

# Ecological correlates of population genetic structure: a comparative approach using a vertebrate metacommunity

Mollie K. Manier\* and Stevan J. Arnold

Department of Zoology, 3029 Cordley Hall, Oregon State University, Corvallis, OR 97331-2914, USA

Identifying ecological factors associated with population genetic differentiation is important for understanding microevolutionary processes and guiding the management of threatened populations. We identified ecological correlates of several population genetic parameters for three interacting species (two garter snakes and an anuran) that occupy a common landscape. Using multiple regression analysis, we found that species interactions were more important in explaining variation in population genetic parameters than habitat and nearest-neighbour characteristics. Effective population size was best explained by census size, while migration was associated with differences in species abundance. In contrast, genetic distance was poorly explained by the ecological correlates that we tested, but geographical distance was prominent in models for all species. We found substantially different population dynamics for the prey species relative to the two predators, characterized by larger effective sizes, lower gene flow and a state of migration-drift equilibrium. We also identified an escarpment formed by a series of block faults that serves as a barrier to dispersal for the predators. Our results suggest that successful landscape-level management should incorporate genetic and ecological data for all relevant species, because even closely associated species can exhibit very different population genetic dynamics on the same landscape.

**Keywords:** barriers to dispersal; *Bufo boreas*; landscape genetics; microsatellites; *Thamnophis elegans*; *Thamnophis sirtalis*

## 1. INTRODUCTION

Identification of ecological and evolutionary factors that control local differentiation of populations is crucial for understanding fine-scale microevolutionary processes. This knowledge also allows more accurate prediction of the impact of conservation policy on genetic structure and evolutionary potential. Molecular genetic data typically average population structure over long time periods, while ecological studies usually sample populations over relatively shorter time-intervals. A combination of these two approaches can illuminate the processes driving genetic differentiation. Such a combined framework can then be used as a basis for sound management on landscape and ecosystem scales.

Recent analyses of population structure have incorporated landscape features in a variety of ways (e.g. Jacquemyn *et al.* 2004; Whiteley *et al.* 2004; Namroud *et al.* 2005; Spear *et al.* 2005). However, despite the growing prominence of landscape features as causal factors, most studies focus on a single species or on a group of closely related species (Hirau & Kudo 2004; Petren *et al.* 2005). In this study, we analyse the response of three species to a common landscape. Because these three species interact as competitors

and predator-and-prey, we can explore the role of various ecological factors, including species interactions, as determinants of population genetic structure. Our results highlight the role of habitat characteristics and biological interactions, as well as landscape features, in the determination of population genetic structure.

In this study system (in Lassen County, CA, USA), the western toad (*Bufo boreas*), the terrestrial garter snake (*Thamnophis elegans*) and the common garter snake (*Thamnophis sirtalis*) occupy ponds, lakes and flooded meadows. *Bufo boreas* breeds in shallow water, and its tadpoles and metamorphs are eaten by both snake species (Kephart 1982; Kephart & Arnold 1982). The two snake species exhibit source-sink dynamics in this study system and show correlated genetic structure among the sites where they coexist (Manier & Arnold 2005). If the shared landscape played a major role in the observed pattern, one would expect an unrelated but ecologically relevant species that coexists on the landscape (e.g. *B. boreas*) to exhibit a similar pattern of genetic differentiation and isolation by distance.

We looked for barriers to dispersal using a clustering analysis that detected genetically similar groups of geographically proximate populations. We also used stepwise regression to identify ecological factors that influence genetic differentiation as expressed by effective population size ( $N_e$ ), migration rate ( $m$ ) and genetic distance for each species. Explanatory variables used in the analyses included habitat features, species abundance and nearest-neighbour characteristics.

\* Author and address for correspondence: Division of Biological Sciences, Stanford University, Hopkins Marine Station, Oceanview Boulevard, Pacific Grove, CA 93950, USA (manier@stanford.edu).

The electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2006.3678> or via <http://www.journals.royalsoc.uk>.

Table 1. Names, abbreviations and characteristics of study sites. (Type refers to relative depth: M, meadow; L, lake and LS, Eagle Lake shoreline. Sampling effort is the number of collector trips over 5 years.  $N_c$  is the number of individuals caught over 5 years. d and n indicate sampling during the day for snakes and at night for toads.  $N$  is the sample size for the microsatellite analysis. Asterisk denotes informal name, not official geographical place name.)

site	code	elevation (m)	perimeter (km)	type	sampling effort	<i>B. boreas</i>		<i>T. elegans</i>		<i>T. sirtalis</i>	
						$N_c$	$N$	$N_c$	$N$	$N_c$	$N$
Antelope Mountain Pond*	AMP	1920	0.52	M	15	0	—	32	30	20	19
Ashurst Lake	ASH	1950	3.5	M	9d, 4n	7	—	30	30	39	38
Blue Water	BLW	1770	0.425	L	11d, 10n	18	18	25	27	2	—
Bullard Lake	BUL	1860	3.3	M	13	0	—	67	67	35	35
Cleghorn Reservoir	CLG	1880	3.6	M	15	1	—	34	44	4	—
Colman Lake	COL	1965	2.2	L	28	0	—	49	24	35	27
Deans Meadow	DNS	1980	0.95	L	18	1	—	32	26	29	27
Feather Lake	FEA	1740	5.3	L	15	0	—	3	—	25	24
Gallatin Shoreline*	GAL	1575	2.3	LS	27	0	—	113	56	0	—
Gordon Lake	GOR	1850	3.1	M	4	0	—	2	—	43	42
Jacks Lake	JCK	1690	2	L	9	0	—	28	27	1	—
Little Cleghorn Reservoir	LTC	1880	3	M	15	0	—	40	18	8	—
Mahogany Lake	MAH	2065	1.7	L	39	0	—	167	91	94	26
McCoy Flat Reservoir	MCY	1700	18.9	M	25	0	—	20	16	3	—
Mosquito Flat	MOS	1890	1.6	M	2d, 2n	24	24	9	—	1	—
Nameless Meadow*	NML	1915	1.5	M	33	0	—	95	29	30	29
Papoose Meadows	PAP	1645	5	M	45	7	—	306	140	16	—
Pikes Point	PIK	1555	3.2	LS	53d, 4n	32	32	108	48	0	—
Pine Valley Meadow*	PVM	1730	9.1	M	9	0	—	72	70	83	83
Rocky Point	RKY	1580	4.1	LS	21	0	—	13	19	1	—
Roney Corral	RON	1825	1.2	L	28d, 8n	28	28	26	24	25	29
Camp Stanford	STF	1870	0.458	L	8d, 8n	15	15	47	45	31	32
Summit Lake	SUM	1890	2.5	M	13	0	—	27	27	27	22
Upper Gooch Valley Pond*	UGV	1920	0.2	M	7d, 8n	33	33	0	—	0	—

