8-2007

Population Genetics of the Carinate Pillsnail, *Euchemotrema hubrichti*: Genetic Structure on a Small Spatial Scale

Frank E. Anderson
Southern Illinois University Carbondale, feander@siu.edu

Recommended Citation

Title: Population genetics of the carinate pillsnail, *Euchemotrema hubrichti*: genetic structure on a small spatial scale

Frank E. Anderson

Department of Zoology and Center for Ecology, Life Science 2, 1125 Lincoln Drive, Southern Illinois University, Carbondale, IL 62901-6501 USA

Corresponding author: F. E. Anderson, Department of Zoology, Life Science 2, 1125 Lincoln Drive, Southern Illinois University, Carbondale, IL 62901-6501 USA
Fax: (618) 453-2806
e-mail: feander@siu.edu

Running title: *Euchemotrema hubrichti* population genetics

Abstract

Euchemotrema hubrichti (Pilsbry, 1940) is a polygyrid land snail known only from the Larue Pine Hills region of southwestern Illinois, USA. Despite the restricted range of the species, observations in the field suggested that patterns of genetic variation within *E. hubrichti* might show geographic structure. To assess the dispersal potential of *E. hubrichti*, a mark-recapture experiment was performed at two sites <100 meters apart. To evaluate the geographic distribution of genetic variation, a region of the mitochondrial cytochrome c oxidase subunit I gene was sequenced from 53 snails collected from sites spanning the species range. Mark-recapture experiments suggested that population sizes at two sites were low and migration between sites was not detected. AMOVA and nested clade analysis (NCA) revealed significant genetic structure among sites. Two sites in particular appeared to be isolated from the others based on nested clade analysis. AMOVA also detected significant genetic differentiation among sites, but partitioning the data into two groups (populations 1 and 2) based on the NCA results did not explain a significant amount of genetic variation. Coalescent methods suggested that population sizes and migration rates for populations 1 and 2 are unequal. These patterns are likely due to the low dispersal ability of *E. hubrichti* and patchiness of preferred *E. hubrichti* habitat rather than an ancient vicariant event. These preliminary findings have implications both for the conservation of this species and for the population genetics of any spatially restricted endemic species with low dispersal ability.

Introduction

Gene flow or lack thereof can be of critical importance in the maintenance of genetic diversity in endangered or threatened species. A major factor determining the amount of gene flow within and among animal populations is dispersal ability (vagility) (Bohonak 1999). In terrestrial habitats, small flightless animals without a phoretic, airborne or aquatic dispersal phase might be expected to show higher levels of population differentiation—i.e., lower levels of gene flow—than more vagile animals like birds or mammals, even in the absence of historic barriers to gene flow (Irwin 2002). Dispersal ability also has implications for species conservation, as low-vagility species could be particularly sensitive to environmental perturbations as well as habitat loss or degradation. Furthermore, sensitivity to environmental perturbation may be exacerbated if the species is rare and has a small range or a patchy distribution.

Land snails are often presumed to be low-vagility organisms, although there is some evidence that they can disperse via occasional hitchhiking on other organisms (Dundee *et al.* 1967; Gittenberger *et al.* 2006; Rees 1965). High levels of cryptic genetic diversity have been found in several land snail species (e.g., Arnaud *et al.* 1999; Thomaz *et al.* 1996; Watanabe & Chiba 2001), suggesting that the limited dispersal ability of land snails can have an important effect on the geographic distribution of genetic variation. Land snail species with restricted ranges and/or patchy distributions may therefore be of particular conservation concern.

The carinate pillsnail, *Euchemotrema hubrichti* (Pilsbry, 1940), is a polygyrid land snail with a highly restricted range and a patchy distribution (Anderson 2005). *Euchemotrema hubrichti* was described based on shell material collected from Pleistocene talus at the base of a limestone bluff in the Larue–Pine Hills region of southwestern Illinois, USA. Initially, the species was believed to be extinct (Pilsbry 1940), because no live individuals had been found. Hubricht (1943) subsequently described his first encounter with live carinate pillsnails, and much later noted the distribution of the species and described the habitat of *E. hubrichti* as "crevices of shaded cliffs, often more than 20 ft above the ground" (Hubricht 1985). These snails seem to prefer the xeric tops of relatively exposed limestone outcrops, where they often can be found in high abundance aestivating beneath limestone slabs (Anderson 2005; Anderson & Smith 2005).

