

An analysis of selection on a colour polymorphism in the northern leopard frog

ERIC A. HOFFMAN,* FREDERICK W. SCHUELER,† ADAM G. JONES‡ and MICHAEL S. BLOUIN§
*Department of Biology, University of Central Florida, Orlando, Florida 32816, USA, †Bishops Mills Natural History Centre, RR#2 Bishops Mills, Ontario, Canada K0G 1T0, ‡Department of Biology, Texas A & M University, College Station, Texas 77843, USA, §3029 Cordley Hall, Oregon State University, Department of Zoology, Corvallis, Oregon 97331, USA

Abstract

In this study, we investigated the role of selection in the maintenance of a dorsal colour polymorphism in natural populations of the northern leopard frog, *Rana pipiens*. We determined genetic structure both spatially and temporally from a suite of putatively neutral molecular markers and tested whether or not the colour locus exhibited patterns of genetic variation that differed from those of the neutral loci. Spatial genetic structure at the colour locus was indistinguishable from structure at neutral loci [95% confidence intervals of F_{ST} (neutral) = (0.07, 0.35), F_{ST} (colour locus) = 0.114]. In the temporal analysis, we found that the variance among populations in the change in allele frequency at the colour locus (equal to 0.004) lies within the 95% confidence intervals for the variance among populations in changes in allele frequencies at neutral loci. In light of our inability to show evidence for the selective maintenance of the colour polymorphism, we used computer simulations to infer the power of our spatial and temporal techniques to detect selection. The computer simulations showed that although the strength of selection (s) would need to be relatively strong to have been detected by the temporal approach ($s = 0.1$ – 0.4 , depending on the model tested), the spatial analysis would have detected all but weak selection ($s = 0.01$ – 0.04 , depending on the model tested). This study illustrates the importance of using a locus comparison approach to detect evidence for selective maintenance before conducting studies to measure the selective mechanisms maintaining a polymorphism.

Keywords: anuran, F_{DIST2} , genetic structure, ISSR, microsatellite, population genetics

Received 30 October 2005; revision accepted 06 February 2006

Introduction

A central goal of population genetics is to understand the origin and maintenance of genetic and phenotypic variation in natural populations. However, most studies investigating the maintenance of phenotypic polymorphisms posit a selective mechanism without investigating nonselective mechanisms for the persistence of the polymorphism. An alternate approach is to obtain evidence that selection has a role in maintaining a polymorphism before speculating about possible selective mechanisms or investing a great deal of effort in testing potential selective scenarios. One method to establish that a polymorphic trait is selectively maintained, as opposed to persisting as a neutral polymorphism governed by mutation and genetic drift, is to

study multiple polymorphisms simultaneously (Cavalli-Sforza 1966). The availability of molecular markers readily provides numerous polymorphic characters enabling a general approach (hereafter called the locus comparison approach) for testing whether the level of genetic variation among populations can be explained by drift alone or if selection must be invoked.

The locus comparison approach has primarily been applied in spatial analyses with suites of molecular markers to determine post hoc if any loci are putatively under selection. For example, studies have found differences in the patterns of genetic variation between different classes of molecular markers (e.g. Karl & Avise 1992) or even within a single type of molecular marker (e.g. Bowcock *et al.* 1991; Beaumont & Nichols 1996; Vitalis *et al.* 2001; Akey *et al.* 2002; Storz & Nachman 2003). Most of the studies employing the locus comparison approach to test for selection on phenotypic traits have examined quantitative characters under

Correspondence: Eric A. Hoffman, Fax: (407) 823 5769; E-mail: eahoffma@mail.ucf.edu

selection (Prout & Barker 1993; Spitze 1993; Long & Singh 1995; Bonnin *et al.* 1996; Yang *et al.* 1996; Lynch *et al.* 1999) or highly heritable traits thought to be under the control of few loci (King & Lawson 1995; Mithon *et al.* 1995). In general, these studies compared the difference between F_{ST} and Q_{ST} to infer the nature of selection responsible for quantitative variation. Few researchers have used the locus comparison approach to examine particular alleles at individual loci under selection, and these studies typically involved loci under intense artificial selection from agents such as insecticides (Chevillon *et al.* 1995; Mithon *et al.* 1995; Taylor *et al.* 1995) rather than sources of more natural selection, such as predation. Moreover, few studies have utilized both spatial and temporal sampling schemes to uncover the relative importance of microevolutionary forces (e.g. Charbonnel *et al.* 2002 and references therein). In the few studies that have been conducted, however, the combination of spatial and temporal analyses has been enlightening.

Colour or pattern polymorphisms provide excellent systems in which to evaluate whether a null hypothesis of genetic drift can be rejected in favour of an alternative hypothesis of selection to explain patterns of genetic variation in natural populations. Of the numerous studies attempting to measure selective mechanisms maintaining colour polymorphisms, only three recent studies have taken the locus comparison approach. Each of these studies (Gillespie & Oxford 1998; Andres *et al.* 2000, 2002) found that genetic variation among populations (F_{ST}) at the colour locus was significantly less than variation among populations at putatively neutral loci. Such a result suggests that some form of balancing selection is keeping allele frequencies similar among populations.

In the present study, we look for a genetic signature of selective maintenance of a dorsal colour polymorphism in populations of the northern leopard frog, *Rana pipiens*. This species hibernates in large bodies of water, breeds in ponds and swamps during the spring, and metamorphoses from early summer to early fall. By metamorphosis, tadpoles exhibit a conspicuous polymorphism in which individuals develop either green or brown permanent colouration on their dorsum. This striking colour dimorphism is controlled by a one-locus, two-allele system in which the green allele is dominant to the brown allele (Fogleman *et al.* 1980). Thus, the easily scored phenotypic frequencies can be used to obtain estimates of genotypic frequencies.

Past studies of mechanisms maintaining colour polymorphisms in anurans provide a plausible hypothesis as to how the polymorphism could be selectively maintained. Observations of snakes, herons, egrets, and birds of prey stalking adult and juvenile *R. pipiens* along with evidence of antipredator behaviour towards some of these predators (Heinen & Hammond 1997), led Hoffman & Blouin (2000) to conclude that direct selection on the colour polymorphism by visually

orientated predators is the most plausible mechanism maintaining this polymorphism. Unfortunately, *R. pipiens* individuals range widely and are viewed by predators against different backgrounds in each season, making direct measurement of selection problematic.

Previous studies have suggested that the green-brown colour polymorphism in *R. pipiens* is indeed selectively maintained. Schueler (1979) found correlations between colouration and habitat and concluded that green frogs are favoured in more forested habitats. Additionally, Corn (1981) found a correlation, from a single season, between colour and time of larval metamorphosis in Colorado, such that brown frogs emerged from ponds earlier, when less green vegetation is present, than did green frogs.

In this study, we aimed to determine if there was a genetic signature of selection on the colour locus in *R. pipiens*. Our analyses are based on spatial and temporal versions of the locus comparison approach, applied using data from putatively neutral loci (from molecular markers) compared to the colour locus. First, we used both a large and small-scale spatial approach to determine if the colour locus exhibited a pattern of genetic variation among populations that is different from that at neutral loci. At a large spatial scale, we might expect genetic variation among populations at the colour locus to be less than that observed at neutral loci (smaller F_{ST}) owing to stabilizing selection on the colour locus. Across a small spatial scale, we might expect either lower or higher F_{ST} depending on whether stabilizing selection is acting on the colour locus or if selection is acting to maintain different optimal allele frequencies in each population despite migration. Second, we used temporally spaced samples, obtained by resampling sites where one of us (FWS) had collected in the 1970s, to determine whether or not observed changes in allele frequencies through time at the colour locus are different from changes through time at simulated neutral bi-allelic loci. Finally, assuming balancing selection has maintained the polymorphism, we used computer simulations to determine the values of the selection coefficient (assuming frequency-dependent selection or overdominance) that would be detectable with our techniques. Our approach to investigating the selective maintenance of the colour polymorphism is novel because we use a combination of spatial and temporal analyses coupled with simulation-based analyses of power. Therefore, we can assess the relative strength of each analysis, under various analysis parameters and models of selection, and this study should serve as a model for future studies of selection in nature.

Materials and methods

Samples

Samples of *Rana pipiens* comprised two main classes: contemporary population collections and historical collections.

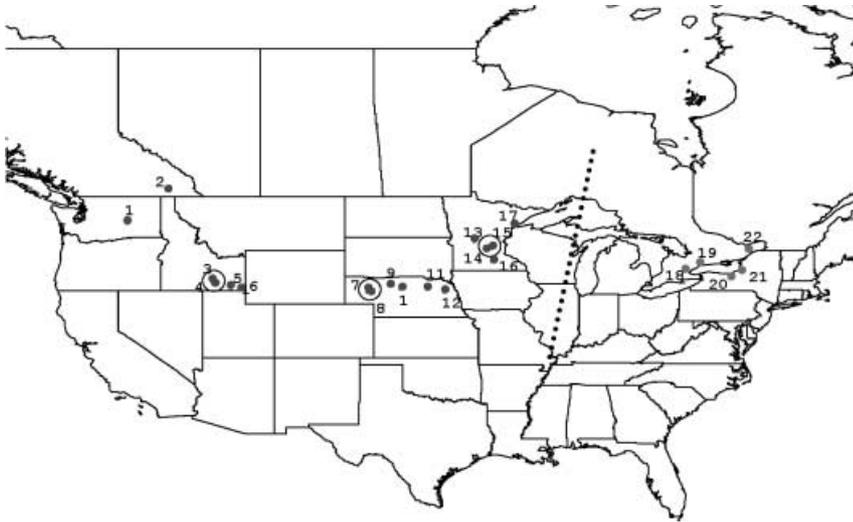


Fig. 1 Map of sampling localities. Dotted line indicates the east/west regional split. Circles around pairs of populations indicate populations used for small-scale geographical analyses. All eastern populations were used for temporal as well as spatial analysis. Exact population locations are listed in Table 1.