## 2. MATERIAL AND METHODS

### (a) Microsatellite analysis

Genetic methods for *T. elegans* and *T. sirtalis* were reported by Manier & Arnold (2005). We evaluated variation at nine microsatellite loci for *T. elegans* and at seven for *T. sirtalis*. Analysis of 16 *B. boreas* microsatellite loci followed the same procedures, which are briefly outlined below. We collected a hind toe tip from 150 adult *B. boreas* at five sites (table 1) within an area of 1000 km<sup>2</sup> (figure 1), and stored them in Drierite, an anhydrous calcium sulphate desiccant. Simandle (2005) provided genotypic data for the Pikes Point locality. Sixteen microsatellite primer sets were obtained from Simandle *et al.* (2006). PCR profiles consisted of 94°C for 2 min followed by 36 cycles of 94, 55 and 72°C for 30 s each, ending with 72°C for 2 min.

All statistical analyses for *B. boreas* genetic data were as described by Manier & Arnold (2005). These analyses included exact tests for departure from Hardy–Weinberg equilibrium (Guo & Thompson 1992; Markov chain parameters: 5000 dememorizations, 500 000 steps per chain) calculated in ARLEQUIN v. 2.000 (Schneider *et al.* 2000) and tests for linkage disequilibrium (Slatkin & Excoffier 1996; Markov chain parameters: 5000 dememorizations, 1000 batches, 5000 iterations per batch) performed in GENEPOP (Raymond & Rousset 1995). Significance levels were adjusted for multiple comparisons (Rice 1989). Number of alleles and observed and expected heterozygosities in each population and in all populations were calculated in GENEPOP (Raymond & Rousset 1995).

Overall and population pairwise estimates of  $F_{ST}$  were calculated using analysis of molecular variance (AMOVA;

Excoffier *et al.* 1992) in ARLEQUIN v. 2.000. Significance was assessed after 16 000 permutations for global estimates and 3000 permutations for pairwise estimates, with  $p$ -values adjusted with the sequential Bonferroni correction. We regressed genetic distance ( $F_{ST}/(1-F_{ST})$ ; Rousset 1997) on log geographical distance to look for patterns of migration-drift equilibrium (Hutchison & Templeton 1999). A Mantel test (Mantel 1967; Mantel & Valand 1970; Manly 1997) implemented in ARLEQUIN v. 2.000 was used to estimate the correlation between genetic distance and log geographical distance (significance over 10 000 permutations).

Estimates of effective population size and migration rate were obtained with MIGRATE v. 1.7.6.1 (Beerli & Felsenstein 2001). Assumptions, protocols and parameters used for the analysis are as outlined by Manier & Arnold (2005).  $F_{ST}$  estimates were used as starting values for the initial analysis. For all other analyses, ending parameters of the previous run were used as starting values for the next run until the results equilibrated at approximately the same values. Ten short chains with 10 000 sampled genealogies each and two long chains with 100 000 sampled genealogies each were run for each analysis. One of every 20 constructed genealogies was sampled and multiple long chains were combined for estimates. We used adaptive heating with temperature specifications of 1.0, 1.2, 1.5 and 3.0.

Effective population size was expressed as  $\Theta = 4N_e\mu$ , where  $N_e$  is effective population size and  $\mu$  is mutation rate. Assuming a mutation rate typical for vertebrates of  $10^{-4}$  per locus per generation (Dallas 1992; Edwards *et al.* 1992; Weber & Wong 1993; Banchs *et al.* 1994; Ellegren 1995),  $N_e$  was obtained from  $\Theta$ . Estimates of migration were expressed as  $4N_em$ , where

