

INFERRING POPULATION HISTORY AND DEMOGRAPHY USING MICROSATELLITES, MITOCHONDRIAL DNA, AND MAJOR HISTOCOMPATIBILITY COMPLEX (MHC) GENES

David H. Bos,^{1,2} David Gopurenko,¹ Rod N. Williams,¹ and J. Andrew DeWoody¹

¹Department of Forestry and Natural Resources, Purdue University, West Lafayette, Indiana 47906

²E-mail: dbos@purdue.edu

Received August 15, 2007

Accepted February 11, 2008

Microsatellites and mitochondrial DNA (mtDNA) have traditionally been used in population genetics because of their variability and presumed neutrality, whereas genes of the major histocompatibility complex (MHC) are increasingly of interest because strong selective pressures shape their standing variation. Despite the potential for MHC genes, microsatellites, and mtDNA sequences to complement one another in deciphering population history and demography, the three are rarely used in tandem. Here we report on MHC, microsatellite, and mtDNA variability in a single large population of the eastern tiger salamander (*Ambystoma tigrinum tigrinum*). We use the mtDNA mismatch distribution and, on microsatellite data, the imbalance index and bottleneck tests to infer aspects of population history and demography. Haplotype and allelic variation was high at all loci surveyed, and heterozygosity was high at the nuclear loci. We find concordance among neutral molecular markers that suggests our study population originated from post-Pleistocene expansions of multiple, fragmented sources that shared few migrants. Differences in N_e estimates derived from haploid and diploid genetic markers are potentially attributable to secondary contact among source populations that experienced rapid mtDNA divergence and comparatively low levels of nuclear DNA divergence. We find strong evidence of natural selection acting on MHC genes and estimate long-term effective population sizes (N_e) that are very large, making small selection intensities significant evolutionary forces in this population.

KEY WORDS: *Ambystoma*, effective population size, mismatch distribution, M -ratio, natural selection, neutral evolution.

Demographic factors play a key role in shaping the patterns of neutral genetic variation of a population. These patterns of variation can be interpreted in a well-established theoretical context of molecular evolution applied to various demographic models (Charlesworth et al. 2003). Consequently, population genetic studies have generally relied upon mitochondrial DNA (mtDNA) or microsatellites because both marker types are polymorphic and they often do not deviate significantly from expectations of neutral evolution (Charlesworth et al. 1994; although see Ballard and Whitlock 2004).

Increasingly, nonneutral markers are also being evaluated by empirical population geneticists, often in conjunction with neutral

genetic markers (e.g., Boyce et al. 1997; Aguilar et al. 2004; Campos et al. 2006). Of particular note are major histocompatibility complex (MHC) genes, which function in the immune system and present short peptides to T cells, thereby distinguishing between foreign and native protein products. Because of their immunological (and possibly other) functions, nucleotide variation at MHC genes may be directly tied to individual fitness and survivorship (Hill et al. 1991; Carrington and O'Brien 2003). In other words, much of the molecular variation among alleles is adaptive and is maintained by natural and/or sexual selection (Potts et al. 1994).

Natural selection can lead to rapid changes in patterns of genetic variation in populations. Changes in populations driven

by natural selection often occur in an episodic manner, making selection a potentially strong but often transient force (Elena et al. 1996). As such, loci may experience long periods of near-neutral evolution dominated by genetic drift, but also experience recurring episodes of evolution driven by natural selection. The effects of natural selection are also partially dependent on the long-term effective size (N_e) of a population, which determines the relative strength of a selection coefficient (Satta et al. 1994). Kimura and Ohta (1969) showed that selection is a potent evolutionary force only when $N_e(s) \geq 1$, where s is the selection coefficient. Thus, the evolutionary response (e.g., changes in allele frequencies or numbers of alleles) to an advantageous allele is more effective in a population with large N_e than a population with small N_e .

Because the response to selection is based on N_e , and because N_e is governed in part by population parameters (such as the operational sex ratio or temporal variance in size), demography can confound the inference of natural selection (Williamson et al. 2005). Therefore, studies of natural selection must account for demography by comparing different types of mutations (e.g., nonsynonymous vs. synonymous substitutions) or different types of loci (Nielsen 2001). Regarding the latter, one approach is to compare adaptive MHC and neutral microsatellite (or mtDNA) variation. Such comparisons rely on the genomewide effects of drift or gene flow as opposed to the locus-specific effects of selection (Hedrick et al. 2001). For example, similarities in MHC and microsatellite variability may indicate that selection has been weak or inconsistent, but increased levels of MHC variability relative to microsatellites is often taken as evidence of selection (e.g., Landry and Bernatchez 2001; Cohen 2002; Wegner et al. 2003).

Very few studies have simultaneously gauged the relative levels of genetic variation at microsatellites, mtDNA, and MHC loci despite their ability to complement one another in deciphering the recent evolutionary history of a population. Here, we provide a comparative analysis of these three classes of genetic marker sampled from a single breeding population of the eastern tiger salamander (*Ambystoma tigrinum tigrinum*). Eastern tiger salamanders are obligate pond breeding amphibians that are not known to hybridize with coexisting species. Although *A. t. tigrinum* have limited capabilities for dispersal, the species has a wide range, and the generation time varies depending on the location of the population; however, the population under study is thought to have a generation time of two years. In parts of the species' range previously disturbed by Pleistocene glacial maxima, population fragmentation likely resulted from reduced gene flow and population sizes. If so, the impact of these historical events should be manifest in extant genetic lineages. Population genetic and phylogeographic studies have extensively sampled from the species' south, eastern, and western range limits (Routman 1993; Templeton et al. 1995; Church et al. 2003), often in areas unaffected by Pleistocene

glaciation. However, the north and central parts of the range remain genetically uncharted.

In this study, we sample a population in the central part of the range, near the putative limits of Wisconsin glaciation to infer the origin, history, and demography of this population. Specifically, we test (1) whether it originated after the retreat of Pleistocene ice sheets, (2) if the population is currently or has been historically fragmented, (3) whether substantial changes in population size occurred, and (4) if the signature of selection is constrained by past population dynamics. The use of multiple types of molecular markers is essential to distinguish between similar patterns that could result from locus-specific events (such as a selective sweep) and genomewide events (such as bottlenecks) that simultaneously affect all loci, although to potentially different degrees. We predict that if this population has been influenced by glacial advance and retreat, a pattern of genetic variation consistent with a range expansion will emerge from all neutral markers. This pattern should be accompanied by signatures of demographic changes depending on migration rates and the level of recovery and severity of potential bottlenecks. In contrast, if the data indicate a population at mutation–drift equilibrium, then the population is unlikely to have originated recently and may represent a stable refugium unaffected by Pleistocene glaciations. By integrating data from multiple molecular markers, we provide a detailed description of the history and demography of a population, illustrating how microevolution shapes population genetic diversity.

Materials and Methods

DNA ISOLATION

To assess gene diversity, tissue samples (consisting of tail or toe clips) from 100 *A. t. tigrinum* were collected from a single site in a large wetland complex at the Purdue Wildlife Area in central Indiana (USA). Salamanders were captured using drift fences and pitfall traps, then uniquely marked via toe-clipping. Tissue samples were digested using a standard lysis procedure (extraction buffer: 20 mM Tris-Cl, 400 mM NaCl, 5 mM EDTA, 1% SDS and 5 μ l proteinase K (10 mg/mL)). Phenol–chloroform extractions were performed on the cell lysate followed by isopropanol precipitation and a 70% ethanol wash (Sambrook and Russell 2001). Genomic DNA was resuspended in 100 μ l TLE buffer (10 mM Tris-Cl, 0.1 mM EDTA).