Table 1 Population identification (from Fig. 1), state or province of population origin, analysis for which each population was used (L, large spatial area; S, small spatial area; T, temporal variation), number of samples per population, number of green frogs (spatial analysis; see text), geographical coordinates of each population (map datum = WGS84), and museum catalogue number from the Canadian Museum of Nature for museum specimens

Population	Location	Analysis	<i>N</i>	<i>N</i> _{green}	Latitude	Longitude	Catalogue no.
1	WA	L	29	22	47.0761	-119.3537	
2	BC	L	27	27	49.05	-116.5017	
3	ID	S, L	23	6	42.6120	-113.2473	
4	ID	S, L	31	11	42.6198	-113.2837	
5	ID	L	23	16	42.2568	-112.0132	
6	ID	L	22	13	42.4005	-100.7893	
7	NE	S, L	31	28	41.9393	-102.3872	
8	NE	S, L	22	17	41.9652	-102.339	
9	NE	L	25	17	41.9467	-102.4334	
10	NE	L	23	19	42.2371	-100.0883	
11	NE	L	34	19	42.1734	-98.1558	
12	NE	L	26	20	42.0006	-97.00636	
13	MN	L	32	31	45.7097	-94.89306	
14	MN	S, L	38	36	45.1264	-94.0306	
15	MN	S, L	32	31	45.1917	-93.6917	
16	MN	L	34	31	44.2153	-93.5931	
17	MN	L	36	33	46.9208	-92.1556	
18 (1971)	ONT	T	30	25	43.521	-79.997	23890-1 – 29, 22687-1,2
18 (2001)	ONT	L, T	41	36			
19 (1971)	ONT	T	43	37	44.178	-78.965	2456-1 – 22, 24257-1 – 22
19 (2001)	ONT	L, T	40	40			
20 (1971)	NY	T	36	22	42.989	-76.772	23876-1 – 32, 23877-1 – 10
20 (2001)	NY	L, T	48	40			
21 (1971)	NY	T	54	41	43.4679	-76.01	23842-1 – 15, 23851-1 – 22, 23856-1 – 17
21 (2001)	NY	L, T	23	20			
22 (1979)	ONT	T	39	35	45.0687	-75.653	24332-1 – 39
22 (2001)	ONT	L, T	38	37			

Contemporary population tissue collections consisted entirely of toe clips collected during the summers of 1997–2001 and preserved by desiccation in 1.5 mL tubes filled with drierite desiccant. We recorded colour and collected

tissue from 22 to 48 randomly caught adult and subadult *R. pipiens* from 22 populations across the species range (Table 1, Fig. 1). Historical samples were preserved by F. Schueler from four sites in 1971, and one site in 1979

(Schueler 1982). The samples were preserved as dried skins (Schueler 1981) and catalogued at the Canadian Museum of Nature (CMN; Ottawa, Ontario). Colour locus information from specimens was derived from the field notes associated with each specimen. In all, we collected genetic and colour locus information from 881 frogs in 22 populations (five of which had two collecting time periods; Fig. 1) with an average of 33 frogs per population (Table 1). For analysis, we split samples into eastern and western regions because of a large genetic split in the species' mtDNA phylogeny (Hoffman & Blouin 2004a). By analysing populations within regions, we decreased the probability of homoplasy in the genetic markers and thus increased the likelihood of accurately measuring genetic structure.

For the molecular analysis, total genomic DNA was extracted following a standard phenol/chloroform technique (Sambrook *et al.* 1989). Two molecular marker types were used to acquire neutral-locus genetic structure. Seven microsatellite loci identified from *R. pipiens* (Rpi100, Rpi101, Rpi103, Rpi104, Rpi106, Rpi107, Rpi108) were used under the conditions described in Hoffman *et al.* (2003), and two microsatellite markers originally developed from *Rana pretiosa* (RP193, Monsen & Blouin 2003; RP415, M. Blouin, unpublished data) were used under the conditions in Hoffman & Blouin (2004b). Microsatellite loci were tested for HWE using exact tests in GENEPOP version 3.3 (Raymond & Rousset 1995). Additional markers were developed using the intersimple sequence repeat (ISSR) technique. This technique uses random primers anchored in microsatellite regions of DNA and produces a large number of dominant loci that can be scored reliably (Zietkiewicz *et al.* 1994). ISSR loci were optimized such that each PCR was carried out in a 25 μ L reaction under the following conditions: 3 mM MgCl₂, 1.0 mM each dNTP, 1 \times PCR enhancer (MasterAmp Enhancer, Epicentre), 0.2 μ M primer 812 (UCB primer stock), 0.5 U *Taq* DNA polymerase, and 20 ng DNA. The temperature profile consisted of an initial 3 min denaturation step (94 °C), 40 cycles at 94 °C for 30 s, 46 °C for 45 s, and 72 °C for 2 min, followed by a final extension (72°) for 7 min. All PCRs were carried out in a PerkinElmer 9600. DNA concentrations were quantified with a Hoefer DyNAQuant 200 fluorometer (Hoefer Scientific Instruments). Both microsatellite and ISSR loci were scored with an internal lane standard on an ABI 3100 to increase reliability of scored markers. Data were scored using GENOTYPER 3.2 (Applied Biosystems).

Estimating allele frequencies from dominant loci

Because microsatellite markers showed that each population was in HWE (see Results) the allele frequencies from the ISSR loci and at the colour locus were calculated according to Lynch & Milligan (1994). Lynch and Milligan show that simply taking the square root of q^2 (the frequency of the

absent phenotype at a dominant locus) can give a downwardly biased estimate of q (the allele frequency at that locus). However, as long as Nq^2 (with N equal to the number of sampled individuals per population) is greater than 2, then their alternative method for estimating q (equation 2a in Lynch & Milligan 1994) is less biased (or essentially unbiased if Nq^2 is greater than 3). For the ISSR loci, we estimated all allele frequencies with Nq^2 values greater than 2 as described above. All loci where Nq^2 was less than 2 (approximately 1% of samples) were removed from the analysis as suggested by Lynch & Milligan (1994). For the colour locus, allele frequencies were estimated using two methods. First, we estimated allele frequencies in each population by taking the square root of the frequency of the brown frogs (homozygous recessive state) as the frequency of the brown allele. The frequency of the green allele was calculated as one minus the frequency of the brown allele. Second, we estimated allele frequencies adjusted according to Lynch & Milligan (1994). Both of these estimates of the colour locus allele frequencies were used in our analyses. In general, the simple estimates based on the square root of q^2 give a more liberal estimate of F_{ST} and the Lynch & Milligan (1994) estimates are more conservative (lower F_{ST}). However, the type of estimate never changed the interpretation of the results (see Results).

Spatial analyses

Spatial analysis was carried out using the procedure of Beaumont & Nichols (1996) as implemented by the computer program FDIST2 (located at <http://www.rubic.rdg.ac.uk/~mab/software.html>). Briefly, this program calculates F_{ST} and heterozygosity according to the methods described by Cockerham & Weir (1993). Then, the expected F_{ST} is calculated from the data as the average among loci weighted by their heterozygosity. To build the null model in this study the data set included all the microsatellite and ISSR loci. Coalescent simulations were then performed using samples of the same size and number as the data (i.e. number of populations in the analysis and median sample size of all populations), assuming an island model (with 100 islands) and an infinite alleles mutational model. F_{ST} and heterozygosity were calculated, and this whole procedure was repeated 40 000 times. The confidence limits for the distribution of F_{ST} as a function of heterozygosity was characterized by estimating where the least extreme 95% of the simulated data points were expected to lie (Beaumont & Nichols 1996). By plotting the observed values for the loci used in the genetic analysis, we determined whether any loci were outside these 95% confidence limits.

Using this Beaumont & Nichols (1996) technique, we conducted two different analyses. First, we tested for variation across broad geographical ranges. This large-scale analysis included all populations in the west region or the east

region, respectively (see Fig. 1). Second, we tested for variation across three small geographical ranges (mean distance between each population pair = 12 km; Fig. 1). By analysing populations across this broad scope of distances we hoped to increase the likelihood of finding evidence for selection. In all of these analyses, the colour locus was removed when simulating the expected distribution, and then plotted on the distribution to determine if it fell outside of the 95% confidence intervals of the simulated distribution.

