetic diversity. Biological Conservation **98:**201–210.

Semlitsch, R. D., and J. R. Bodie. 2003. Biological criteria for buffer zones around wetlands and riparian habitats for amphibians and reptiles. Conservation Biology **17:**1219–1228.

Smith, M. A., and D. M. Green. 2005. Dispersal and the metapopulation paradigm in amphibian ecology and conservation: are all amphibian populations metapopulations? Ecography **28:**110–128.

Spear, S. F., C. R. Peterson, M. D. Matocq, and A. Storfer. 2005. Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). Molecular Ecology **14:**2553–2564.

Storfer, A. 1999. Gene flow and population subdivision in the streamside salamander, *Ambystoma barbouri*. Copeia **1:**174–181.

Stuart, S. N., J. S. Chanson, N. A. Cox, B. E. Young, A. S. L. Rodrigues, D. L. Fischman, and R. W. Waller. 2004. Status and trends of amphibian declines and extinctions worldwide. Science **306:**1783–1786.

Storfer, A., M. A. Murphy, J. S. Evans, C. S. Goldberg, S. Robinson, S. F. Spear, R. Dezzani, E. Delmelle, L. Vierling, and L. P. Waits. 2007. Putting the 'landscape' in landscape genetics. Heredity **2006:**1–15.

Trenham, P. C., W. D. Koenig, and H. B. Shaffer. 2001. Spatially autocorrelated demography and interpond dispersal in the salamander *Ambystoma californiense*. Ecology **82:**3519–3530.

Van Oosterhout, C., W. F. Hutchinson, P. M. Willis, and P. Shipley. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes **4:**535–538.

Wauters, L. A., Y. Hutchinson, D. T. Parkin, and A. A. Dhondt. 1994. The effects of habitat fragmentation on demography and on the loss of genetic variation in the red squirrel. Proceedings of the Royal Society of London, Series B **255:**107–111.

Weir, B. S., and C. C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. Evolution **38:**1358–1370.

Zamudio, K. R., and A. M. Wieczorek. 2007. Fine-scale spatial genetic structure and dispersal among spotted salamander (*Ambystoma maculatum*) breeding populations. Molecular Ecology **16:**257–274.