MOLECULAR METHODS

To assay mtDNA variation in *A. t. tigrinum*, we sequenced a portion of the NADH dehydrogenase subunit 2 mitochondrial gene (*ND2*). This gene was chosen because substitution rates are elevated at *ND2* relative to other regions of the *Ambystoma* mtDNA genome (Samuels et al. 2005) and because our pilot work revealed intraspecific variation at *ND2* but none in the D-loop. PCR

with *ND2* primers (*ND2-F*, 5'-GCGAGCAACAGAAGCAGCTA-3' and *ND2-R*, 5'-GCATAAATCCGGTTGTTGGT-3') yielded a 635 base pair (bp) amplicon. Final PCR conditions included 0.4 μ M of each primer, \sim 50 ng of total DNA, 1.25 mM dNTPs, 2.5 mM MgCl₂, 1 \times PCR buffer (10 mM Tris-Cl, 50 mM KCl, 0.5 μ g bovine serum albumin), and 0.2 units of *Taq* polymerase (New England Biolabs, Ipswich, MA) in a final volume of 10 μ L. The thermal profile consisted of 60 sec at 94°C followed by 33 cycles of 15 sec at 94°C, 15 sec at 59°C, and 15 sec at 72°C, concluding with a 5-min extension at 72°C. Ethanol precipitated PCR products were rehydrated and bidirectionally sequenced using BigDye[®] chemistry (Applied Biosystems, Foster City, CA).

Six polymorphic microsatellites were genotyped as described in Gopurenko et al. (2006), Parra-Olea et al. (2007), and Williams and DeWoody (unpubl. data). These methods consist of multiplex PCR amplification using forward primers labeled with fluorescent dye. Amplicons from each sample were scored for size via electrophoresis on an ABI 377 or AB3730XL sequencer (Applied Biosystems) using associated GeneScan, Genotyper, and GeneMapper software.

The lone *A. t. tigrinum* MHC class II β gene (*Amti-DAB*) is expressed, has classical function, and has been the subject of natural selection (Bos and DeWoody 2005). A 264-bp fragment of this gene containing the peptide-binding region (PBR) was PCR amplified from whole genomic DNA as in Bos and DeWoody (2005). PCR products were cleaned using spin columns and then sequenced directly using BigDye chemistry and an Applied Biosystems 3730XL automated sequencer. Only high-quality sequences were retained for analysis, and every individual was sequenced bidirectionally. In Sequencher ver. 4.12 (GeneCodes, Ann Arbor, MI), we aligned sequences from each individual and assigned genotypes to homozygotes and single-site heterozygotes. In multiple-site heterozygotes, unknown phase orientations were resolved using PHASE, a methodology shown to be accurate for reconstructing MHC haplotypes (Stephens et al. 2001; Stephens and Donnelly 2003; Bos et al. 2007).

STATISTICAL ANALYSES

Several statistics were calculated to summarize genetic diversity at each locus. mtDNA sequences were aligned using BioEdit ver. 7.0.1 (Hall 1999) and sorted using GENALEX ver. 6 (Peakall and Smouse 2006) to identify individual haplotypes and their frequencies. Nucleotide content among haplotypes was examined using MEGA ver. 3.1 (Kumar et al. 2001). Summary statistics included measures of haplotype diversity ($\hat{H} \pm \text{SD}$; Nei 1987) and nucleotide diversity (average over all sites: $\pi_n \pm \text{SD}$; Tajima 1983) which were estimated using ARLEQUIN ver. 3.01 (Schneider et al. 2000). For the microsatellite loci, we calculated allele frequencies, observed heterozygosity, and tested for genotypic disequilibrium and deviation from Hardy–Weinberg expectations

using GENEPOP ver. 3.4 (Raymond and Rousset 1995). Upon establishment of *Amti-DAB* genotypes, we calculated allele frequencies and characterized polymorphic sites. Allelic diversity was gauged as mean nucleotide diversity and mean pairwise amino acid differences using dnaSP ver. 4.0 (Rozas et al. 2003).

We investigated aspects of population origin, history and demography using both mtDNA and microsatellite data. We used the mismatch distribution (Rogers and Harpending 1992) among *ND2* sequences to evaluate various models of population growth, stability, and spatial distribution. We compared observed distributions with the strong positive skewed distribution expected under mutation–drift equilibrium (Watterson 1975), the unimodal wave distribution expected for sudden population growth (Rogers and Harpending 1992), and the bimodal distribution expected for a spatial expansion with low migration rates (Ray et al. 2003) as implemented in ARLEQUIN ver. 3.01 (Schneider et al. 2000). We note, however that there is inherent variation in the expected distributions under different demographic scenarios. Occasionally these distributions overlap, and the analysis of a single population whose mismatch distribution is influenced by the natural stochasticity of evolution should be interpreted cautiously. To quantify the correspondence of the observed distribution to different demographic scenarios, the expected distribution is numerically calculated and parameters for the observed mismatch distribution (e. g., for the spatial expansion: $\tau = 2T\mu$, $M = 2Nm$, and $\theta = 2N\mu$) are estimated by minimizing the sum of squared deviations (SSD) (Schneider and Excoffier 1999; Excoffier 2004). Significance was assessed by comparing the observed SSD with a distribution of values drawn from data simulated using parameters optimized from the observed data under each demographic scenario. A metric analogous to the mismatch distribution but comparing the distribution of length differences among microsatellite alleles was also used to detect demographic expansion in this population (Kimmel et al. 1998; King et al. 2000). We calculated the variance estimator (θ_V) and the homozygosity estimator (θ_H) from equations (1) and (3) of Kimmel et al. (1998). The difference in the natural logarithm of these estimators, averaged over all microsatellite loci, is the imbalance index ($\ln\beta$). This statistic relies on an imbalance between the allele size variance and heterozygosity to detect population growth. The imbalance index should be < 1 for expanding populations whose preexpansion history is stable, but is characteristically > 1 in populations with a reduction in size that precedes a detectable growth phase. The $\ln\beta$ metric was chosen because of its power and ability to detect historic signals of population expansion (King et al. 2000).

Tajima's *D* (Tajima 1989) and Fu's *F_S* (Fu 1997) were calculated from mtDNA haplotypes and used to test for deviations from the neutral equilibrium condition and to refine hypotheses from the mismatch distribution. We estimated Tajima's *D* as a test of selective neutrality and *F_S* to provide an indication of population

growth or hitchhiking. Significance was assessed by comparing observed D and F_S values to the distribution of 10,000 neutral coalescent simulations. *Ambystoma tigrinum tigrinum* mtDNA haplotypes were also connected as a network cladogram using TCS ver. 1.21, which estimates the most parsimonious pathway of descent among haplotypes based upon shared mutations and without assuming a strictly bifurcating process (Clement et al. 2000). We used the microsatellite data to calculate F_A , the homozygosity index used as the basis for the Ewens–Watterson test (Watterson 1977). The Ewens–Watterson test relies on testing the observed allele frequencies compared to those expected under mutation–drift equilibrium, and is therefore different from a test of Hardy–Weinberg proportions. This test is useful for detecting recent or current deviations from mutation–drift equilibrium due to selection or demographic changes.