Temporal analyses

The goal of the temporal analysis was to determine whether the allele frequency at the colour locus had changed more or less over time than could be explained by genetic drift alone. The estimate of genetic drift expected at the colour locus in each population was based on the effective population size (N_e) of each population, the starting allele frequency of the colour locus, the final allele frequency at the colour locus, and the number of generations that had ensued between sampling periods in the populations sampled. The starting allele frequency of the colour locus was calculated from the museum data (see Table 2). In obtaining the final allele frequency at the colour locus we could not use the same frequencies as in the spatial analysis, because we had to account for a discrepancy in the way frog colour was scored originally in 1971 as compared to the 2001 samples. In 1971, the museum specimens were recorded as brown even if they contained small amounts of green. However, it is now known that the presence of any green on the dorsum is indicative of a genetically green frog (Corn 1982). Therefore, to maximize similarity in scoring methods between the 1971 and 2001 samples, green frogs that were scored as predominantly brown (in the 2001 collection by FWS) were scored as brown frogs (as was originally done by FWS in the 1970s).

Table 2 Table of parameters used in simulations of temporal change in allele frequency. P_0 is the observed allele frequency at the colour locus at generation 0. N_e is the effective population size (see text). 'Generations' denotes the number of generations of simulated allele frequency change. P' is the observed allele frequency at the colour locus after n ($n = 15$ or 11) generations. ΔP is the observed change in allele frequency at the colour locus. The observed variance in ΔP is 0.004

Population	P_0	N_e		Generations	P'	ΔP
		Low	High			
18	0.59	205	420	15	0.53	-0.06
19	0.63	324	588	15	0.72	0.09
20	0.38	469	1820	15	0.36	-0.02
21	0.50	102	410	15	0.54	0.04
22	0.68	243	1019	11	0.64	-0.04

The number of generations between sampling periods was calculated assuming 2 years per generation (Ryan 1953; Gilbert *et al.* 1994). To determine the extent that N_e influences the simulations, we used both low and high estimates of N_e as calculated using the temporal method [using the Waples (1989) (high estimate) and Wang (2001) (low estimate) approaches (see Table 2) as calculated in Hoffman *et al.* (2004)]. Hoffman *et al.* (2004) concluded that these methods provide the most accurate estimates of N_e and that the effective sizes of typical *R. pipiens* populations are in the range of a few hundreds to a few thousand frogs.

For each population, we used estimated N_e , the starting allele frequency of the colour locus, and the number of generations between sampling periods in a Monte Carlo simulation of genetic drift for a bi-allelic locus. In a given generation, we drew a uniformly distributed random number between zero and one for each chromosome present in the simulated diploid population. If the number was between zero and the frequency of allele A in the parents, then gamete A was produced. If the number was between the frequency of A and one, then gamete B was produced. This procedure was repeated until the total number of gametes equalled $2N_e$. The new frequency of allele A (after one generation of drift) was calculated from this pool of chosen gametes. In addition to running simulations in the absence of migration (as described above), we ran simulations in which migration (estimated from the observed F_{ST} ; assuming an island model) was allowed to homogenize allele frequencies (after the drift phase of the model). Although the assumptions of an island model are seldom met (Whitlock & McCauley 1999) the island model should provide a liberal estimate of the effect of migration on our analyses. We repeated this process for multiple generations to simulate drift over the desired span of time (Hedrick 2000).

To determine whether the change at the colour locus from the five observed populations deviated significantly from expectations under drift (or drift and migration), we modelled the simultaneous evolution of five populations with parameters equal to the observed populations (for model parameters see Table 2). From these simulations, we calculated the change in the allele frequency (Δp) for each population and the variance in Δp (Δp_{var}) among all five simulated populations. The simulation was repeated 10 000 times to generate a distribution of Δp_{var} under the hypothesis of neutrality. The observed estimate of Δp_{var} at the colour locus from the five populations was then compared to the distribution of simulated Δp_{var} . An observed value within the 2.5% tail at the low end of the distribution would indicate significantly low values of Δp_{var} likely owing to stabilizing selection or unidirectional selection (in the same direction) for all five populations. An observed value within the 2.5% tail at the high end of the distribution would indicate significantly high values of Δp_{var} likely owing to distinct directional selection pressures in different

populations. An observed value falling within the 95% body of the distribution would indicate that Δp_{var} at the colour locus could not be differentiated from a null hypothesis of variation as expected from genetic drift (or drift and migration).

Simulations to determine detectable selection coefficients

Spatial simulations. To determine our ability to detect selection over a large spatial scale, we simulated the effects on F_{ST} of selection in an array of evolving populations. We focused on the effects of two likely mechanisms of balancing selection: negative frequency-dependent selection and overdominance (Allen 1988; Hedrick 2000). The guiding principle behind these analyses is that balancing selection will tend to promote convergence in allele frequencies among populations. Therefore, if the value of the upper 0.9 quantile of the simulated F_{ST} , for a given value of the selection coefficient (s) is less than the observed 0.025 quantile for the distribution of neutral alleles (calculated from FDIST2 for a value of expected heterozygosity (H_E) at the colour locus), then that value of s is the minimum strength of selection coefficient that would be statistically detectable in the present analysis. In other words, 90% of the simulated F_{ST} values under selection would fall outside the 95% confidence interval of observed F_{ST} derived from neutral markers. For these simulations, the number of simulated populations was 17 (the same as in the large scale spatial analysis), effective population size was set to include a range of effective sizes typical for *R. pipiens* ($N_e = 250, 1000$; Hoffman *et al.* 2004), and the starting allele frequency was set to 0.5. We ran simulations in both the presence and absence of migration. The migration rate, m , was estimated from F_{ST} of the large-scale spatial analysis, assuming an island model of migration. We varied the values of selection coefficients to simulate a wide array of selection strengths.

The basic drift model described above was modified to include a change in allele frequencies due to one of two types of selection before the random sampling of gametes each generation. We assume that (i) one allele is completely dominant to the other; (ii) s_1 and s_2 represent the selection coefficients associated with the dominant phenotype and the recessive phenotype, respectively; and (iii) the fitness of each phenotype decreases linearly as a function of its frequency. The equation for the single-generation change in the allele frequency for frequency-dependent selection is:

$$\Delta q = q^2(1 - q)[(s_1 + s_2)q^2 - s_1]/\bar{W}, \quad (\text{eqn 1})$$

where

$$\bar{W} = (1 + s_1) - 2s_1q^2 + (s_1 + s_2)q^4. \quad (\text{eqn 2})$$

For overdominance, assuming a two-allele system with fitnesses of 1, $1 - s_1$ and $1 - s_2$ for the heterozygote and two

homozygous genotypes, respectively, the equation for the change in the allele frequency over time (from Futuyma 1986) is:

$$\Delta q = [pq(s_1p - s_2q)]/(1 - s_1p^2 - s_2q^2). \quad (\text{eqn 3})$$

The simulation allowed allele frequencies to evolve according to selection and drift (and migration when present) for a random number of generations (between 500 and 700 generations in our analyses). In this type of simulation, with no mutation, all populations will eventually go to fixation, so we chose a number of generations that allowed the populations to achieve a quasi-equilibrium due to selection and drift. The number of generations required to reach quasi-equilibrium was determined by observing the dynamics of F_{ST} under different parameter values. In all cases, the populations appeared to reach quasi-equilibrium in less than 500 generations. Seventeen allele frequencies (representing the observed 17 populations) were obtained at the end of the simulation, and F_{ST} was calculated for that set of populations. In all cases, we assumed that the selection coefficient s_1 was equal to s_2 . For each value of interest for s_1 and s_2 , the simulation was run 200 times to obtain estimates of mean F_{ST} and 80% confidence intervals for a given selection coefficient for both frequency-dependent selection (Fig. 5a) and overdominance (Fig. 5b).

Temporal simulations

To determine the strength of selection necessary to cause detectable temporal variation between neutral loci and the colour locus, we again simulated the effects of frequency-dependent selection and overdominance. General model parameters were identical to those used to estimate Δp_{var} in the original temporal analysis simulation. The number of populations simulated was five, N_e was set to the both the high and low value estimated for each population (Table 2), and the equilibrium allele frequency was set to the observed starting allele frequency in each population. Once again, we ran these simulations with and without migration. For frequency-dependent selection and overdominance, the equation for the change in allele frequency over time (Δq) was the same as in the spatial simulations. The value of s_1 was set to a value of interest (ranging from -0.025 to -0.50 for frequency-dependent selection and 0.025 – 0.50 for overdominance), and the value of s_2 was calculated assuming that each population was initially at equilibrium. Thus, the populations were allowed to have different equilibria and different values of s_2 , but s_1 was held constant across populations as a general indicator of the strength of selection. The equilibrium gene frequency for frequency-dependent selection (from Wright 1984) is:

$$\hat{q} = \sqrt{s_1/(s_1 + s_2)}, \quad (\text{eqn 4})$$

so

$$s_2 = (s_1 - \hat{q}^2 s_1) / \hat{q}^2, \quad (\text{eqn 5})$$

whereas for overdominance (from Futuyma 1986):

$$\hat{q} = s_1 / (s_1 + s_2), \quad (\text{eqn 6})$$

so

$$s_2 = [s_1(1 - \hat{q})] / \hat{q}. \quad (\text{eqn 7})$$

The simulation allowed each population's allele frequencies to evolve according to drift and selection (and migration when present) as in the spatial simulations of selection for the number of generations observed (11 or 15). At the end of the simulated number of generations Δp was measured in each population and Δp_{var} was calculated for all populations. The entire simulation was repeated 10 000 times, and the upper and lower 95% confidence limits were calculated for each value of the selection coefficient, as in the temporal simulations of drift described above. Selection coefficients producing simulation runs in which the upper 95% confidence limit of the simulated Δp_{var} was less than the observed Δp_{var} at the colour locus (0.004) indicated the minimum value of the selection coefficient that would have been statistically detectable in this study. In other words, we can use our data to reject the selection models that produced 95% confidence intervals that did not encompass the observed value of Δp_{var} from the natural populations.