The restricted distribution of this species may qualify it for federal threatened or endangered status. *Euchemotrema hubrichti* is considered a Regional Forester Sensitive species by the United States Department of Agriculture Forest Service. It has a NatureServe (www.natureserve.org) Global Status of G1—critically imperiled across its entire range—and an IUCN Red List of Threatened Species designation of DD ("data deficient") (Mollusc Specialist Group 1996). An IUCN status of "DD" indicates that there is insufficient information to assess the extinction risk of a taxon. Fortunately, the known range of *E. hubrichti* lies entirely within the LaRue–Pine Hills/Otter Pond Research Natural Area, a part

of the Shawnee National Forest where collections are prohibited. The total known range of *E. hubrichti* extends roughly north to south along the edge of the Larue-Pine Hills limestone bluff system for approximately 5 km. Within this region, the distribution of *E. hubrichti* appears to be rather patchy. *E. hubrichti* is found in high abundance (>100 individuals) at only three sites and in low abundance at a few other sites (Anderson 2005).

If *E. hubrichti* is truly a low-vagility organism, geographic population structure could arise even across relatively small distances, particularly given the apparent preference of these snails for patchily distributed exposed limestone outcrops. If geographically based genetic structure does exist in *E. hubrichti*, an understanding of this structure could be valuable for species conservation. This leads to two questions: 1) Is *Euchemotrema hubrichti* actually a low-vagility organism? 2) If so, is this low vagility reflected in the genetic structure of the species?

Several approaches were used to address these questions. First, mark-recapture methods were used to estimate the number of snails at two proximate sites (< 100 meters apart) and to determine if migration of snails between the two sites could be directly detected over a period of three years. To investigate population structure, a region of the mitochondrial cytochrome oxidase subunit I (COI) gene was PCR amplified and sequenced from snails collected throughout the Larue-Pine Hills region. The sequence data were analyzed using nested clade analysis (NCA) to investigate the geographic component of population structure. Results of the NCA were evaluated further using a combination of standard population genetic methods (AMOVA) and coalescent-based methods.

Methods

Mark-recapture

Population sizes, overwinter survivorship and migration rates at two sites (C and D; Figure 1) were estimated via mark-recapture methods. When the mark-recapture work was initiated, only sites C and D were known to harbor large numbers of *E. hubrichti* individuals. Snails were collected by hand from each site, and shells of live snails were marked with different colors of nail polish to denote collection site and year. Population size estimates for each site were combined across years using the Schnabel method; samples taken during the summer of 2003 were also analyzed with the Jolly-Seber method. The Schnabel method makes stringent assumptions (i.e., no birth, death, immigration or emigration have occurred), but specimens do not have to be individually marked. The Jolly-Seber method allows for an open population, but requires that snails be individually marked so that date of last capture can be recorded (Krebs 1999).

Specimen collection, DNA extraction, PCR amplification and sequencing

53 snails were collected from seven sites in the Larue–Pine Hills Research Natural Area in Union County, southwestern Illinois (Figure 1). Rarity of snails at several sites is the primary reason for the small sample sizes; *all* snails captured at "low abundance" sites (A, B, E and G) were assayed for COI. Small pieces of foot tissue (~1–2 mm^2) were isolated from live snails using sterile razor blades, as recommended by Holland and Hadfield (2002). Most of the snails were marked and allowed to recover for several days in the lab prior to being released alive at the sites from which they were collected. Four voucher specimens were retained for morphological study and deposited at the Field Museum of Natural History (catalogue number 307844) as part of a redescription of this species (Anderson & Smith 2005). Total DNA was extracted using the DNeasy DNA isolation kit (QIAGEN). A ~650-bp region of the mitochondrial COI gene was amplified via PCR (25-μL reactions, annealing temperature = 42°C) using HotStar Master Mix (QIAGEN) and primers from Folmer *et al*. (1994), along with a negative control. PCR amplicons were gel-purified from a 1% TBE agarose gel using a MinElute Gel Purification Kit (QIAGEN) to remove primer dimers, and gel-purified amplicons were sequenced in both directions using the PCR primers with BigDye Terminator Ready Reaction v.3.0 mix (one-eighth reactions). Sequencing reactions were cleaned using Centri-Sep spin columns (Princeton Separations) and run out on an ABI 377 automated sequencer. All kits were used following manufacturer's protocols.