The population under study may have undergone past changes in demographics, such as a bottleneck, for which the genetic signature is still manifest. Two different microsatellite-based approaches were used to test for signatures of recent population bottlenecks. First, M -ratios were used to test for bottlenecks using a two-phase model with 90% single-step mutations and an average $\Delta_g = 3.5$ (Garza and Williamson 2001). The M -ratio test relies on allele frequencies and gaps between existing allelic states that serve as a signal of past reductions in population size. M_{crit} values represent significant cutoffs below which a bottleneck is inferred. M_{crit} is sensitive to underlying model parameters, therefore we chose two likely θ values of 10 and 100 based to represent a degree of model uncertainty. Second, we used the program BOTTLENECK ver. 1.2.02 (Cornuet and Luikart 1996), again using a two-phase model with 90% single-step mutations. BOTTLENECK identifies very recent reductions in effective size by comparing heterozygosity from observed data to heterozygosity in a simulated population at neutral mutation–drift equilibrium. A Wilcoxon signed-rank test was used to determine if a statistically significant number of loci displayed a heterozygote excess compared to expectations based on the observed number of alleles.

The long-term effective number of females (N_{ef}) in the population was calculated from mtDNA data using two different estimates of the female-specific theta ($\theta_f = 2N_{ef}\mu$). For both estimates, a range of $ND2$ substitution rates (1.6×10^{-8} to 3.7×10^{-9} substitutions/site/year) was used to reflect uncertainty in this parameter (Mueller 2006). N_{ef} was calculated from the range of substitution rates and the most probable estimate of θ_f derived from LAMARC (Kuhner et al. 2005) using Bayesian inference with 10 initial search chains of 10,000 iterations each. For final estimates of θ , we performed two final chains of 1,000,000 iterations. The second N_{ef} estimate used θ_f derived from the distribution of pairwise nucleotide differences (i.e., the mismatch distribution; Rogers and Harpending [1992]). N_{ef} is expected to be roughly $1/4$ that of N_e derived from genomic DNA because of

the different modes of inheritance for nuclear and mtDNA, as well as the haploid nature of the mitochondrial genome (Ballard and Whitlock 2004).

Long-term effective population size was directly estimated based on microsatellite heterozygosity using Nei's (1987) formula

$$N_e = (1/[1 - H_E]^2 - 1)/8\mu$$

and assuming mutation rates (μ) from 2.5×10^{-4} to 1.0×10^{-5} mutations/locus/generation. The population parameter theta ($\theta = 4N_e\mu$) was also used to estimate effective size based on all six microsatellite loci. RoyChoudhury and Stephens (2007) have shown that the homozygosity estimator (θ_H) of Kimmel et al. (1998) is an accurate and unbiased estimator of the scaled population mutation rate. Therefore, we use this estimator of θ to calculate the long-term effective population size using the above mutation rates.

To test for signals of historic natural selection in mtDNA and MHC sequences, we calculated the ratio of nonsynonymous (d_N) to synonymous (d_S) substitution rates (d_N/d_S) for each marker, with values greater than 1 serving as an indicator of positive natural selection. We also calculate the selection intensity (S), using the formula

$$S = N^2\omega/\sqrt{2}$$

from Satta et al. (1994, eqn. 6), where N = the mean number of nonsynonymous substitutions, and $\omega = d_N/d_S$. Satta et al. (1994) suggest multiple methods of approximating N . We use methods I and IV, which approximate N using the observed number of nonsynonymous substitutions in the sample, and the number of alleles in the sample, respectively (see Satta et al. [1994] for logic and underlying assumptions regarding use of each of these methods).

The Ewens–Watterson homozygosity statistic was calculated on MHC sequences to compare the empirical allele frequency distribution to those expected under mutation–drift equilibrium. This test measures the intensity of natural selection in recent time scales as opposed to the historic time scales evaluated by nucleotide substitution patterns (Garrigan and Hedrick 2003).

Results

GENETIC VARIATION

A total of nine mtDNA haplotypes ($ND2$ sequences) were identified among the sample of 100 adult *A. t. tigrinum*. There was no evidence of pseudogenes among sample sequences, and a total of 16 polymorphic sites were observed among the nine haplotypes (Table 1). Haplotype frequencies varied from 0.01 to 0.24 (Fig. 1) and overall haplotype diversity was 0.84 ± 0.01 . Pairwise sequence divergence ranged from 0.17 to 1.71% (mean = 0.93%) and overall nucleotide diversity was 0.009 ± 0.005 . Thirteen segregating sites were silent with substitutions occurring at third

Table 1. Variable mtDNA sites among nine ND2 haplotypes (Hap 1–9) sampled from 100 *A. t. tigrinum*. *N* = number of each haplotype in sample. Δ = the type of substitution in which *s* = transition and *v* = transversion; boldface represents replacement substitution. Codon refers to first, second, or third-position substitutions (1, 2, 3).

| Gene: | ND2 | | | | | | | | | | | | | | | <i>N</i> | GenBank Acc. # | |
|----------|---------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------------|----------|
| | Nucleotide position | | | | | | | | | | | | | | | | | |
| | 0 | 0 | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 4 | 5 | | |
| | 0 | 1 | 9 | 3 | 5 | 8 | 6 | 6 | 7 | 1 | 1 | 1 | 3 | 8 | 2 | 3 | | |
| | 3 | 8 | 7 | 2 | 6 | 0 | 2 | 6 | 3 | 2 | 5 | 8 | 0 | 7 | 0 | 7 | | |
| Hap 8 | T | G | C | G | A | C | C | A | A | C | C | T | G | C | C | T | 15 | EF571304 |
| Hap 7 | . | . | . | . | . | . | . | . | . | . | . | . | A | . | . | . | 16 | EF571305 |
| Hap 5 | . | . | A | . | . | . | . | . | . | . | . | . | . | . | . | . | 11 | EF571306 |
| Hap 6 | . | . | . | . | . | . | . | . | . | . | . | C | . | . | . | . | 5 | EF571307 |
| Hap 9 | . | . | . | . | . | T | . | . | . | . | . | . | . | . | . | . | 2 | EF571308 |
| Hap 1 | C | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | 2 | EF571309 |
| Hap 2 | . | A | . | A | . | T | T | C | . | . | T | . | . | T | T | C | 19 | EF571310 |
| Hap 4 | . | A | . | . | G | . | . | C | G | . | T | . | . | T | T | . | 24 | EF571311 |
| Hap 3 | . | A | . | . | G | . | . | C | G | A | T | . | . | T | T | . | 6 | EF571312 |
| Δ | <i>s</i> | <i>s</i> | <i>v</i> | <i>s</i> | <i>s</i> | <i>s</i> | <i>s</i> | <i>v</i> | <i>s</i> | <i>v</i> | <i>s</i> | <i>s</i> | <i>s</i> | <i>s</i> | <i>s</i> | <i>s</i> | | |
| codon | 3 | 3 | 1 | 3 | 3 | 3 | 1 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | 3 | | |

codon positions; the remaining three sites harbored replacement substitutions at first and second codon positions (position 97 (H ↔ N); position 262 (H ↔ Y); and position 266 (Q ↔ P)).

The six microsatellite loci were highly polymorphic with a mean of 17.5 alleles per locus (range: 4–29) and mean heterozygosity of 0.75. There was no evidence of genotypic disequilibrium among pairs of loci and a multilocus test of deviation from Hardy–Weinberg equilibrium was not significant (*P* = 0.19; Table 2); tests based on each individual locus indicate that they are all in Hardy–Weinberg equilibrium.