Results

Spatial analysis

We found very few departures from Hardy–Weinberg equilibrium (HWE) at the microsatellite loci. Only two populations exhibited loci that deviated significantly from HWE after a sequential Bonferroni correction (Rice 1989). In population 20, two loci deviated from HWE, and in population 2, one locus deviated from HWE. Because of the rarity of loci out of HWE, we included all loci from all populations in the *FDIST2* analyses. Nineteen ISSR loci showed strong and repeatable electropherograms with *GENOTYPER* 3.2 and were scored for use in the *FDIST2* analyses. As dominant markers, the ISSR loci were scored such that the presence of the band signified either the homozygous dominant or the heterozygous state, and the absence of the band indicated the homozygous recessive state. Hardy–Weinberg equilibrium could not be directly assessed, but was assumed given the population data from the microsatellites. Not all ISSR loci were present in all populations. Therefore, when we conducted analyses involving populations in which an ISSR locus was absent,

we dropped that locus from the analysis, changing the total number of loci used among analyses.

In the western region, expected F_{ST} at neutral loci was calculated from the data as 0.184 for 17 populations (with 26 loci) and a median sample size of 29. Figure 2 shows the 95% confidence limits for the relationship between F_{ST} and heterozygosity for simulated neutral loci. The estimates of F_{ST} as a function of heterozygosity at the colour locus (simple, square root of q^2 , estimate of $F_{\text{ST}} = 0.197$; Lynch and Milligan estimate of $F_{\text{ST}} = 0.114$) appear squarely within the 95% confidence limits of the expected distribution. In the eastern region, expected F_{ST} for the neutral markers

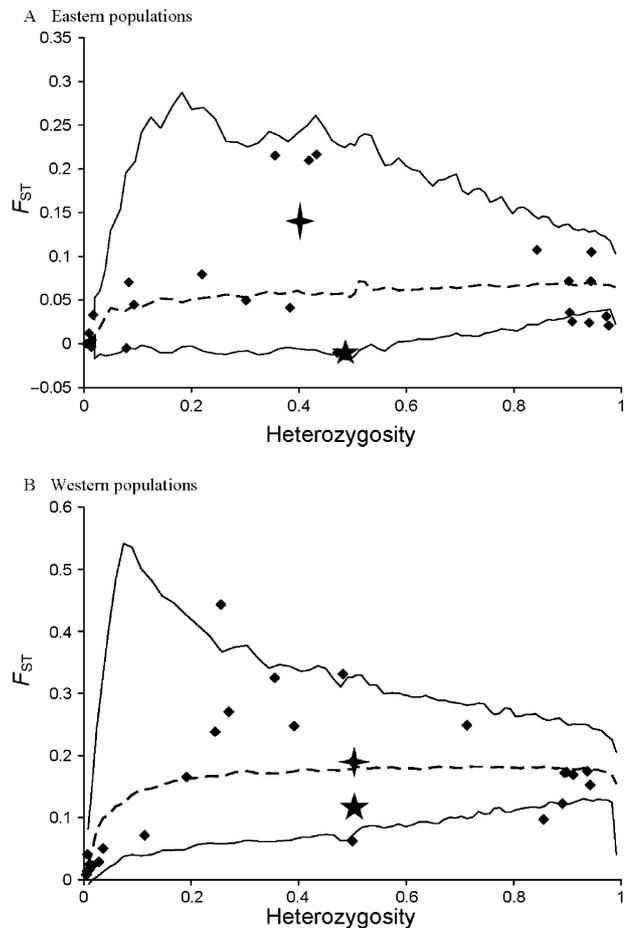


Fig. 2 *FDIST2* results of F_{ST} for simulated neutral markers as a function of heterozygosity in the large-scale spatial analysis. The upper and lower lines indicate 0.975 and 0.025 quantiles, the middle line indicates the median F_{ST} . Points on the figure indicate values of F_{ST} and heterozygosity for molecular markers used to create the distribution. The stars indicate the colour locus F_{ST} (five-pointed stars indicate the Lynch and Milligan estimates, four-pointed stars indicate the square root of q^2 estimate of F_{ST} and heterozygosity). Panel (a) indicates the east region and panel (b) indicates the west region. Note that the removal of the two most divergent populations (owing to the Lynch and Milligan correction) causes the striking change of colour locus F_{ST} in the east region.

was calculated from the data as 0.0711 for the five eastern populations (with 25 loci) and a median sample size of 40. As in the western region, the estimates of F_{ST} as a function of heterozygosity at the colour locus (simple estimate of $F_{ST} = 0.1398$, Lynch and Milligan estimate of $F_{ST} = -0.010$) also fall well within the expected distribution (Fig. 2a). Here, these values of F_{ST} at the colour locus are quite different because of the removal of the two most divergent populations owing to the Lynch and Milligan correction. Therefore, we found no evidence that selection was acting to homogenize or disrupt allele frequencies given both our liberal estimate of the colour locus F_{ST} (the estimate based on the square root of q^2) and our more conservative estimate of the colour locus F_{ST} (the Lynch and Milligan estimate). In both the east and west regions, some putatively neutral markers used to create the expected distribution fell outside the 95% expected range. The majority of these markers are microsatellites and likely fall outside the expected distribution as a consequence of their large numbers of alleles and hence high heterozygosities. The values observed for three loci in the western distribution cannot be explained in this way. However, these loci may fall outside the expected distribution due to chance (we would expect approximately two loci to fall outside the 95% confidence limits by chance alone) or may represent loci linked to parts of the genome influenced by selection.

For the pairwise small-scale analysis, mean expected F_{ST} at neutral loci was calculated as 0.016 (with 18 loci), 0.013 (with 23 loci), and 0.049 (with 14 loci), respectively, for the population pairs in Nebraska, Minnesota, and Idaho. In each of these comparisons, two populations were analysed with median sample sizes of 26 (Nebraska), 35 (Minnesota), and 26 (Idaho). In each comparison plot of F_{ST} as a function of heterozygosity, the colour locus again falls in the middle of the expected distribution (Fig. 3; mean F_{ST} of colour locus: Nebraska, square root of q^2 estimate $F_{ST} = 0.041$, Lynch and Milligan estimate $F_{ST} = 0.044$; Minnesota, square root of q^2 estimate $F_{ST} = -0.010$, Lynch and Milligan estimate $F_{ST} = -0.008$; Idaho, square root of q^2 estimate, Lynch and Milligan estimate $F_{ST} = 0.092$). Once again, we find no evidence that selection is either homogenizing allele frequencies or maintaining different allele frequencies in each population. However, it is important to note that the 95% confidence intervals of the expected distribution overlapped zero, making homogenizing selection undetectable by this technique in these small-scale population comparisons (see Discussion).

Temporal analysis

Only seven microsatellite loci and none of the ISSR loci amplified consistently with the historical DNA samples (microsatellite loci that did not amplify were Rpi104 and Rpi107). Therefore, only seven microsatellite loci were used in the temporal analysis. The maximum-likelihood