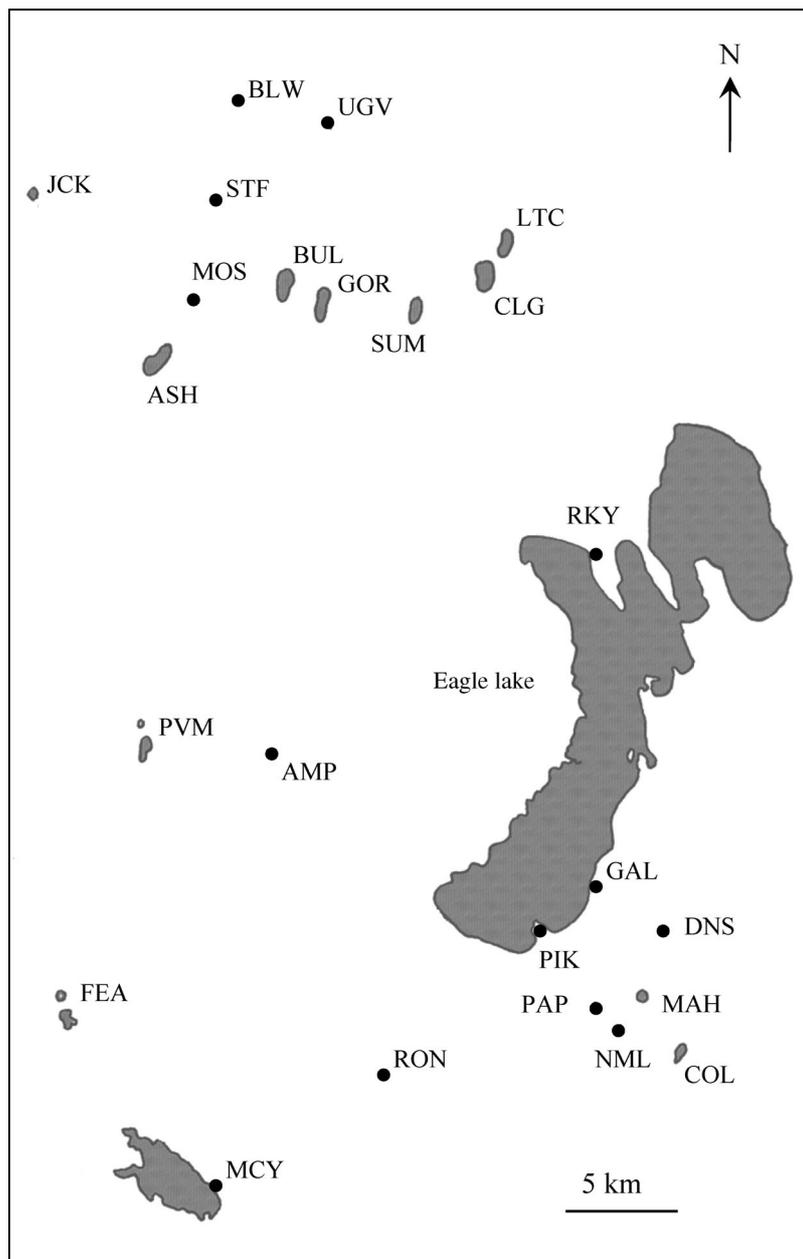


Figure 1. Map of the study area showing sampled sites. (For a key to the site codes, see table 1.)

$m$  is migration rate and  $N_e$  refers to the receiving population. Using  $N_e$  estimates from the previous step,  $m$  from each population into all other populations was calculated.

#### (b) Barriers to gene flow

For each species, spatial clustering of variation at micro-satellite loci and, hence, potential barriers to dispersal were identified by partitioning genetic variation within and among populations and regions using AMOVA. The program SAMOVA v. 1.0 (spatial analysis of molecular variance; Dupanloup *et al.* 2002) was used to identify clusters of populations. For each species, SAMOVA used decimal degree latitude and longitude for each population (obtained using MAPTECH TERRAIN NAVIGATOR v. 3.02) to create a user-defined number of maximally differentiated groups of geographically proximate populations. For example, if three groups were defined, SAMOVA chose three spatially clustered groups of populations with the highest within-group relatedness and among-group divergence. Since SAMOVA sets the missing value threshold for rejecting a locus relatively low (5%), we

tested the groups again using ARLEQUIN v. 2.000 (rejection threshold of 50%). Thus, we used SAMOVA to generate hypotheses about population clusters and tested these hypotheses in ARLEQUIN. For each hypothesis, we confirmed that the proposed groupings were spatially realistic and, if necessary, regrouped the sites to reflect reality. The  $F$ -statistic analogues,  $\Phi_{ST}$ ,  $\Phi_{SC}$  and  $\Phi_{CT}$  were estimated for each hypothesis and their significance levels were determined with 16 000 permutations (Excoffier *et al.* 1992).  $\Phi_{ST}$  represents the correlation between randomly drawn pairs of alleles within populations relative to the whole array of populations,  $\Phi_{CT}$  is the correlation between randomly drawn pairs of alleles within a group of populations relative to the whole array of populations and  $\Phi_{SC}$  is the correlation between randomly drawn pairs of alleles within populations relative to a group of populations (Excoffier *et al.* 1992).

#### (c) Ecological analysis

Multiple linear regression was used to identify ecological variables that best predicted population differentiation

Table 2. Abbreviations and descriptions of variables used in stepwise regression analysis.

abbreviation	variable description
PERIM	site perimeter (km)
TYPE	site type (1, meadow; 2, lake; 3, lakeshore)
ELEV	elevation (m)
BUFO	census size of <i>B. boreas</i>
ELEG	census size of <i>T. elegans</i>
SIRT	census size of <i>T. sirtalis</i>
SNAKES	summed census size of <i>T. elegans</i> and <i>T. sirtalis</i>
NN-DIST	distance to nearest neighbour (km)
NN-PERIM	nearest-neighbour site perimeter (km)
NN-TYPE	nearest-neighbour site type (1, meadow; 2, lake; 3, lakeshore)
NN-ELEV	nearest-neighbour elevation (m)
NN-BUFO	nearest-neighbour <i>B. boreas</i> census size
NN-ELEG	nearest-neighbour <i>T. elegans</i> census size
NN-SIRT	nearest-neighbour <i>T. sirtalis</i> census size
NN-SNAKES	nearest-neighbour summed census size of <i>T. elegans</i> and <i>T. sirtalis</i>
PERIMDIFF	difference in perimeter between two sites (km)
TYPEDIFF	difference in type between two sites
ELEVDIFF	difference in elevation between two sites (m)
BUFODIFF	difference in <i>B. boreas</i> census sizes between two sites
ELEGDIFF	difference in <i>T. elegans</i> census sizes between two sites
SIRTDIFF	difference in <i>T. sirtalis</i> census sizes between two sites
SNAKEDIFF	difference in summed <i>T. elegans</i> and <i>T. sirtalis</i> census sizes between two sites
DIST	distance between sites (km)

( $F_{ST}$ ), effective population size ( $N_e$ ) and one-way migration rate ( $m$ ). Independent variables (table 2) described habitat characteristics (site perimeter, site type and elevation) and census size of the focal and/or interacting species for each of 24 sites and its nearest neighbour (table 1). Log pairwise distance between sites or nearest-neighbour distance was also included to describe spatial structure. Site perimeter (km), elevation (m) and pairwise distances between sites (km) were estimated using MAPTECH TERRAIN NAVIGATOR v. 3.02. *Site perimeter* represented the maximum circumference of a pond or meadow or the length of Eagle Lake shoreline. This variable approximated a biologically relevant measure of habitat size, because shallow water is important to all the three species for breeding, foraging and/or thermoregulation. Snakes in this system are most often found around pond or meadow edges, and toad egg masses and larvae are likewise usually situated in shallow water. Meadow perimeter was defined as the maximum flooded waterline. *Site type* was characterized as meadow, lake or Eagle Lake shoreline (lakeshore) based on average depth and resulting degree of annual change in water level. *Abundance* represented all individuals caught in the months of June and July from 1999 to 2004, adjusted by sampling effort (number of collectors over all sampling trips). Garter snakes were caught by hand while searching at water edges and under cover objects, typically between midmorning and late afternoon. Ponds harbouring breeding toad

populations were identified by the presence of tadpoles and metamorphs, and adult toads were captured by hand or net at night. Since toads were not sampled during the breeding season, these estimates probably underestimate true census sizes. However, since all populations were sampled within the same time period every year, these numbers probably accurately reflect the sizes of populations relative to one another.