At the *Amti-DAB* locus, no indels were detected and there was no evidence for pseudogene amplification (i.e., stop codons, frameshifts, or mutations in conserved residues essential for classical MHC class II protein function; Bos and DeWoody [2005]). A total of 40 polymorphic nucleotide sites were identified among

eight distinct alleles (Table 3). The most common allele had a frequency of 0.31 and the rarest allele had 0.03 (Fig. 1). Pairwise amino acid differences among alleles varied from 0.4 to 13.5% (mean = 8.1%). Substitutions occurred at all codon positions, but predominately the first position (18 of 40 sites). Excluding four nucleotide sites with a complex pattern of substitution, three segregating sites led to synonymous (silent) substitutions whereas 33 segregating sites were nonsynonymous.

POPULATION ORIGIN AND HISTORY

The *ND2* mismatch distribution was bimodal and somewhat ragged (Fig. 2). These data were statistically inconsistent with the neutral model, and also with the model of sudden increase in

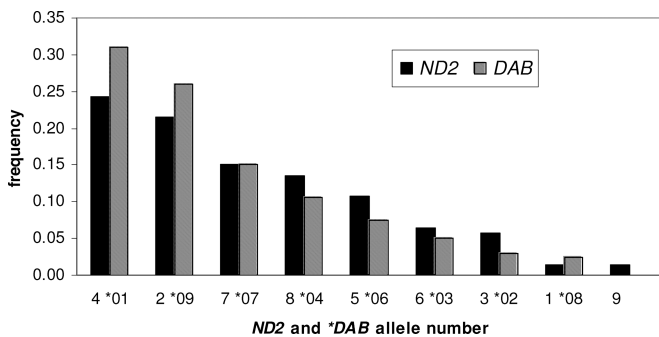


Figure 1. Observed allele frequencies in our sample of MHC (*Amti-DAB*) and mtDNA (*ND2*) genes. The *DAB* gene has eight alleles and the *ND2* gene has nine alleles (haplotypes).

Table 2. Microsatellite variation at six loci (*AT55-7*, *Atex65*, 52.1, 52.34, 52.6, and 60.3) among 100 *A. t. tigrinum*. *K* is the number of observed alleles; inbreeding coefficient (*F_{is}*) as per Weir and Cockerham (1984); *H_{obs}* and *H_{exp}* refer to observed and expected heterozygosity (Nei 1978), and *P* (*HW*) is the probability of deviation from Hardy–Weinberg equilibrium. *EW_{exp}* and *EW_{obs}* are the observed and expected Ewens–Watterson values, respectively.

| Locus: | <i>AT55-7</i> | <i>Atex65</i> | 52.1 | 52.34 | 52.6 | 60.3 |
|-------------------------|---------------|---------------|--------|--------|--------|--------|
| <i>K</i> | 3 | 25 | 10 | 20 | 2 | 6 |
| <i>F_{is}</i> | + 0.065 | +0.019 | −0.062 | +0.010 | +0.133 | +0.010 |
| <i>H_{obs}</i> | 0.535 | 0.931 | 0.832 | 0.871 | 0.436 | 0.871 |
| <i>H_{exp}</i> | 0.571 | 0.951 | 0.783 | 0.879 | 0.507 | 0.879 |
| <i>P</i> (<i>HW</i>) | 0.511 | 0.237 | 0.159 | 0.761 | 0.235 | 0.761 |
| <i>EW_{obs}</i> | 0.431 | 0.053 | 0.219 | 0.125 | 0.500 | 0.531 |
| <i>EW_{exp}</i> | 0.705 | 0.120 | 0.314 | 0.156 | 0.831 | 0.471 |

Table 3. Variable nucleotide sites among eight identified haplotypes (DAB01–DAB09) at *Amti-DAB*. Sites identical to haplotype 8 at a nucleotide position represented as a dot, changes as per indicated base difference (A, C, G, and T). *N* = number of each haplotype in sample. Δ = class of mutational change, where *s* = transition and *v* = transversion, and \diamond = complex nucleotide site; boldface represents amino acid replacement substitution. Codon (1,2,3) refers to amino acid coding position.

| Gene: | Amti-DAB | | | | | | | | | | | | | | | | | | | | | | | | N | GenBank Acc. # | | | | | | |
|-----------|----------|------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------------|----------|----------|----------|----------|----------|----------|
| Position: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | | | | | | | | |
| DAB01 | G | G | C | A | G | T | G | A | T | C | A | T | C | A | G | A | G | C | G | C | C | G | A | G | C | C | T | A | A | 61 | DQ071905 | |
| DAB02 | C | T | G | G | . | . | G | . | G | . | C | . | . | . | A | G | . | . | . | . | . | A | G | G | T | T | . | . | 6 | DQ071906 | | |
| DAB03 | C | A | G | G | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | 10 | DQ071907 | | |
| DAB04 | C | A | G | G | . | . | A | T | . | A | G | G | . | . | A | G | . | . | . | . | A | G | G | T | T | . | . | 21 | DQ071908 | | | |
| DAB06 | C | . | . | A | C | T | A | A | . | A | . | . | . | . | A | T | . | . | A | G | T | G | T | . | C | G | T | . | 15 | DQ071910 | | |
| DAB07 | C | . | . | C | T | A | T | . | . | C | . | C | A | C | T | C | T | . | . | A | A | . | . | G | C | A | . | G | T | T | 30 | DQ071911 |
| DAB08 | . | . | . | . | A | . | . | . | . | C | . | C | T | C | T | A | C | . | A | . | T | . | . | . | A | . | G | T | T | 5 | DQ071912 | |
| DAB09 | C | A | G | G | . | . | A | T | . | A | G | G | . | . | G | . | . | A | G | . | . | A | . | . | A | G | G | T | T | 52 | DQ071913 | |
| Δ | <i>v</i> | \diamond | <i>v</i> | <i>s</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | <i>v</i> | | | | |
| codon | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 1 | 2 | 3 | 2 | 1 | 2 | 1 | 2 | 2 | 1 | 3 | 1 | 2 | 2 | 2 | 3 | 2 | 2 | | | |

a panmictic population (both rejected, $P < 0.05$). On the other hand, the observed SSD of the mtDNA mismatch distribution is within the distribution of values obtained under a spatial expansion model with multiple ancestral sources sharing few migrants ($P = 0.19$; optimized parameters: $\tau = 7.37$, $M = 2.74$, and $\theta = 2.11$). Tajima’s D is consistent with a lack of selection or other deviations from mutation–drift equilibrium on $ND2$ ($D = 2.07$, $P = 0.98$). From our microsatellite dataset, we calculate $\ln\theta_H$ and $\ln\theta_V$ of 1.59 and 6.21, respectively, resulting in an imbalance index of 4.62, indicating that the population may have experienced a growth phase following an earlier reduction. F_{IS} values were slightly positive at 5 of the 6 microsatellite loci, with the lone exception of locus 52.1 (Table 2). In all cases, F_{IS} values were not significantly different from zero. Finally, analysis of Fu’s F_S ($F_S = 5.36$, $P = 0.95$) reveal no evidence of excess rare alleles at the *Amti-DAB* locus that could be caused by sudden population growth.