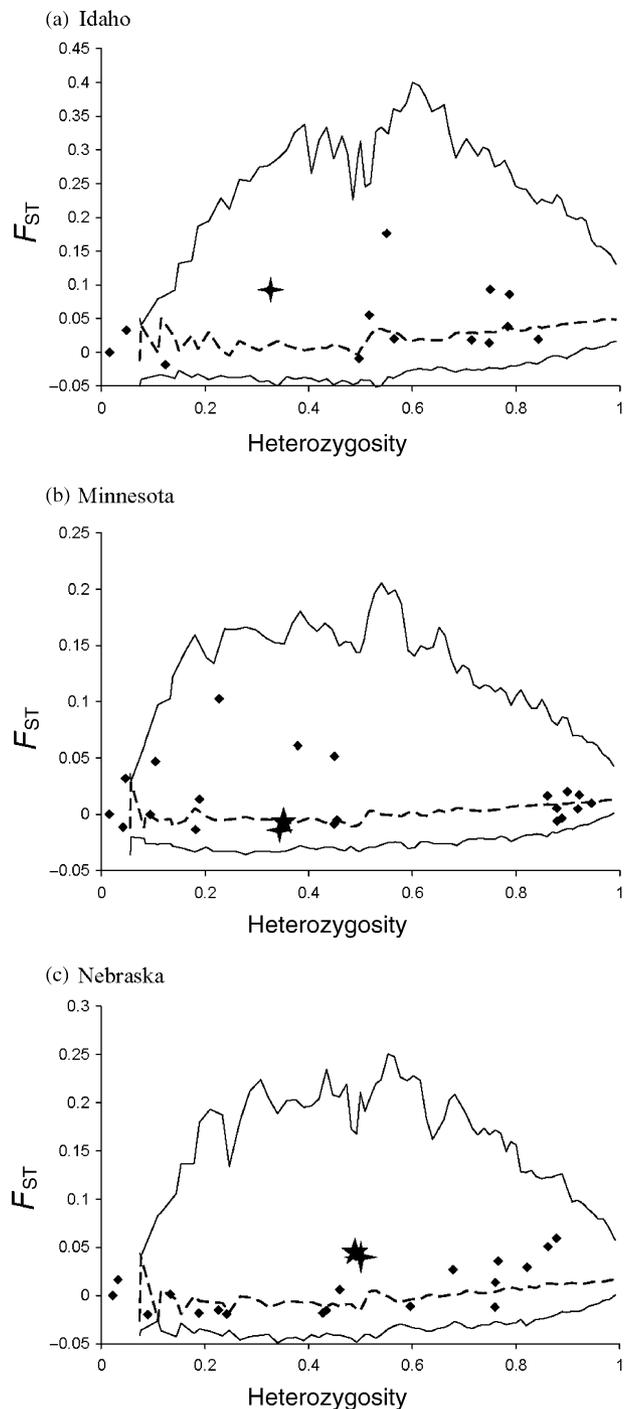


Fig. 3 F_{ST} results of F_{ST} as a function of heterozygosity in the small-scale spatial analysis. The upper and lower lines indicate 0.975 and 0.025 quantiles, the middle line indicates the medium F_{ST} . Points on the figure indicate values of F_{ST} and heterozygosity for molecular markers used to create the distribution. The stars indicate the colour locus F_{ST} (five-pointed stars indicate the Lynch and Milligan estimates, four-pointed stars indicate the square root of q^2 estimate of F_{ST} and heterozygosity). Panel (a) indicates the Idaho populations, panel (b) indicates the Minnesota populations, and panel (c) indicates the Nebraska populations.

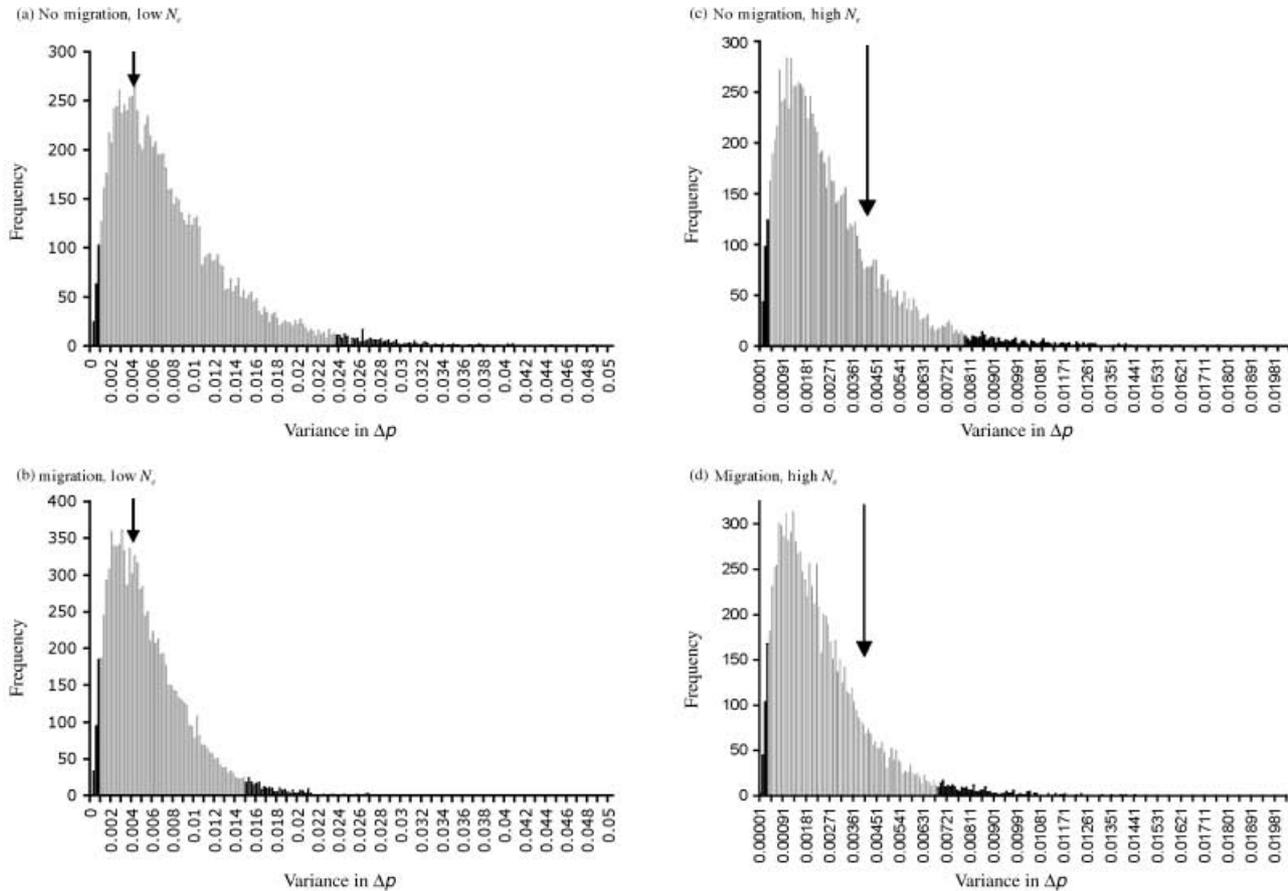


Fig. 4 Graph of variance of Δp due to genetic drift with low N_e without migration (a) and with migration (b) or with high N_e without migration (c) and with migration (d) for five populations simulated 10 000 times with parameters similar to those of the observed populations (see text). Arrow indicates the value of the point estimate for the variance in Δp at the colour locus among the five observed populations. Darkened tails of graph indicate values outside the 95% least extreme simulated values. As indicated by the location of the arrow, the observed Δp_{var} at the colour locus lies in the middle of the distribution of populations simulated in the absence of selection.

estimates of N_e are listed in Table 2. Figure 4 displays the distribution of Δp_{var} (with and without migration) when the parameters were used to simulate genetic drift at a bi-allelic locus for each of five populations. For the low N_e regime, 95% of Δp_{var} values fell between 0.00089 and 0.0231 in the absence of migration and fell between 0.00067 and 0.0150 in the presence of migration. For the high N_e regime, 95% of Δp_{var} values fell between 0.00027 and 0.0079 in the absence of migration and fell between 0.00025 and 0.0068 in the presence of migration. The Δp_{var} from the five observed populations for the colour locus was 0.004. Once again, the observed value of change at the colour locus (Δp_{var}) falls well within the distribution expected from putatively neutral loci.

Simulations to determine detectable selection coefficients

Spatial simulations indicated that global F_{ST} responds as expected to increasing values of selection coefficients

(Fig. 5). Moreover, in the presence of migration (Fig. 5b, d) the value of F_{ST} with no selection is equal to F_{ST} observed from the molecular markers, whereas in the absence of migration (Fig. 5a, c) F_{ST} tends toward 1.0 with low selection and low N_e , indicating that our parameters and simulations are accurately describing the system. F_{ST} does not reach 1.0 for all simulations with no migration and no selection because large values of N_e limit the change in allele frequencies owing to genetic drift. As expected, migration acts to homogenize allele frequencies and thus decreases F_{ST} relative to the same simulation parameters without migration (see Fig. 5). The simulations with migration are likely a more accurate simulation of the actual populations, because previous studies have suggested that migration does indeed connect leopard frog populations (Hoffman & Blouin 2004a; Hoffman *et al.* 2004). Therefore, the simulations without migration represent a conservative model of the effects of selection on F_{ST} . Finally, the simulations indicated that the difference between the lower and upper estimates of N_e is