Regression models included abundance of the focal and interacting species. Since the two garter snake species eat the same prey (Kephart 1982; Kephart & Arnold 1982), they were viewed as interacting species. *Bufo boreas* was not included in the snake models, because the diets of both predators can include other more abundant prey species, decreasing the ecological importance of *B. boreas*. For *B. boreas*, the interacting species variable was total garter snake abundance. Study sites whose nearest neighbour was not sampled for genetic variation were excluded from the analysis (FEA, JCK and MCY). The stepwise regression procedure selected the final model.

Of the three population genetic parameters used as response variables,  $m$  and genetic distance were expressed as pairwise matrices, while  $N_e$  had a single value per population. Regression models describing the matrix response variables consisted of pairwise population differences in each of the independent variables described previously. For example, ELEVDIFF was a matrix describing pairwise differences in elevation between sites. Since migration measures similarity between two populations (e.g. the proportion of population A that originated from population B), it was converted to a measure of dissimilarity (the proportion of population A that did not originate from population B) using the formula  $\rho = 1 - m$  (Sokal & Rohlf 1995). Elements of the genetic distance matrix, representing the distance between two populations, were calculated as  $F_{ST}/(1 - F_{ST})$  (Rousset 1997). Explanatory matrix variables of the  $F_{ST}/(1 - F_{ST})$  response matrix represented absolute differences between population pairs, because  $F_{ST}/(1 - F_{ST})$  is a symmetrical matrix that characterizes the overall genetic distance between two populations. For example, ELEVDIFF, in this case, was a symmetrical matrix of absolute differences in elevation between sites. In contrast, explanatory matrix variables of the  $\rho$  response matrix contained information on the directionality of the difference, because the  $\rho$  matrix was an asymmetrical full matrix of bi-directional dissimilarities. Each element of the asymmetrical variable matrices represented the value of the receiving population minus the value of the source population. In this model, ELEVDIFF was an asymmetrical matrix of differences, with negative values indicating that the receiving site was lower in elevation than the source site.

Stepwise regression of  $N_e$  was performed in SAS v. 9.2 (SAS Institute 2002). Analyses of  $\rho$  or  $F_{ST}/(1 - F_{ST})$  matrices were implemented using forward and backward selections in PERMUTE! v. 3.4 alpha 9 (Legendre *et al.* 1994) with a  $p$  to enter or remove of 0.1. In all cases, forward selection and backward elimination chose the same variables. Significance of the coefficients and associated  $R^2$  values were determined using a permutation method for distance matrices (Legendre *et al.* 1994). PERMUTE! gives  $p$ -value and  $R^2$  for forward-selected variables and backward-eliminated models but does not provide coefficients and standard errors for selected variables. These statistics were obtained using SAS, which identified the same models as PERMUTE!

### 3. RESULTS

#### (a) Population genetics

One *B. boreas* locus deviated significantly from Hardy–Weinberg equilibrium in one population (MOS) after sequential Bonferroni correction (Rice 1989). All loci were in linkage equilibrium in all populations. Numbers of alleles ranged from 6 to 27 per locus with 4–16 per population per locus. Population averages for numbers of alleles over all loci ranged from 7 to 10 (see table 5 in the electronic supplementary material). Genetic differentiation of *B. boreas* populations was low but significant ( $F_{ST}=0.024$ ;  $p<0.0001$ ). Pairwise  $F_{ST}$  values are given in table 6 in the electronic supplementary material. We found significant isolation by distance (Mantel  $r=0.678$ ;  $p=0.016$ ), with a regression plot suggesting a migration–drift equilibrium (Hutchison & Templeton 1999; see figure 2 in the electronic supplementary material).

*Bufo boreas*  $N_e$ 's were high relative to the garter snakes (Manier & Arnold 2005) with an average  $\Theta=4N_e\mu$  of 0.497 (see table 7 in the electronic supplementary material). This  $\Theta$  yields an  $N_e$  of approximately 1240. UGV had the largest  $N_e$  (1765), followed by PIK (1560), while STF had the smallest  $N_e$  (765). Despite high  $N_e$ ,  $m$  was very low. The average  $4N_e m$  between populations was 3.12 or 0.0007 migrants per generation. UGV was the largest source of migrants (0.005), implying that 0.5% of individuals in other populations originated from UGV. We found no evidence of a large source population driving migration in *B. boreas* as observed in both garter snakes (Manier & Arnold 2005).

#### (b) Barriers to dispersal

*Thamnophis elegans* populations clustered into at most three groups by separation of COL and DNS ( $p=0.010$ ), then BLW ( $p=0.019$ ) from the other populations. *Thamnophis sirtalis* populations formed a maximum of six groups with COL, DNS and MAH splitting off first ( $p=0.006$ ), followed by COL splitting from this smaller group ( $p=0.019$ ). Next ASH ( $p=0.007$ ), SUM ( $p=0.018$ ), and BUL and GOR ( $p=0.013$ ) broke from the main cluster. Thus, population structure of both snakes was characterized by strong differentiation of a group that included COL and DNS. These two populations are separated from the others by a 300 m escarpment formed by a series of block faults. Structuring of *B. boreas* populations was not statistically significant but suggested a weak separation of northern (BLW, UGV, MOS and STF) and southern (PIK and RON) populations ( $p=0.066$ ). The two clusters of points at smaller and larger geographical distances in figure 2 in the electronic supplementary material correspond with distances within and between the groups of populations identified by SAMOVA, respectively (table 3).

#### (c) Ecological correlates

Stepwise regression results are shown in table 4. For *T. elegans*, variation in  $N_e$  was explained only by abundance ( $F_{1,20}=24.72$ ,  $p<0.0001$ ,  $R^2=0.57$ ). For *T. sirtalis*,  $N_e$  was explained first by abundance, then by a positive correlation with elevation and a negative correlation with nearest-neighbour elevation ( $F_{3,20}=24.78$ ,  $p<0.0001$ ,  $R^2=0.81$ ). In other words, *T. sirtalis* populations with higher  $N_e$  were characterized by higher census sizes, higher elevations and lower elevation neighbours.