There was no evidence of previous population bottlenecks as indicated by M -ratio or BOTTLENECK analyses. The mean multilocus M -ratio was 0.71 and the M_{crit} values (which indicate the point below which M scores indicate a historical bottleneck) were 0.68 and 0.71 for different values of θ . In the BOTTLENECK analysis, three loci (52.34, 60.3, and 52.1) displayed a heterozygote deficiency whereas the other three loci exhibited a heterozygote excess. Under the two-phase model and equilibrium, each individual locus has a roughly equal chance of having a heterozygote deficiency or excess. Accordingly, the Wilcoxon signed-rank test provided no evidence for the heterozygote excess expected from a recently bottlenecked population ($P = 0.66$). The tripartite “star” network of $ND2$ sequences suggests this population has not recently experienced any severe or prolonged bottlenecks (Fig. 3).

EFFECTIVE POPULATION SIZE

The most probable estimate of θ_f for $ND2$ data was 0.006 (95% CI 0.005–0.008) using Bayesian estimates of θ_f , and 0.003 using θ_f estimated from the best-fit model for the mismatch distribution. For the former, the N_{ef} values obtained using the most probable estimate of θ_f range from 1.9×10^5 to 7.8×10^5 , depending on the mutation rate; for the latter, N_{ef} estimates range from 9.5×10^4 to 4.1×10^5 (Fig. 4). Using the microsatellite data, θ was estimated as 4.92 using the estimator of Kimmel et al. (1998). N_e based on this estimator of θ and the above mutation rates ranged from 4.92×10^3 to 1.23×10^5 . In comparison, heterozygosity estimates of N_e ranged from 7.7×10^3 to 1.9×10^5 (Fig. 4).

NATURAL SELECTION

The observed d_N/d_S for the *Amti-DAB* gene was 3.12. Using the same techniques, the mtDNA sequence data have a d_N/d_S value of 0.067. The selection intensity (S) on *Amti-DAB* was estimated as 2,401.75 using method I (patterns of substitution) and 141.15

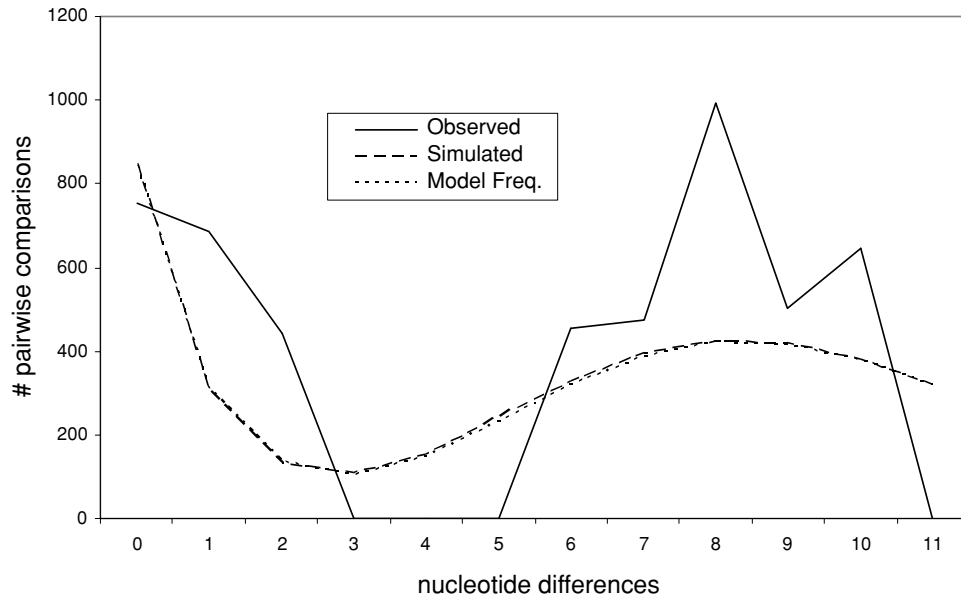


Figure 2. Observed distribution of pairwise differences (the mismatch distribution) of *ND2* haplotypes between individuals. The observed distribution is consistent with this spatial expansion model ($P = 0.19$). The expected distributions calculated both numerically (Model Frequency expected distribution line) and with simulated data (Simulated expected distribution line) using optimized model parameters ($\tau = 7.37$, $M = 2.74$, and $\theta = 2.11$) from observed data are also shown. Expected distributions are calculated using the simplifying assumptions of a continent-island model with low Nm (Excoffier 2004).

using method IV (allele frequencies). The intensities of selection for the *ND2* sequence data were $S = 0.43$ and 3.05 for methods I and IV, respectively. The observed F_A (Watterson's homozygosity statistic) for the MHC data was 0.23 , significantly less than the expected value of 0.38 , suggesting an allele frequency distribu-

tion that is more uniform than expected under neutrality (Ewens 1972).

Discussion

POPULATION ORIGIN AND HISTORY

Numerous studies have compared MHC and microsatellite data across populations, but this study is among the first to simultaneously use genetic variation at microsatellites, mtDNA, and MHC genes to evaluate a population's evolutionary history. We use a large sample from a single wild population to obtain accurate estimates of allele frequencies and to avoid the confounding effects of "admixing" data from multiple demes (Wright and Gaut 2005). Our study site consists of a large complex of permanent and ephemeral wetlands separated by short distances (~100 m) that can sustain a large population of amphibians. Salamanders at our study site have been continuously studied since 2002. Since then, more than 600 adults have been sampled and over that time period, the recapture rate of adult salamanders was very low (<2%), an indication that the contemporary census population may be large. The genetic data agree, but unlike the natural history data, the molecules can help decipher the evolutionary origin and demographic history of this population.

The *ND2* gene mismatch distribution is useful in this regard and provides a clear indication of evolutionary history (Ray et al. 2003). The bimodal and uneven distribution of the *ND2* data is

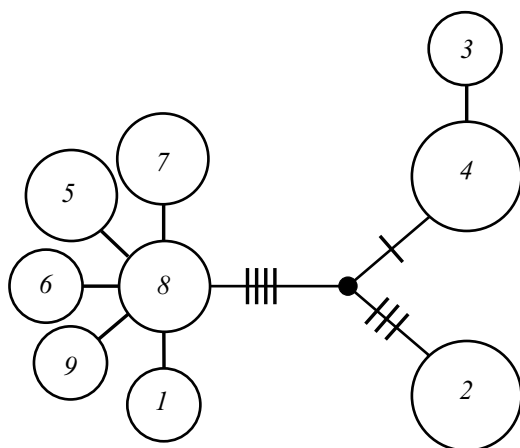


Figure 3. Haplotype network for the *ND2* gene as estimated by TCS ver. 1.2. Connections between haplotypes 1–9 (open circles) equal one mutational step and crossbars represent additional mutations; the missing intermediate haplotype is indicated by a solid circle. Haplotype numbers correspond to those in Table 2, and size of open circles represent high, medium, and low frequency haplotypes.