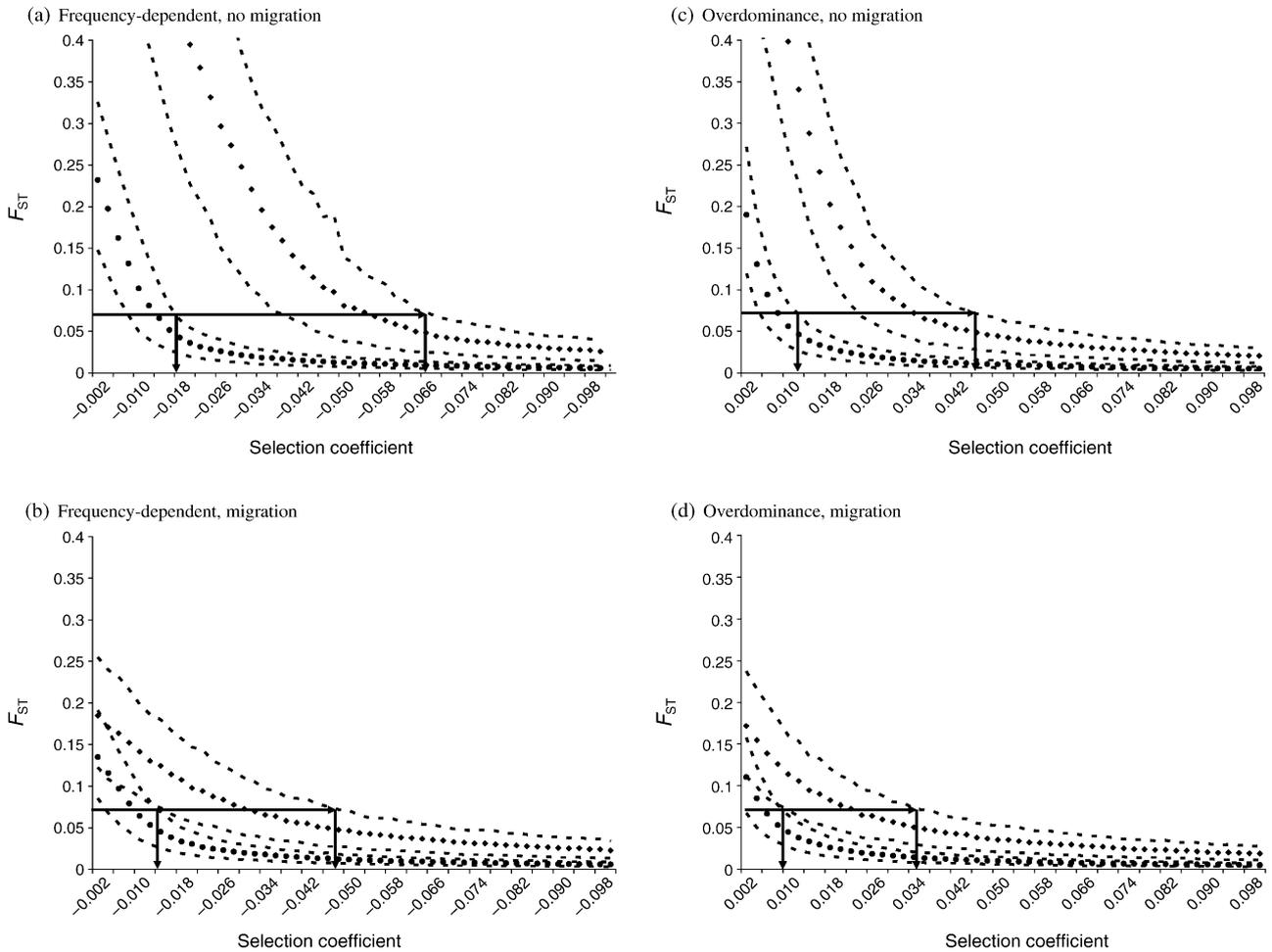


Fig. 5 Graphs of F_{ST} as a function of selection coefficient from simulations of (a) frequency-dependent selection and no migration (b) frequency-dependent selection and migration (c) overdominance and no migration, and (d) overdominance and migration. In all graphs, diamonds and circles (with corresponding dashed lines) indicate mean F_{ST} (and 80% confidence limits) from simulations run with N_e set to 250 and 1000, respectively. The horizontal dashed line indicates the value of the lower 0.025 quantile of observed value of F_{ST} from the colour locus (0.071, see Fig. 2 western populations) and the vertical arrows indicate the value of selection where 90% of the simulated F_{ST} values, given that strength of selection, would fall outside the 95% confidence interval of observed F_{ST} derived from neutral markers. Therefore, values of selection greater than those marked by the arrows would be sufficient to cause detectable differences in F_{ST} between neutral loci and loci under balancing selection (see text).

about a three- to sixfold increase in the selection coefficient (Fig. 5). The simulations indicated that if the strength of selection was stronger than would be produced by selection coefficients (i.e. s_1 and s_2) of about -0.014 (with migration and $N_e = 1000$) under frequency-dependent selection or 0.008 (with migration and $N_e = 1000$) under overdominance, then 90% of simulated population differentiation (F_{ST}) values would be less than the 0.025 quantile observed for neutral loci ($F_{ST} = 0.077$, given a value of $H_{E'}'$ similar to that at the colour locus; see Fig. 2). Therefore, if any selection is acting on the colour polymorphism, the intensity of selection is likely weaker than what would be provided by selection coefficients of approximately -0.014 (for frequency dependence) or 0.008 (for overdominance). Our most conservative

estimates of the strength of selection that would have been detectable in our analysis are about -0.064 (without migration and $N_e = 250$) under frequency-dependent selection or about 0.044 (without migration and $N_e = 250$) under overdominance (Fig. 5).

The simulation-based analysis of the temporal data suggests that we have less capacity to detect selection using the temporal method than using the spatial analyses. Surprisingly, in the temporal analysis the presence or absence of migration has almost no effect on the results. Simulations using the lower values of N_e indicated that coefficients of selection greater than -0.4 for frequency-dependent selection or 0.4 for overdominance would provide upper 95% confidence limits of Δp_{var} that do not overlap the observed

Table 3 Coefficient of selection (s_1) and upper and lower 95% confidence intervals for Δp_{var} from temporal simulations that used one of two models of selection: frequency-dependent selection or overdominance. Simulation parameters and N_e regime are identical to those given in Table 2, except the selection coefficient was included in the model (see text). The point estimate of Δp_{var} for the colour locus was 0.004. Therefore, boldfaced values of the selection coefficient indicate models of selection that can be rejected on the basis of our empirical observations

Model of selection	N_e regime	Selection coefficient	Δp_{var} no migration		Δp_{var} with migration	
			Lower 95% CI	Upper 95% CI	Lower 95% CI	Upper 95% CI
Frequency-dependence	Low	-0.05	0.00062	0.01753	0.00050	0.01196
		-0.1	0.00045	0.01386	0.00040	0.00959
		-0.2	0.00030	0.00901	0.00025	0.00670
		-0.3	0.00019	0.00583	0.00017	0.00476
		-0.4	0.00013	0.00402	0.00013	0.00347
Frequency-dependence	High	-0.5	0.00010	0.00295	0.00009	0.00261
		-0.025	0.00023	0.00669	0.00022	0.00571
		-0.05	0.00020	0.00584	0.00020	0.00508
		-0.075	0.00018	0.00500	0.00017	0.00445
		-0.1	0.00015	0.00457	0.00015	0.00389
Overdominance	Low	-0.125	0.00013	0.00390	0.00013	0.00329
		-0.15	0.00012	0.00348	0.00011	0.00298
		0.05	0.00062	0.01667	0.00049	0.01162
		0.1	0.00045	0.01317	0.00039	0.00930
		0.2	0.00028	0.00807	0.00025	0.00621
Overdominance	High	0.3	0.00018	0.00549	0.00017	0.00431
		0.4	0.00013	0.00400	0.00012	0.00334
		0.5	0.00011	0.00315	0.00010	0.00272
		0.025	0.00024	0.00659	0.00022	0.00585
		0.05	0.00021	0.00576	0.00019	0.00494
		0.075	0.00018	0.00477	0.00016	0.00417
		0.1	0.00015	0.00419	0.00013	0.00384
		0.125	0.00013	0.00352	0.00013	0.00330
		0.15	0.00012	0.00319	0.00010	0.00293

value of Δp_{var} at the colour locus (Table 3). With higher values of N_e on the other hand, coefficients of selection greater than -0.10 for frequency-dependent selection or 0.10 for overdominance would provide upper 95% confidence limits of Δp_{var} that do not overlap the observed value of Δp_{var} at the colour locus (Table 3).

Discussion

Detecting evidence for selection

This study exemplifies the use of the locus comparison approach on spatial and temporal samples to detect selection. In this particular case of a colour polymorphism in *Rana pipiens*, we found no evidence for any kind of selection on the colour locus. However, our approach can be applied to any putatively selected locus. One of the major advances in our study is the introduction of a simulation-based method to determine the lower limits of selection differentials that would be detectable by the locus comparison method. Thus, we can say that if stabilizing selection is responsible for maintaining the colour polymorphism in *R. pipiens*, then

the strength of selection must be relatively weak. Our recommendation is that future studies testing for evidence of selection via the locus comparison approach use these same types of simulations to investigate the power of their analyses.

One interesting result of this study is the indication that the spatial technique is much more likely to detect selection than the temporal method in our system. For our data, the temporal method requires a selection coefficient approximately 10–20 times greater than that for the spatial method in order for selection to be detectable. The reason that the spatial analysis is more powerful stems from the nature of our samples. The spatial analysis involves samples collected across a wide geographical range, and distinct populations have been separated for long enough periods of time for them to diverge genetically at neutral loci. Consequently, the spatial analysis is able to detect all but weak balancing selection in the western region. On the other hand, the temporal analysis only covers at most 15 generations, which is not much time for divergence at neutral loci to occur. Thus, selection would have to be quite strong for the selected loci to display a pattern different

from neutrality over such a short time period. One message of this study is that for the locus comparison method to work well on temporal samples, the amount of time elapsed between the historical and present-day samples should probably be much greater than 15 generations unless selection is extremely strong.