For *B. boreas*,  $N_e$  was larger when the nearest neighbour was deeper ( $F_{2,20}=31.92$ ,  $p<0.0001$ ,  $R^2=0.78$ ).

Interpretation of migration associations is facilitated by the fact that variable estimates are identical for  $\rho$  and  $m$ , but with signs reversed. For example, in *T. elegans*, the parameter estimate for ELEGDIFF is  $-2.16\times 10^{-4}$  when  $\rho$  is the dependent variable (table 4) and  $2.16\times 10^{-4}$  when  $m$  is the dependent variable (data not shown).  $m$  for *T. elegans* was most associated with differences in abundance of *T. elegans*, then abundance of *T. sirtalis*, geographical distance, elevation and habitat type ( $p=0.001$ ,  $R^2=0.34$ ). In short, *T. elegans* migrates to sites that primarily have more conspecifics, but also have fewer *T. sirtalis*, are nearby, lower in elevation and shallower. *Thamnophis sirtalis* migration occurs more to populations with higher *T. sirtalis* abundance ( $p=0.001$ ,  $R^2=0.46$ ). *Bufo boreas* migrated towards sites that were deeper, had fewer snakes, were at higher elevation and were larger and nearby ( $p=0.001$ ,  $R^2=0.66$ ).

Variation in genetic distance was explained primarily by geographical distance. After accounting for this variable, genetic distance between *T. elegans* populations increased with larger differences in elevation ( $p=0.002$ ,  $R^2=0.14$ ). Genetic distance of *T. sirtalis* populations was greater with smaller differences in *T. elegans* abundance and more similar site perimeter sizes ( $p=0.001$ ,  $R^2=0.35$ ), and *B. boreas* genetic distance was related only to geographical distance ( $p=0.007$ ,  $R^2=0.46$ ).

Relative to their representation, habitat descriptors were more important than species abundances and nearest-neighbour characteristics. Regression models of  $N_e$  always included abundance and were explained by 1–3 variables that explained 57–81% of the variation. Interspecific interactions did not influence  $N_e$  for any species. Overall,  $m$  was determined by 1–5 variables of habitat features, abundance and interspecific interactions. In particular, migration occurred from sites with higher densities to sites with lower densities of both the focal and interacting species, and the selected models explained 34–66% of the variation. Genetic distance was explained by 1–3 variables, which accounted for 14–46% of the variation.

### 4. DISCUSSION

#### (a) Ecological genetics

Multiple regression analysis identified habitat variables, biotic features and nearest-neighbour distances as elements that were statistically important in explaining population genetic parameters for the three study species. Models for *T. elegans* explained much less of the variation in population genetic parameters than for the other species (table 4), suggesting that factors other than those examined here are more important for genetic differentiation in that species. One possible explanation is that selection plays a major role in driving divergence of *T. elegans* populations. In separate work, we show that phenotypic differentiation between lakeshore and meadow populations is a key aspect of adaptation in this study system but is not reflected in microsatellite loci (M. K. Manier, C. M. Seyler and S. J. Arnold, unpublished data; see also Bronikowski & Arnold 1999).

Physical barriers to dispersal play an important role in this study system. Spatial clustering of both snake species identified a genetically distinct group of populations that

Table 3. Results of genetic clustering analysis. (Best results for each pre-defined number of groups are shown. Only the number of groups which gave significant or almost significant  $\Phi_{CT}$  is shown.)

number of groups	structure tested	total variance	$\Phi$ -statistic	s.e.	$p$ -value
<i>B. boreas</i>					
1	(BLW, MAR, MOS, PIK, RON, STF)	58.43	$\Phi_{ST}=0.024$	<0.0001	<0.0001
2	(BLW, MAR, MOS, STF) (PIK, RON)	58.96	$\Phi_{ST}=0.035$ $\Phi_{SC}=0.007$ $\Phi_{CT}=0.028$	<0.0001 <0.0001 0.0019	<0.0001 <0.0001 0.066
<i>T. elegans</i>					
1	(ANT, ANS, ASH, BLW, BUL, CLG, COL, DNS, GAL, JCK, LTC, MAH, MCY, NML, PAP, PIK, RKY, RON, STF, SUM)	24.49	$\Phi_{ST}=0.024$	<0.0001	<0.0001
2	(ANT, ANS, ASH, BLW, BUL, CLG, GAL, JCK, LTC, MAH, MCY, NML, PAP, PIK, RKY, RON, STF, SUM) (COL, DNS)	25.52	$\Phi_{ST}=0.039$ $\Phi_{SC}=0.022$ $\Phi_{CT}=0.018$	<0.0001 <0.0001 0.0008	<0.0001 <0.0001 0.010
3	(ANT, ANS, ASH, BLW, BUL, CLG, GAL, JCK, MAH, MCY, NML, PAP, PIK, RKY, RON, STF, SUM) (COL, DNS) (BLW)	25.35	$\Phi_{ST}=0.033$ $\Phi_{SC}=0.022$ $\Phi_{CT}=0.011$	<0.0001 <0.0001 0.0011	<0.0001 <0.0001 0.019
<i>T. sirtalis</i>					
1	(ANT, ANS, ASH, BUL, COL, DNS, FEA, GOR, MAH, NML, RON, STF, SUM)	21.12	$\Phi_{ST}=0.035$	<0.0001	<0.0001
2	(ANT, ANS, ASH, BUL, FEA, GOR, NML, RON, STF, SUM) (COL, DNS, MAH)	21.44	$\Phi_{ST}=0.054$ $\Phi_{SC}=0.025$ $\Phi_{CT}=0.029$	<0.0001 <0.0001 0.0007	<0.0001 <0.0001 0.006
3	(ANT, ANS, ASH, BUL, DNS, FEA, GOR, NML, RON, STF, SUM) (DNS, MAH) (COL)	21.44	$\Phi_{ST}=0.049$ $\Phi_{SC}=0.027$ $\Phi_{CT}=0.023$	<0.0001 <0.0001 0.0011	<0.0001 <0.0001 0.019
4	(ANT, ANS, ASH, BUL, DNS, FEA, GOR, NML, RON, STF, SUM) (DNS, MAH) (COL) (ASH)	21.40	$\Phi_{ST}=0.048$ $\Phi_{SC}=0.023$ $\Phi_{CT}=0.026$	<0.0001 <0.0001 0.0006	<0.0001 <0.0001 0.007
5	(ANT, ANS, ASH, BUL, DNS, FEA, GOR, NML, RON, STF, SUM) (DNS, MAH) (COL) (ASH) (SUM)	21.30	$\Phi_{ST}=0.044$ $\Phi_{SC}=0.024$ $\Phi_{CT}=0.020$	<0.0001 <0.0001 0.0011	<0.0001 <0.0001 0.018
6	(ANT, ANS, ASH, FEA, NML, RON, STF, SUM) (DNS, MAH) (COL) (ASH) (SUM) (BUL, GOR)	21.23	$\Phi_{ST}=0.041$ $\Phi_{SC}=0.020$ $\Phi_{CT}=0.022$	<0.0001 <0.0001 0.0009	<0.0001 <0.0001 0.013