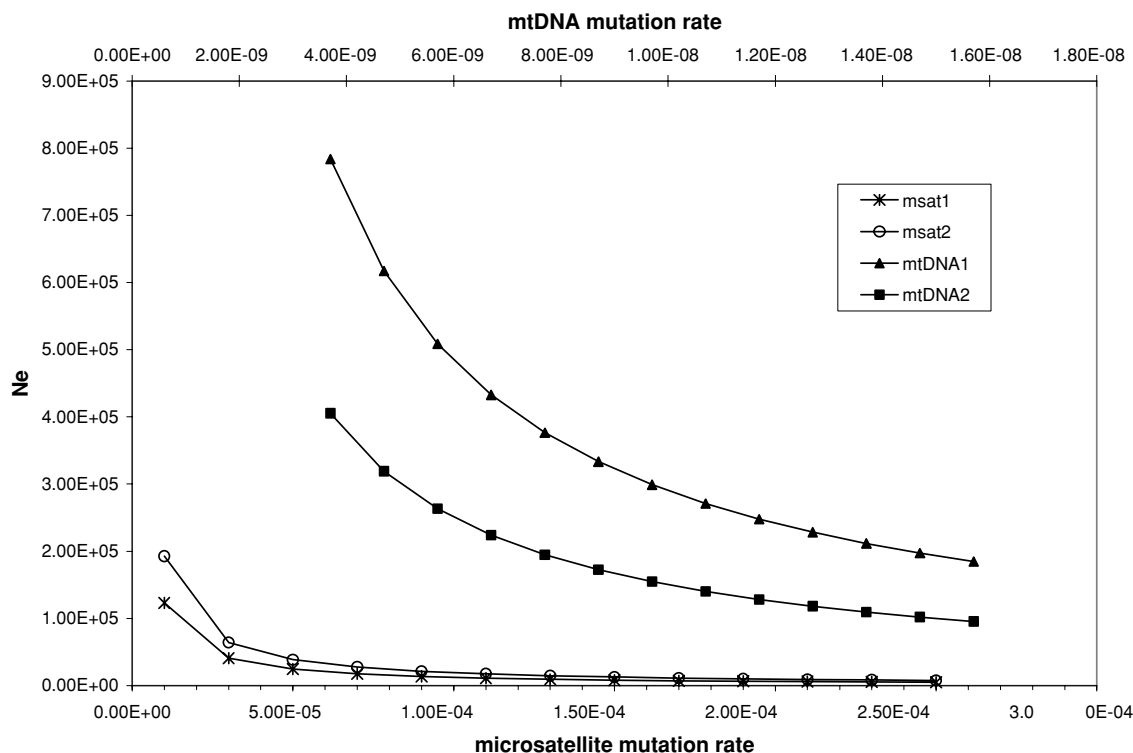


Figure 4. Effective population size estimates using microsatellite and mtDNA data. Various estimates are obtained using a range of substitution rates that are plausible for these data. Microsatellite method 1 relies on θ estimates from LAMARC, whereas method 2 is the heterozygosity method of Nei (1987). MtDNA method 1 relies on θ_f estimates from LAMARC, and method 2 uses the θ_f estimate of the mismatch distribution model.

most consistent with a demographic growth phase in a population that originated due to historical expansion from multiple demes that share few migrants (low Nm). The shape of the distribution is also sensitive to the age of the expansion, with older expansion events leading to a right-shifted unimodal peak, and relatively recent expansion to a bimodal distribution. We note further that a spatial expansion from multiple demes differentially affects the numbers of alleles and the average pairwise distance of alleles, producing conditions that are not favorable to detection of population size changes with the D and F_S statistics (Ray et al. 2003). Thus, the nonsignificant values for these statistics are not unexpected for populations that were derived from multiple demes because an assumed interrelationship between the number and distance of alleles is not met. Evidence for this interpretation can also be seen in the $ND2$ haplotype network, where multiple haplotypes are evolutionarily isolated from one another by multiple substitutions. The imbalance index based on the microsatellite data also reveals the signature of recent increases in population size, consistent with the mtDNA data.

The bimodal mismatch distribution and the imbalance index both support a model of spatial expansion (within the past $\sim 4,000$ generations; Ray et al. 2003). In *A. t. tigrinum*, such timing is consistent with events that occurred over the past 8,000 to 10,000

years. Our study site was almost certainly influenced by glaciation over the past 18,000 to 14,000 years, notably the Wisconsin glaciation (Pielou 1991). Thus it seems likely that this population originated during range expansion events following the retreat of the Wisconsin ice sheets at the end of the Pleistocene epoch, and occurred in areas exposed and rendered ecologically suitable for (re)colonization.

A population originally established from Pleistocene refugia may have experienced a reduction in population size at some point. This reduction could have occurred during the initial founding of the population and/or during fragmentation and isolation of Pleistocene refugia. The microsatellite imbalance index does indicate that before the population reached its contemporary (large) size, it was likely preceded by a reduction in size. However, the M -ratio and BOTTLENECK analyses failed to detect patterns consistent with recent bottlenecks in the contemporary gene pool. The discrepancy of these results may be explained by the severity and age of the bottleneck and/or the stability of the molecular signal (King et al. 2000; Garza and Williamson 2001). Also, the M -ratio and BOTTLENECK tests are known to be especially sensitive to gene flow, which rapidly degrades the signal of past bottlenecks (Busch et al. 2007). For example, BOTTLENECK is well suited for detection of recent reductions in N_e leading to shifts

from HWE, but is insensitive to earlier events as a population approaches neutral mutation–drift equilibrium (Cornuet and Luikart 1996). If our population did indeed originate via migration events from multiple sources then most signals of a previous bottleneck may be obscured, especially in light of the clear signal of recent population expansion (i.e., mismatch distribution and imbalance index). In fact, the estimates of long-term effective size at both mitochondrial and nuclear markers are consistent with a long-term large population.

EFFECTIVE POPULATION SIZE

N_e can be calculated directly from θ if mutation rates are known. For salamanders, a large range of mtDNA substitution rates have been estimated (Tan and Wake 1995; Caccone et al. 1997; Mueller et al. 2004; Samuels et al. 2005) but there is little published data on microsatellite mutation rates (although see Gopurenko et al. 2006). Therefore, we estimated N_e and N_{ef} using a range of plausible mutation rates for both microsatellites and mtDNA to represent these uncertainties. The long-term N_e estimates that result from our calculations are large and surprisingly, the mtDNA estimate is larger than the microsatellite-based estimates (Fig. 4).

In theory, this disparity between N_e and N_{ef} could result from behavior, historical demography, mutation, or some combination thereof. For example, *Ambystoma* breeding behavior leads to very high reproductive skew that reduces the number of successfully breeding males (Kelly 2001; Gopurenko et al. 2007). *Ambystoma tigrinum tigrinum* have male biased operational sex ratios and pronounced reproductive skew (Gopurenko et al. 2006), which should inflate N_{ef} while reducing N_e . Nevertheless, ecological factors alone are unlikely to account for the observed disparity between N_e and N_{ef} .

Historical events may also explain the apparent discord between mtDNA and microsatellite-based estimates of long-term population size. For instance, if the historical source populations that contributed to the origin of the contemporary study population were isolated from each other except for male-biased dispersal, then the populations would more rapidly diverge at mtDNA loci than nuclear loci because mtDNA inheritance is matrilineal in this taxon (D. Gopurenko and J. A. DeWoody, unpubl. data). The secondary contact (i.e., isolate breaking) of these isolates upon the founding of our study population would merge a relatively homogenous pool of nuclear genes, but mtDNA lineages would be differentiated. This evolutionary scenario would further contribute to the reversal in expected size for N_e and N_{ef} because our N_{ef} calculations may reflect the collective genetic diversity and divergence of all source populations.