Our conclusion that we had great power to detect balancing selection in the spatial analysis appears to be in conflict with the results of a recent study conducted by Beaumont & Balding (2004), who concluded that it is difficult to detect balancing selection using the Beaumont & Nichols (1996) approach. The conclusions of Beaumont & Balding (2004) are based on an analysis of the conditions that cause the values obtained for the lower 95% confidence limit from the Beaumont and Nichols (1996) approach to fall close to, or even below zero. Clearly, stabilizing selection is impossible to detect when the lower confidence limit for F_{ST} falls below zero, and we observed just such a phenomenon in our data from the east region as well as in the small-scale population analysis (see Figs 2 and 3). However, the Beaumont & Balding (2004) criticism is not a problem for the western region analysis because the lower 95% confidence interval lies well above zero. Indeed, the advantage of using the simulations is that they show clearly that relatively weak selection would be sufficient to drive F_{ST} below the lower 95% confidence interval. Thus, even though Beaumont & Balding (2004) are correct that balancing selection often will be difficult to detect using the Beaumont & Nichols (1996) technique, there are certainly circumstances under which it will be detectable (as in our western region). Furthermore, it will be clear from the analysis of the neutral loci whether or not the approach is capable of detecting balancing selection for any particular set of markers, because the location of the lower confidence limit relative to the origin is easily determined.

Two potentially important parameters in the simulations of selection are the migration rate and the effective population size, because they can be difficult to measure in natural populations, and they certainly affect the dynamics of drift, population divergence and selection. The inclusion of migration had a relatively minor effect on the interpretation of our simulations. In general, migration homogenizes allele frequencies (as does balancing selection), so weaker selection is necessary to pull the colour locus F_{ST} out of the expected distribution of F_{ST} generated by the neutral loci. However, the effect in our case was relatively small in the spatial analysis and almost nonexistent in the temporal analysis. Thus, for most situations in which balancing selection is suspected, the inclusion of migration in the power simulations may not be necessary, because the simulations without migration will provide a conservative estimate of the power to detect selection. One important aspect of the power simulations is that they require at least a ballpark estimate of the effective population size, because the rate at

which populations diverge by genetic drift as well as the efficiency of selection are affected by the N_e . However, it requires large errors, such as one order of magnitude or more, for the results of the simulations to be seriously misleading.

Evolution of the colour polymorphism in R. pipiens

Empirically, this study measured differences in genetic structure owing to the signature of selection at a locus controlling a green/brown dorsal colour polymorphism in the northern leopard frog. The results of our temporal and spatial analyses indicate that patterns of genetic variation at the colour locus of the northern leopard frog do not deviate from those expected at a locus evolving in response to genetic drift alone. Hence, not all striking colour polymorphisms may be selectively maintained. When investigating the maintenance of these polymorphisms, one should first rule out a nonselective, neutral hypothesis and thus obtain supporting evidence of selection, rather than diving straight into investigations of selective mechanisms.

The results of this study beg the question of why Schueler (1979) and Corn (1981) found evidence in support of a selective maintenance of the colour polymorphism (see Introduction). Both of these studies (Schueler 1979; Corn 1981) were correlational in nature and may have reported coincidental correlations. Schueler's study provides a testable hypothesis, namely that green frogs are more frequent at sites with greater forest area and topographic relief below the collection site, and brown frogs are more frequent in areas of marsh and large lakes (Schueler 1979). Future studies could test this hypothesis either temporally by revisiting the original sites to see if the correlation still holds true, or spatially by testing these variables across a broader range of populations over a greater spatial scale than originally tested. The study that found variation in dates of metamorphosis among colour morphs (Corn 1981) can be explained by at least one other hypothesis. Because Corn (1981) did not directly measure length of larval period, we do not know absolute differences in length of larval period of metamorphosed frogs. Since females lay eggs in masses, emergence dates of siblings (of the same colour) are likely to be similar. Thus, a few brown females laying eggs early in the season could be the cause of the observation that a greater percentage of early metamorphs were brown.

Do our results signify that selection is not acting on the colour locus in the northern leopard frog? Not necessarily – the results only indicate that genetic variation at the colour locus does not differ from expectations under the null hypothesis of genetic drift. It is possible that selection is still maintaining the colour polymorphism but we did not detect it. However, if among-site stabilizing selection were maintaining the polymorphism, then we would expect genetic variation across large distances to be larger at

neutral loci than at the colour locus. If each population in the study had a different equilibrium allele frequency at the colour locus, we might expect increased variation among populations regardless of distance separating populations. The result of this pattern of selection would be increased genetic structure at the colour locus. However, under these circumstances, one might expect to observe evidence of local adaptation by divergent allele frequencies at the colour locus for populations separated by small distances.

If the green/brown colour polymorphism is indeed selectively maintained, then two likely scenarios could explain why we did not detect it. First, the selective mechanism maintaining the polymorphism could be something other than overdominance or frequency-dependent selection. Therefore, the simulation results, which indicated that we should detect all but very weak selection, could be based on the wrong model of selection. Under these circumstances, we cannot predict the relative strength of the selection coefficient we could detect. Additionally, it is possible that there are population differences in selection regimes through time and/or space. However, testing for such a pattern would be impossible given the wide range of expectations that could arise from such a process. Second, the selection may be so weak that although the polymorphism is actually selectively maintained, the allele frequencies at the colour locus fluctuate such that allele frequency variation is indistinguishable from that caused by genetic drift. Kingsolver *et al.* (2001) recently reviewed the strength of selection in natural populations on quantitative traits and concluded that the median absolute value of selection gradients for quadratic selection (i.e. disruptive or stabilizing selection) was 0.10. Additionally, Endler (1986) concluded that selection coefficients for polymorphic traits in natural populations greater than 0.1 are probably common.

General conclusions

In the past, the approach to studying selection has been to begin with the assumption that interpopulation differences result from selection, and then studying the mechanisms by which selection maintains those differences. However, the results of this study clearly indicate that the assumption of a selective maintenance should be revisited.

What can we conclude about the locus comparison approach as a means to detect evidence of natural selection on single-locus polymorphisms? One important feature of the locus comparison approach is that it is applicable for a wide range of species and is not predestined to find selective differences between morphological and molecular markers. Four studies have analysed selection on a colour polymorphism using the locus comparison approach, one on the happy-face spider (Gillespie & Oxford 1998), two on

damselflies (Andres *et al.* 2000; Andres *et al.* 2002) and this study on northern leopard frogs. The three nonfrog studies did indeed find evidence for balancing selection. The locus comparison approach has two general advantages over the classic approach of measuring various mechanisms of selection as a method of identifying selective maintenance. First, the locus comparison approach is less likely to conclude that a polymorphism is selectively maintained owing to spurious correlations. Many studies in the literature report evidence of selective maintenance owing to habitat matching (e.g. Calvar & Bradley 1991; Forsman & Shine 1995), correlations with fitness traits (e.g. Shine *et al.* 1998), seasonal change in morph frequency (e.g. Osawa & Nishida 1992), and differential influence of predators (e.g. Stimson & Berman 1990; Sandoval 1994). The locus comparison approach may show that in some of these species there actually is no evidence for selective maintenance (as was found for *R. pipiens*). Second, the locus comparison approach is more likely to find evidence for selection when the mechanism maintaining selection is unknown. For instance, Fincke (1994) tested various hypotheses for the selective maintenance of a colour polymorphism in a damselfly and concluded that there was no evidence for the selective maintenance of the polymorphism. However, Andres *et al.* (2000) and Andres *et al.* (2002) did indeed find evidence for the selective maintenance of similar colour polymorphisms in other damselflies using the locus-comparison approach. Overall, the locus comparison approach provides a strong, valid method to detect even relatively weak selection in natural populations. Therefore, this method, rather than ad hoc studies of various selective mechanisms, should be the first step in the study of selection in nature.

Acknowledgements

We would like to thank M. Steigerwald for assistance in sampling the CMN collections. The historical collections were supported by National Research Council of Canada grant A5999 to J. D. Rising at the Department of Zoology of the University of Toronto, and by the Herpetology Section of the National Museum of Natural Sciences (now CMN) under the direction of F. Cook. We would also like to thank K. Mosen for comments on early versions manuscript. This manuscript was improved by the comments of P. Jarne, G. Hewitt, and four anonymous reviewers. This research was funded by USGS contract 00HQAG0027 to MB and WDFW grant 58300726 to EH.

References

- Akey JM, Zhang G, Zhang K, Jin L, Shriver MD (2002) Interrogating a high-density SNP map for signatures of natural selection. *Genome Research*, **12**, 1805–1814.
- Allen JA (1988) Frequency-dependent selection by predators. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **319**, 485–503.