included COL and DNS. A series of block faults have created a 300 m escarpment that separates these sites from NML and PAP to the west. This landscape feature appears to be a significant barrier to dispersal for both snake species, driving landscape-level patterns of migration-drift disequilibrium and similarity in genetic structure (Manier & Arnold 2005). Removing COL and DNS from the analysis restored migration-drift equilibrium (data not shown) and reduced the correlation in pairwise  $F_{ST}$  between the two species ( $p=0.06$ ). Clustering of *B. boreas* populations, though not statistically significant, differs from that of its predators in suggesting a north-south division.

Multiple regression analysis identified different ecological correlates, as well as some commonalities, for each species. Population census size best predicted  $N_e$  for all three species, suggesting an  $N_e-N_c$  association. Migration rates were generally explained by species abundance as well as by census size for the garter snakes and predator abundance for the toad. The trend for both snakes to migrate to sites with higher snake abundance may reflect convergence on a common resource. Nevertheless, this result is unexpected, given a historical signal of source-sink dynamics in both snake species (Manier & Arnold 2005). This discrepancy may reflect inaccuracies in the

abundance estimates or, most probably, it may simply result from trying to account for historical population divergence using contemporary measurements of ecological factors. Escape from predation may cause *B. boreas* to emigrate from sites with more snakes. Genetic distance for all the species was most explained by geographical distance, though  $R^2$  values were relatively low. These results reflect the growing realization that  $F_{ST}$  alone may lack important information that is present in independent estimates  $N_e$  and  $m$ . For example, biological associations (both intra- and interspecific) seem to play a larger role than environmental variables in explaining the variation in  $N_e$  and  $m$ . This pattern is not found in  $F_{ST}$ .

The amount of variation that can be explained in studies of ecological and evolutionary processes is an issue of current debate (Møller & Jennions 2002; Peek *et al.* 2003). The  $R^2$  values obtained in this study are consistent with the findings of Peek *et al.* (2003), who argue that the proportion of explained variance can be relatively high, up to 50%. The regression models from this study explained between 11 and 81% of the variance in population genetic parameters, with an average of 50%. Although higher  $R^2$  values can result from the inclusion of more explanatory variables, we observed no relationship between the

Table 4. Results of stepwise regression analysis. (Effective population size ( $N_e$ ), residency or  $1 - m$  ( $\rho$ ) and genetic distance ( $F_{ST}/1 - F_{ST}$ ) for each species, showing variables as  $X_1, X_2$ , etc. with parameter estimates (s.e.),  $R^2$  contribution and  $p$ -values for each explanatory variable.)

Y	species	X1	X2	X3	X4	X5	$R^2$	$p$ -value
$N_e$	<i>T. elegans</i>	ELEG 62.73 (12.62)					0.565	<0.0001
	<i>T. sirtalis</i>	SIRT 19.92 (2.43)	ELEV 0.62 (0.17)	NN-ELEV -0.30 (0.17)			0.814	<0.0001
$\rho$	<i>B. boreas</i>	BUFO 148.22 (24.10)	NN-TYPE 394.68 (89.53)				0.781	<0.0001
	<i>T. elegans</i>	ELEGDIFF -2.16 $\times 10^{-4}$	SIRTDIFF 2.04 $\times 10^{-5}$	DIST 2.12 $\times 10^{-4}$	ELEVDIFF 8.31 $\times 10^{-7}$	TYPEDIFF 2.01 $\times 10^{-4}$	0.342	0.001
$F_{ST}/(1 - F_{ST})$	<i>T. sirtalis</i>	SIRTDIFF 0.001; 0.252	(4.05 $\times 10^{-6}$ ) 0.001; 0.030	(5.54 $\times 10^{-5}$ ) 0.001; 0.026	(2.43 $\times 10^{-7}$ ) 0.001; 0.021	(4.93 $\times 10^{-5}$ ) 0.001; 0.013	0.449	0.001
	<i>B. boreas</i>	TYPEDIFF -3.42 $\times 10^{-4}$	SNAKEDIFF 2.46 $\times 10^{-5}$	ELEVDIFF -3.01 $\times 10^{-6}$	PERIMDIFF -2.70 $\times 10^{-4}$	DIST 1.22 $\times 10^{-4}$	0.659	0.001
$F_{ST}/(1 - F_{ST})$	<i>T. elegans</i>	DIST 6.23 $\times 10^{-3}$	ELEVDIFF 4.74 $\times 10^{-5}$	0.001; 0.207	0.001; 0.108	0.011; 0.097	0.114	0.002
	<i>T. sirtalis</i>	DIST 1.28 $\times 10^{-3}$	ELEGDIFF -4.01 $\times 10^{-3}$	PERIMDIFF -6.82 $\times 10^{-3}$			0.352	0.001
$F_{ST}/(1 - F_{ST})$	<i>B. boreas</i>	DIST 8.86 $\times 10^{-4}$	0.001; 0.158	0.001; 0.139	0.001; 0.055		0.464	0.007
			(2.39 $\times 10^{-4}$ ) 0.003; 0.464					

number of variables selected by the stepwise regression procedure and the total amount of variation explained ( $r=0.128$ ,  $p=0.743$ ).