Another untested possibility is that the high mutation rate of microsatellites compared to mtDNA has obscured some of the signal on the historical origin of the study population due to size homoplasy. Size homoplasy is particularly acute in populations

with a large N_e (Estoup et al. 2002). Thus, microsatellite-derived N_e estimates may not reflect the composite origin of this population as well as N_{ef} . Regardless of the ultimate cause of the discord between N_e and N_{ef} , all estimators indicate that the long-term size of this population is very large and thus the evolutionary potential of natural selection is strong.

NATURAL SELECTION

Relative efficiency of natural selection is a hallmark in populations of large size. The d_N/d_S was 3.12 at *Amti-DAB* but only 0.07 at *ND2*, indicating that positive selection on substitution pattern is evident at the MHC locus, but purifying selection dominates substitutions at the mitochondrial locus. As expected, the intensity of positive selection on *Amti-DAB* is similar to those for MHC genes in other species (e.g., in mammals selection intensity ranges from $S = 210$ to 3800 for methods I and IV [Satta et al. 1994; Richman et al. 2003]), and provides ample evidence of positive or balancing selection acting at this locus. In contrast, patterns of substitutions at *ND2* provide no convincing evidence of positive or balancing selection. Note that calculations of the selection intensity rely on parameters (number of nonsynonymous sites and the d_N/d_S ratio) that are independent of demographic effects (Nielsen 2001) so the discrepancy between the mtDNA and the MHC gene cannot be explained by differences in mode of transmission or effective population size between these loci. These differences in selection contribute to variable patterns of nucleotide divergence; for example, mean pairwise divergence between *Amti-DAB* alleles (8.1%) was nearly an order of magnitude higher than for *ND2* (0.93%).

The signature of natural selection on pattern of substitutions persists for a very long time, but the effect of selection on allele frequencies is more transient (Garrigan and Hedrick 2003). Therefore, differences between estimates using Satta et al.'s (1994) method I and IV may be due to variation in selection intensity over time. In this regard, these two methods may complement each other with regard to evaluating the timing of selection and its effects (e.g., on substitution patterns or allele frequencies). Additionally, the two methods may differ somewhat in their power to detect selection intensity. Method I uses the mean number of nonsynonymous substitutions, whereas method IV uses the number of alleles with at least one nonsynonymous difference in the PBR. Method I should be more sensitive to sequence length and method IV to the number of samples, but the two methods result in similar selection intensity estimates in humans (Satta et al. 1994), indicating that the difference in power between the two methods may be small.

Conclusions

If our large study population has indeed been stable in size over its recent history, then natural selection may have been a powerful evolutionary force on certain loci because the large N_e should

enable a strong response to selection (as evidenced by *Ami-DAB*). On the other hand, the concomitant lack of selection on the *ND2* gene suggests a much more neutral mode of evolution for the mtDNA molecule. Because the relative intensity of natural selection can be confounded by historical demography, we used the molecular data to reconstruct the most likely evolutionary history of the population and conclude that demographics probably played a relatively minor role in shaping contemporary genetic variation.

Combined, the data support a historical model of population establishment that we think best accounts for the patterns of genetic variation seen in our study population. We have demonstrated that a single, large, panmictic population like ours may not totally conform to neutral expectations despite the lack of overt evidence for population substructure or changes in size. Clearly, observed deviations from random expectations should be considered when comparing neutral markers and MHC genes, as should the history and demography of a population. In the future, we anticipate that simultaneous evaluations of adaptive and neutral genetic markers will help illuminate the evolutionary history of populations.

ACKNOWLEDGMENTS

We thank members of the DeWoody laboratory group and two anonymous referees for valuable comments that dramatically improved this manuscript. This research is contribution #ARP2007-18125 from Purdue University, and was supported by a grant from the National Science Foundation (DEB-0514815) to JAD.

LITERATURE CITED

- Aguilar, A., G. Roemer, S. Debenham, M. Binns, D. Garcelon, and R. K. Wayne. 2004. High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proc. Natl. Acad. Sci. USA* 101:3490–3494.
- Ballard, J. W. O., and M. C. Whitlock. 2004. The incomplete natural history of mitochondria. *Mol. Ecol.* 13:729–744.
- Bos, D. H., and J. A. DeWoody. 2005. Molecular characterization of major histocompatibility complex class II alleles in wild tiger salamanders (*Ambystoma tigrinum*). *Immunogenetics* 57:775–781.
- Bos, D. H., S. M. Turner, and J. A. DeWoody. 2007. Haplotype inference from diploid sequence data: evaluating the performance of Bayesian methods using non-neutral MHC sequences. *Hereditas* 144:228–234.
- Boyce, W. M., P. W. Hedrick, N. E. Muggli-Crockett, S. Kalinowski, M. C. T. Penedo, and R. R. Ramey II. 1997. Genetic variation of major histocompatibility complex and microsatellite loci: a comparison in bighorn sheep. *Genetics* 145:421–433.
- Busch, J. D., P. M. Waser, and J. A. DeWoody. 2007. Recent demographic bottlenecks are not accompanied by a genetic signature in two populations of banner-tailed kangaroo rats (*Dipodomys spectabilis*). *Mol. Ecol.* 16:2450–2462.
- Caccone, A., M. C. Milinkovitch, V. Sbordoni, and J. R. Powell. 1997. Mitochondrial DNA rates and biogeography in European newts (Genus *Euproctus*). *Syst. Biol.* 46:126–144.
- Campos, J. L., D. Posada, and P. Moran. 2006. Genetic variation at MHC, mitochondrial and microsatellite loci in isolated populations of Brown trout (*Salmo trutta*). *Conserv. Genet.* 7:515–530.
- Carrington, M., and S. J. O'Brien. 2003. The influence of *HLA* genotype on AIDS. *Annu. Rev. Med.* 54:535–551.
- Charlesworth, B., P. Sniegowski, and W. Stephan. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371:215–220.
- Charlesworth, B., D. Charlesworth, and N. H. Barton. 2003. The effects of genetic and geographic structure on neutral variation. *Annu. Rev. Ecol. Evol. Syst.* 34:99–125.
- Church, S., J. M. Kraus, J. C. Mitchell, D. R. Church, and D. R. Taylor. 2003. Evidence for multiple pleistocene refugia in the postglacial expansion of the eastern tiger salamander, *Ambystoma tigrinum tigrinum*. *Evolution* 57:372–383.
- Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9:1657–1660.
- Cohen, S. 2002. Strong positive selection and habitat-specific amino acid substitution patterns in MHC from an estuarine fish under intense pollution stress. *Mol. Biol. Evol.* 19:1870–1880.
- Cornuet, J. M., and G. Luikart. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144:2001–2014.
- Elena, S. F., V. S. Cooper, and R. E. Lenski. 1996. Punctuated evolution caused by selection of rare beneficial mutations. *Science* 272:1802–1804.
- Estoup, A., P. Jarne, and J. M. Cornuet. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Mol. Ecol.* 11:1591–1604.
- Ewens, W. J. 1972. The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.* 3:87–112.
- Excoffier, L. 2004. Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Mol. Ecol.* 13:853–864.
- Fu, Y.-X. 1997. Statistical test of neutrality of mutations against population growth, hitchhiking, and background selection. *Genetics* 147:915–925.
- Garrigan, D., and P. W. Hedrick. 2003. Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution* 57:1707–1722.
- Garza, J. C., and E. G. Williamson. 2001. Detection of reduction in population size using data from microsatellite loci. *Mol. Ecol.* 10:305–318.
- Gopurenko, D., R. N. Williams, C. R. McCormick, and J. A. DeWoody. 2006. Insights into the mating habits of the tiger salamander (*Ambystoma tigrinum tigrinum*) as revealed by genetic parentage analyses. *Mol. Ecol.* 15:1917–1928.
- Gopurenko, D., R. N. Williams, and J. A. DeWoody. 2007. Reproductive and mating success in the small-mouthed salamander (*Ambystoma texanum*) estimated via microsatellite parentage analysis. *Evol. Biol.* 34:130–139.
- Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41:95–98.
- Hedrick, P. W., K. M. Parker, and R. N. Lee. 2001. Using microsatellite and MHC variation to identify species, ESUs, and MUs in the endangered *Sonoran topminnow*. *Mol. Ecol.* 10:1399–1412.
- Hill, A. V. S., C. E. M. Allsopp, D. Kwiatkowski, N. M. Anstey, P. Twumasi, P. A. Rowe, S. Bennett, D. Brewster, A. McMichael, and B. M. Greenwood. 1991. Common West African HLA antigens are associated with protection from severe malaria. *Nature* 352:595–600.
- Kelly, M. J. 2001. Lineage loss in serengeti cheetahs: consequences of high reproductive variance and heritability of fitness on effective population size. *Conserv. Biol.* 15:137–147.
- Kimmel, M., R. Chakraborty, J. P. King, M. Bamshad, W. S. Watkins, and L. B. Jorde. 1998. Signatures of population expansion in microsatellite repeat data. *Genetics* 148:1921–1930.
- Kimura, M., and T. Ohta. 1969. The average number of generations until fixation of a mutant gene in a finite population. *Genetics* 61:763–771.