- Andres JA, Sanchez-Guillen RA, Rivera AC (2000) Molecular evidence for selection on female color polymorphism in the damselfly *Ischnura graellsii*. *Evolution*, **54**, 2156–2161.
- Andres JA, Sanchez-Guillen RA, Rivera AC (2002) Evolution of female colour polymorphism in damselflies: testing the hypothesis. *Animal Behavior*, **63**, 677–685.
- Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.
- Beaumont MA, Nichols RA (1996) Evaluating loci for use in genetic analysis of population structure. *Proceedings of the Royal Society London. Series B, Biological Sciences*, **263**, 1619–1626.
- Bonnin I, Prosperi J, Olivier I (1996) Genetic markers and quantitative genetic variation in *Medicago truncatula* (Leguminosae): a comparative analysis of population structure. *Genetics*, **143**, 1795–1805.
- Bowcock AM, Kidd JR, Mountain JL *et al.* (1991) Drift, admixture, and selection in human evolution: a study with DNA polymorphism. *Proceedings of the National Academy of Sciences USA*, **88**, 839–843.
- Calvar Bradley JS (1991) Can microhabitat selection explain sex-related color morph frequencies in the grasshopper *Acrida conica* Faricius? *Animal Behavior*, **41**, 1101–1102.
- Cavalli-Sforza LL (1966) Population structure and human evolution. *Proceedings of the Royal Society London. Series B, Biological Sciences*, **164**, 362–379.
- Charbonnel N, Quesnoit M, Razatavonjizay R, Bremond P, Jarne P (2002) A spatial and temporal approach to microevolutionary forces affecting population biology in the freshwater snail *Biomphalaria pfeifferi*. *American Naturalist*, **160**, 741–755.
- Chevillon C, Pasteur N, Marquine M, Heyse D, Raymond M (1995) Population structure and dynamics of selected genes in the mosquito *Culex pipiens*. *Evolution*, **49**, 997–1007.
- Cockerham CC, Weir BS (1993) Estimation of gene flow from *F*-statistics. *Evolution*, **47**, 855–863.
- Corn PS (1981) Field evidence for a relationship between color and developmental rate in the northern leopard frog (*Rana pipiens*). *Herpetologica*, **37**, 155–160.
- Corn PS (1982) *Selection pressures affecting a dorsal color polymorphism in Rana pipiens*. PhD Thesis. Colorado State University, Fort Collins, Colorado.
- Endler JA (1986) *Natural Selection in the Wild*. Princeton University Press, Princeton, New Jersey.
- Fincke OM (1994) Female colour polymorphism in damselflies: failure to reject the null hypothesis. *Animal Behavior*, **47**, 1249–1266.
- Fogleman JC, Corn PS, Pettus D (1980) The genetic basis of a dorsal color polymorphism in *Rana pipiens*. *Journal of Heredity*, **71**, 439–440.
- Forsman A, Shine R (1995) The adaptive significance of colour pattern polymorphism in the Australian scincid lizard *Lampropholis delicata*. *Biological Journal of the Linnean Society*, **55**, 273–291.
- Futuyma DJ (1986) *Evolutionary Biology*, 2nd edn. Sinaur Associates, Inc., Sunderland, Massachusetts.
- Gilbert M, Leclair R, Fortin R (1994) Reproduction of the northern leopard frog (*Rana pipiens*) in floodplain habitat in the Richelieu River, P. Quebec, Canada. *Journal of Herpetology*, **28**, 465–470.
- Gillespie RG, Oxford GS (1998) Selection on the color polymorphism in Hawaiian happy-face spiders: evidence from genetic structure and temporal fluctuations. *Evolution*, **52**, 775–783.
- Hedrick PW (2000) *Genetics of Populations*. Jones and Bartlett Publishers, Sudbury, Massachusetts.
- Heinen JT, Hammond G (1997) Antipredator behaviors of newly metamorphosed green frogs (*Rana clamitans*) and leopard frogs (*R. pipiens*) in encounters with eastern garter snakes (*Thamnophis s. sirtalis*). *American Midland Naturalist*, **137**, 136–144.
- Hoffman EA, Ardren WR, Blouin MS (2003) Nine polymorphic microsatellite loci in the northern leopard frog (*Rana pipiens*). *Molecular Ecology Notes*, **3**, 115–116.
- Hoffman EA, Blouin MS (2000) A review of colour and pattern polymorphisms in anurans. *Biological Journal of the Linnean Society*, **70**, 633–663.
- Hoffman EA, Blouin MS (2004a) Evolutionary history of the northern leopard frog: Reconstruction of phylogeny, phylogeography, and historical changes in population demography from mitochondrial DNA. *Evolution*, **58**, 145–159.
- Hoffman EA, Blouin MS (2004b) Historical data refute recent range contraction as cause of low genetic diversity in isolated frog populations. *Molecular Ecology*, **13**, 271–276.
- Hoffman EA, Schueler FW, Blouin MS (2004) Effective population sizes and temporal stability of genetic structure in *Rana pipiens*, the northern leopard frog. *Evolution*, **58**, 2536–2545.
- Karl SA, Avise JC (1992) Balancing selection at allozyme loci in oysters: implications from nuclear RFLPs. *Science*, **256**, 100–102.
- King RB, Lawson R (1995) Color-pattern variation in Lake Erie water snakes: the role of gene flow. *Evolution*, **49**, 885–896.
- Kingsolver JG, Hoekstra HE, Berrigan JM *et al.* (2001) The strength of phenotypic selection in natural populations. *American Naturalist*, **157**, 245–261.
- Long AD, Singh RS (1995) Molecules versus morphology: the detection of selection acting on morphological characters along a cline in *Drosophila melanogaster*. *Heredity*, **74**, 569–581.
- Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology*, **3**, 91–99.
- Lynch M, Pfrender M, Spitze K *et al.* (1999) The quantitative and molecular genetic architecture of a subdivided species. *Evolution*, **53**, 100–110.
- Mithon R, Raybould AF, Giamoustaris A (1995) Divergent selection for secondary metabolites between wild populations of *Brassica oleracea* and its implications for plant–herbivore interactions. *Heredity*, **75**, 472–484.
- Monsen KM, Blouin MS (2003) Genetic structure in a montane ranid frog: restricted gene flow and nuclear-mitochondrial discordance. *Molecular Ecology*, **12**, 3275–3286.
- Osawa N, Nishida T (1992) Seasonal variation in elytral colour polymorphism in *Harmonia axyridis* (the ladybird beetle): the role of non-random mating. *Heredity*, **69**, 297–307.
- Prout T, Barker JSF (1993) *F* statistics in *Drosophila buzzatii*: selection, populations size and inbreeding. *Genetics*, **134**, 369–375.
- Raymond M, Rousset F (1995) GENEPOP: population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.
- Rice WR (1989) Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.
- Ryan RA (1953) Growth rates of some ranids under natural conditions. *Copeia*, **1953**, 73–80.
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Sandoval CP (1994) Differential visual predation on morphs of *Timema cristinae* (Phasmatodeae: Timemidae) and its consequences for host range. *Biological Journal of the Linnean Society*, **52**, 341–356.

- Schueler FW (1979) *Geographic variation in skin pigmentation and dermal glands in the Northern Leopard Frog, Rana pipiens*. PhD Thesis. University of Toronto, Ontario, Canada.
- Schueler FW (1981) Preserving anuran skins by drying. *Herpetological Review*, **12**, 10–11.
- Schueler FW (1982) Geographic variation in skin pigmentation and dermal glands in the northern leopard frog, *Rana pipiens*. *Publications in Zoology, National Museums of Canada, National Museum of Natural Sciences*, **16**, 1–80.
- Shine R, Ambariyanto, Harlow PS (1998) Mumpuni Ecological divergence among sympatric colour morphs in blood pythons, *Python brongersmai*. *Oecologia*, **116**, 113–119.
- Spitze K (1993) Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics*, **135**, 367–374.
- Stimson J, Berman M (1990) Predator induced colour polymorphism in *Danaus plexippus* L. (Lepidoptera: Nymphalidae) in Hawaii. *Heredity*, **65**, 401–406.
- Storz JF, Nachman MW (2003) Natural selection on protein polymorphism in the rodent genus *Peromyscus*: evidence from ontelocus contrasts. *Evolution*, **57**, 2628–2635.
- Taylor MFJ, Shen Y, Kreitman ME (1995) A population genetic test of selection at the molecular level. *Science*, **270**, 1497–1499.
- Vitalis R, Dawson K, Boursot P (2001) Interpretation of variation across marker loci as evidence of selection. *Genetics*, **158**, 1811–1823.
- Wang J (2001) A pseudo-likelihood method for estimating effective population size from temporally spaced samples. *Genetical Research*, **78**, 243–257.
- Waples RS (1989) A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics*, **121**, 379–391.
- Whitlock MC, McCauley DE (1999) Indirect measures of gene flow and migration: F_{ST} doesn't equal $1/(4Nm + 1)$. *Heredity*, **82**, 117–125.
- Wright S (1984) *Evolution and the Genetics of Populations. Volume 2: The Theory of Gene Frequencies*. The University of Chicago Press, Chicago.
- Yang R-C, Yeh FC, Yanchuk AD (1996) A comparison of isozyme and quantitative genetic variation in *Pinus contorta* ssp. *Latifolia* by F_{ST} . *Genetics*, **142**, 1045–1052.
- Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics*, **20**, 176–183.

E. Hoffman's laboratory focuses on investigating the molecular ecology, population genetics, and evolution of many taxa. F. Schueler is a general naturalist whose herpetological work has centered on the geographic variation, distribution, variation, and life histories of common species, especially Leopard Frogs. A. Jones uses molecular markers to study many aspects of behavior and evolutionary biology. M. Blouin's laboratory focuses on the causes and consequences of genetic structuring, and on applications of methods for pedigree reconstruction, in natural populations. He works on a variety of taxa including fish, amphibians and parasites.