### (b) *Bufo boreas* population genetics

We found low but significant genetic structuring of *B. boreas* populations, with an  $F_{ST}$  of 0.024, comparable to those previously reported for *T. elegans* (0.024) and *T. sirtalis* (0.035; Manier & Arnold 2005). However, populations of *B. boreas* had effective sizes that were substantially larger than those for the garter snakes (average  $N_e=1240$  for *B. boreas*, 375 for *T. elegans* and 325 for *T. sirtalis*), while bi-directional migration rates were much lower (average  $m=0.0007$  for *B. boreas*, 0.0012 for *T. elegans* and 0.0017 for *T. sirtalis*). Thus, comparable  $F_{ST}$  values among all species masked differences in  $N_e$  and  $m$ .

Genetic structuring observed for *B. boreas* is consistent with the common observation of philopatry in anurans (e.g. Berven & Grudzien 1990) and low juvenile dispersal, despite a high capacity for vagility in juveniles and adults (e.g. Goater *et al.* 1993; Sinsch 1997). These life-history observations, along with our results, support a general observation of population differentiation over short geographical distances in anurans (e.g. Rowe *et al.* 2000; Brede & Beebee 2004).

Several lines of evidence point to a more stable population structure for *B. boreas* than for the garter snakes. *Bufo boreas* had a higher expected heterozygosity (average  $H_e=0.77$ ) than *T. elegans* (0.54) and *T. sirtalis* (0.59), which could reflect larger stochastic fluctuations in garter snake population size or more extreme extinction–recolonization dynamics (Newman & Squire 2001). Higher *B. boreas* effective sizes and lower migration rates indicate that populations are relatively closed, with very little gene flow. Previous research revealed source–sink metapopulation dynamics in our garter snake populations (Manier & Arnold 2005). Other studies have found stable genetic structure in anurans (Hoffman *et al.* 2004; Rowe & Beebee 2004). Although population sizes can fluctuate widely (Berven & Grudzien 1990), they can be very large (Crawford 2003). These results are consistent with the general observation that source–sink dynamics in anurans are rare (Berven & Grudzien 1990).

This study represents the first comparative examination of landscape genetics in a metacommunity spanning two trophic levels. Our results contribute to an understanding of how ecological processes at smaller spatio-temporal scales shape population genetic and evolutionary trajectories at larger scales and across species. Conservation management of ecosystems will benefit from these types of studies through more accurate predictions of population structure and connectivity based on both community interactions and habitat features.

Many people assisted with fieldwork, especially J. Scarl, A. Bronikowski, D. Hanken, J. Feder, K. Manier and T. Knight. E. Simandle generously provided primer sequences. M. Blouin, M. Pfrender, E. Hoffman, C. Criscione, P. Beerli, P. Legendre and P. Casgrain gave statistical advice. M. Blouin, A. R. Kiester, P. McEvoy and several anonymous reviewers provided helpful comments on the manuscript. Funding was provided by an EPA STAR Fellowship (U-91552801-5), an NSF GK-12 Teaching Fellowship and an NSF DDIG (DEB-0309017) to

M.K.M., an NSF grant (DEB-9903934) to S.J.A. and M. Pfrender, and an NSF grant (DEB-0323379) to A. Bronikowski and S.J.A.

### REFERENCES

- Banchs, I., Bosh, A., Guimera, J., Lazaro, C., Puig, A. & Estivill, X. 1994 New alleles at microsatellite loci in CEPH families mainly arise from somatic mutations in the lymphoblastoid cell line. *Hum. Mutat.* **3**, 365–372. (doi:10.1002/humu.1380030407)
- Beerli, P. & Felsenstein, J. 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. (doi:10.1073/pnas.081068098)
- Berven, K. A. & Grudzien, T. A. 1990 Dispersal in the wood frog (*Rana sylvatica*): implications for genetic population structure. *Evolution* **44**, 2047–2056. (doi:10.2307/2409614)
- Brede, E. G. & Beebee, T. J. C. 2004 Contrasting population structures in two sympatric anurans: implications for species conservation. *Heredity* **92**, 110–117. (doi:10.1038/sj.hdy.6800391)
- Bronikowski, A. M. & Arnold, S. J. 1999 The evolutionary ecology of life history variation in the garter snake *Thamnophis elegans*. *Ecology* **80**, 2314–2325.
- Crawford, A. J. 2003 Huge populations and old species of Costa Rican and Panamanian dirt frogs inferred from mitochondrial and nuclear gene sequences. *Mol. Ecol.* **12**, 2525–2540. (doi:10.1046/j.1365-294X.2003.01910.x)
- Dallas, J. F. 1992 Estimation of microsatellite mutation rates in recombinant inbred strains of mouse. *Mamm. Genome* **5**, 32–38.
- Dupanloup, I., Schneider, S., Langaney, A. & Excoffier, L. 2002 A simulated annealing approach to define the genetic structure of populations. *Mol. Ecol.* **11**, 2571–2581. (doi:10.1046/j.1365-294X.2002.01650.x)
- Edwards, A., Hammond, H. A., Jin, L., Caskey, C. T. & Chakraborty, R. 1992 Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* **12**, 241–253. (doi:10.1016/0888-7543(92)90371-X)
- Ellegren, H. 1995 Mutation rates at porcine microsatellite loci. *Mamm. Genome* **6**, 376–377. (doi:10.1007/BF00364807)
- Excoffier, L., Smouse, P. E. & Quattro, J. M. 1992 Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**, 479–491.
- Goater, C. P., Semlitsch, R. D. & Bernasconi, M. V. 1993 Effects of body size and parasite infection on the locomotory performance of juvenile toads, *Bufo bufo*. *Oikos* **66**, 129–136.
- Guo, S. W. & Thompson, E. A. 1992 Performing the exact test for Hardy–Weinberg proportions for multiple alleles. *Biometrics* **48**, 361–372. (doi:10.2307/2532296)
- Hirau, A. S. & Kudo, G. 2004 Landscape genetics of alpine-snowbed plants: comparisons along geographic and snowmelt gradients. *Heredity* **93**, 290–298. (doi:10.1038/sj.hdy.6800503)
- Hoffman, E. A., Schueler, F. W. & Blouin, M. S. 2004 Effective population sizes and temporal stability of genetic structure in *Rana pipiens*, the northern leopard frog. *Evolution* **58**, 2536–2545. (doi:10.1554/04-444)
- Hutchison, D. W. & Templeton, A. R. 1999 Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* **53**, 1898–1914. (doi:10.2307/2640449)