- King, J. P., M. Kimmel, and R. Chakraborty. 2000. A power analysis of microsatellite-based statistics for inferring past population growth. *Mol. Biol. Evol.* 17:1859–1868.
- Kuhner, M. K., J. Yamato, P. Beerli, L. P. Smith, E. Rynes, E. Walkup, C. Li, J. Sloan, P. Colacurcio, and J. Felsenstein. 2005. LAMARC v 1.2.1. Univ. of Washington, Seattle.
- Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: Molecular Evolutionary Genetics Analysis software. Arizona State Univ., Tempe.
- Landry, C., and L. Bernatchez. 2001. Comparative analysis of population structure across environments and geographical scales at major histocompatibility complex and microsatellite loci in Atlantic Salmon (*Salmo salar*). *Mol. Ecol.* 10:2525–2539.
- Mueller, R. L. 2006. Evolutionary rates, divergence dates, and the performance of mitochondrial genes in Bayesian phylogenetic analysis. *Syst. Biol.* 55:289–300.
- Mueller, R. L., J. R. Macey, M. Jaekel, D. B. Wake, and J. L. Boore. 2004. Morphological homoplasy, life history evolution, and historical biogeography of plethodontid salamanders inferred from complete mitochondrial genomes. *Proc. Natl. Acad. Sci. USA* 101:13820–13825.
- Nei, M. 1987. *Molecular evolutionary genetics*. Columbia Univ. Press, New York.
- Nielsen, R. 2001. Statistical tests of selective neutrality in the age of genomics. *Heredity* 86:641–647.
- Parra-Olea, G., E. Recuero, and K. R. Zamudio. 2007. Polymorphic microsatellite markers for Mexican salamanders of the genus *Ambystoma*. *Mol. Ecol. Notes* 7:818–820.
- Peakall, R., and P. E. Smouse. 2006. GENELEX 6: genetic analysis in Excel. Population genetics software for teaching and research. *Mol. Ecol. Notes* 6:288–295.
- Pielou, E. C. 1991. *After the ice age*. Univ. of Chicago Press, Chicago.
- Potts, W. K., C. J. Manning, and E. K. Wakeland. 1994. The role of infectious disease, inbreeding and mating preferences in maintaining MHC diversity: an experimental test. *Philos. Trans. R. Soc. Lond.* 346:369–378.
- Ray, N., M. Currat, and L. Excoffier. 2003. Intra-deme molecular diversity in spatially expanding populations. *Mol. Biol. Evol.* 20:76–86.
- Raymond, M., and F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicisim. *J. Heredity* 86:248–249.
- Richman, A., L. G. Herrera, and D. Nash. 2003. Evolution of MHC class II Eb diversity within the genus *Peromyscus*. *Genetics* 164:289–297.
- Rogers, A. R., and H. Harpending. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9:552–569.
- Routman, E. 1993. Population structure and genetic diversity of metamorphic and paedomorphic populations of the tiger salamander, *Ambystoma tigrinum*. *J. Evol. Biol.* 6:329–357.
- RoyChoudhury, A., and M. Stephens. 2007. Fast and accurate estimation of the population-scaled mutation rate, $\{\theta\}$, from microsatellite genotype data. *Genetics* 176:1363–1366.
- Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496–2497.
- Sambrook, J., and D. W. Russell. 2001. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Press, Cold Spring Harbor, New York.
- Samuels, A. K., D. W. Weisrock, J. J. Smith, K. J. France, J. A. Walker, S. Putta, and S. R. Voss. 2005. Transcriptional and phylogenetic analysis of five complete ambystomatid salamander mitochondrial genomes. *Gene* 349:43–53.
- Satta, Y., C. O'hUigin, N. Takahata, and J. Klein. 1994. Intensity of natural selection at the major histocompatibility complex loci. *Proc. Natl. Acad. Sci. USA* 91:7184–7188.
- Schneider, S., D. Roessli, and L. Excoffier. 2000. ARLEQUIN: a software for population genetics data analysis (version 3.01). Univ. of Geneva, Department of Anthropology, Geneva.
- Schneider, S. D. and L. Excoffier. 1999. Estimation of past demographic parameters from the distribution of pairwise difference when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics* 152:1079–1089.
- Stephens, M., and P. Donnelly. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *Am. J. Hum. Genet.* 73:1162–1169.
- Stephens, M., N. J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68:978–989.
- Tajima, F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105:437–460.
- . 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.
- Tan, A.-M., and D. B. Wake. 1995. MtDNA phylogeography of the California newt, *Taricha torosa* (Caudata, Salamandridae). *Mol. Phylogenet. Evol.* 4:383–394.
- Templeton, A. R., E. J. Routman, and C. A. Phillips. 1995. Separating population structure from population history: a cladistic analysis of the geographical distribution of mitochondrial DNA haplotypes in the Tiger Salamander, *Ambystoma tigrinum*. *Genetics* 140:767–782.
- Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* 7:256–276.
- . 1977. Heterosis or neutrality? *Genetics* 35:789–814.
- Wegner, K. M., M. Kalbe, T. B. H. Reusch, and M. Milinski. 2003. Multiple parasites are driving the major histocompatibility complex polymorphism in the wild. *J. Evol. Biol.* 16:224–232.
- Williamson, S. H., R. Hernandez, A. Fledel-Alon, L. Zhu, R. Nielsen, and C. D. Bustamante. 2005. Simultaneous inference of selection and population growth from patterns of variation in the human genome. *Proc. Natl. Acad. Sci. USA* 102:7882–7887.
- Wright, S. I., and B. S. Gaut. 2005. Molecular population genetics and the search for adaptive evolution in plants. *Mol. Biol. Evol.* 22:506–519.

Associate Editor: L. Knowles