- Jacquemyn, H., Honnay, O., Galbusera, P. & Roldán-Ruiz, I. 2004 Genetic structure of the forest herb *Primula elatior* in a changing landscape. *Mol. Ecol.* **13**, 211–219. (doi:10.1046/j.1365-294X.2003.02033.x)
- Kephart, D. G. 1982 Microgeographic variation in the diets of garter snakes. *Oecologia* **52**, 287–291. (doi:10.1007/BF00363852)
- Kephart, D. G. & Arnold, S. J. 1982 Garter snake diets in a fluctuating environment: a seven-year study. *Ecology* **63**, 1232–1236. (doi:10.2307/1938848)
- Legendre, P., Lapointe, F.-J. & Casgrain, P. 1994 Modeling brain evolution from behavior: a permutational regression approach. *Evolution* **48**, 1487–1499. (doi:10.2307/2410243)
- Manier, M. K. & Arnold, S. J. 2005 Population genetic analysis identifies source–sink dynamics for two sympatric garter snake species (*Thamnophis elegans* and *T. sirtalis*). *Mol. Ecol.* **14**, 3965–3976. (doi:10.1111/j.1365-294X.2005.02734.x)
- Manly, B. F. 1997 *Randomization and Monte Carlo methods in biology*, 2nd edn. New York, NY: Chapman & Hall.
- Mantel, N. 1967 The detection of disease clustering and generalized regression approach. *Cancer Res.* **27**, 209–220.
- Mantel, N. & Valand, R. S. 1970 A technique of non-parametric multivariate analysis. *Biometrics* **26**, 547–558. (doi:10.2307/2529108)
- Møller, A. P. & Jennions, M. D. 2002 How much variance can be explained by ecologists and evolutionary biologists? *Oecologia* **132**, 492–500. (doi:10.1007/s00442-002-0952-2)
- Namroud, M.-C., Park, A., Tremblay, F. & Bergeron, Y. 2005 Clonal and spatial genetic structures of aspen (*Populus tremuloides* Michx.). *Mol. Ecol.* **14**, 2969–2980. (doi:10.1111/j.1365-294X.2005.02653.x)
- Newman, R. A. & Squire, T. 2001 Microsatellite variation and fine-scale population structure in the wood frog (*Rana sylvatica*). *Mol. Ecol.* **10**, 1087–1100. (doi:10.1046/j.1365-294X.2001.01255.x)
- Peck, M. S., Leffler, A. J., Flint, S. D. & Ryel, R. J. 2003 How much variance is explained by ecologists? Additional perspectives. *Oecologia* **137**, 161–170. (doi:10.1007/s00442-003-1328-y)
- Petren, K., Grant, P. R., Grant, B. R. & Keller, L. F. 2005 Comparative landscape genetics and the adaptive radiation of Darwin's finches: the role of peripheral isolation. *Mol. Ecol.* **14**, 2943–2957. (doi:10.1111/j.1365-294X.2005.02632.x)
- Raymond, M. & Rousset, F. 1995 GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Hered.* **86**, 248–249.
- Rice, W. R. 1989 Analysing tables of statistical tests. *Evolution* **43**, 223–225. (doi:10.2307/2409177)
- Rousset, F. 1997 Genetic differentiation and estimation of gene flow from *F*-statistics under isolation by distance. *Genetics* **145**, 1219–1228.
- Rowe, G. & Beebee, T. J. C. 2004 Reconciling genetic and demographic estimators of effective population size in the anuran amphibian *Bufo calamita*. *Conserv. Genet.* **5**, 287–298. (doi:10.1023/B:COGE.0000031145.06421.d3)
- Rowe, G., Beebee, T. J. C. & Burke, T. 2000 A microsatellite analysis of natterjack toad, *Bufo calamita* metapopulations. *Oikos* **88**, 641–651. (doi:10.1034/j.1600-0706.2000.880321.x)
- SAS Institute 2002 *SAS/STAT software user's guide, Rel. 9.1*. Cary, NC: SAS Institute, Inc.
- Schneider, S., Roessli, D. & Excoffier, L. 2000 *Arlequin ver 2.000: a software for population genetics data analysis*. Geneva, Switzerland: Genetics and Biometry Laboratory, University of Geneva.
- Simandle, E. T. 2005 Phylogeography and metapopulation structure of the *Bufo boreas* group in the Great Basin. Ph.D. thesis, University of Nevada, Reno, USA.
- Simandle, E. T., Peacock, M. M., Zirelli, L. & Tracy, C. R. 2006 Sixteen microsatellite loci for the *Bufo boreas* group. *Mol. Ecol. Notes* **6**, 116–119. (doi:10.1111/j.1471-8286.2005.01159.x)
- Sinsch, U. 1997 Postmetamorphic dispersal and recruitment of first breeders in a *Bufo calamita* metapopulation. *Oecologia* **112**, 42–47. (doi:10.1007/s004420050281)
- Slatkin, M. & Excoffier, L. 1996 Testing for linkage disequilibrium in genotypic data using the EM algorithm. *Heredity* **76**, 377–383.
- Sokal, R. & Rohlf, F. J. 1995 *Biometry*, 3rd edn. New York, NY: W. H. Freeman.
- Spear, S. F., Peterson, C. R., Matocq, M. D. & Storfer, A. 2005 Landscape genetics of the blotched tiger salamander (*Ambystoma tigrinum melanostictum*). *Mol. Ecol.* **14**, 2553–2564. (doi:10.1111/j.1365-294X.2005.02573.x)
- Weber, J. L. & Wong, C. 1993 Mutation of human short tandem repeats. *Hum. Mol. Genet.* **2**, 1123–1128.
- Whiteley, A. R., Spruell, P. & Allendorf, F. W. 2004 Ecological and life history characteristics predict population genetic divergence of two salmonids in the same landscape. *Mol. Ecol.* **13**, 3675–3688. (doi:10.1111/j.1365-294X.2004.02365.x)