Phylogenetic analysis

Maximum parsimony (MP) and maximum likelihood (ML) phylogenies were inferred for the haplotypes using branch-and-bound searches in PAUP* 4.0b10 (Swofford 2002). To select an appropriate model for ML analysis, fifty-six ML substitution models were evaluated on the MP tree using hierarchical likelihood-ratio tests (hLRT's) and the Akaike Information Criterion (AIC) in Modeltest 3.7 (Posada & Crandall 1998). All ML model parameters were estimated from the data during the tree search. Strength of support was assessed via nonparametric bootstrapping (one hundred pseudoreplicates, each analyzed with branch-and-bound searches under MP and ML) (Felsenstein 1985).

Nested clade analysis and AMOVA

Nested cladistic analysis (NCA; also referred to as "nested clade phylogeographical analysis" or NCPA), as described in a series of papers by Templeton and colleagues (Templeton 1998; 2004; Templeton *et al.* 1992; Templeton *et al.* 1995; Templeton & Sing 1993) is a method for 1) testing the null hypothesis of no association between haplotypes (and clades) and geography and 2) using an inference key for *a posteriori* discrimination among historical events that may have produced significant haplotype/geography associations. NCA has been summarized by several authors (including Paulo *et al.* 2002; Pearse & Crandall 2004; Turner *et al.* 2000; Verovnik *et al.* 2004 and citations therein). A parsimony network was constructed using TCS version 1.21 (Clement *et al.* 2000) and nesting of haplotypes and clades was performed following the methods outlined in Templeton and Sing (1993). Statistical analyses of

geographical associations were performed with GEODIS version 2.4 (Posada *et al.* 2000). The GEODIS program calculates clade (D_c) and nested clade (D_n) distances for all clades. D_c measures the geographic spread of a clade; D_n measures the geographic spread of a group relative to all haplotypes present in the nesting clade. Statistical significance of these values was tested using permutational analysis (1000 resamples). The GEODIS output was evaluated using the inference key from Templeton (2004), available online at http://darwin.uvigo.es/download/geodisKey_14Jul04.pdf.

The geographical pattern of population structure was also assessed using Φ statistics calculated via an analysis of molecular variance (AMOVA, Excoffier *et al.* 1992), which partitions the observed genetic variation into within-population (Φ_{ST}), among-populations-within-groups (Φ_{SC}) and among-group (Φ_{CT}) components, taking both sequence divergence and haplotype frequencies into account. For this study, populations were hierarchically arranged in two ways: as members of a single group including all sites or partitioned into one of two groups based on inferences from nested clade analysis. Tamura distances (Tamura 1992) were used to construct the distance matrix for AMOVA. AMOVA's were performed with the Arlequin version 2.000 software package (Schneider *et al.* 2000). Statistical significance of the variance components and Φ statistics was evaluated by obtaining the null distributions of Φ_{ST}, Φ_{SC} and Φ_{CT} using 10,000 permutations of haplotypes within sites, haplotypes within groups and sites between groups, respectively (Excoffier *et al.* 1992). Population pairwise F_{ST}'s were also calculated using Arlequin.

Coalescent analyses

The IM program (Hey & Nielsen 2004) was used to estimate several population genetic parameters of interest. The program uses an "isolation with migration" model developed by Nielsen and Wakeley (2001) and modified by Hey and Nielsen (2004) to generate coalescent-based estimates of m_x ($= m_x/\mu$, where m_x is the immigration rate of individuals into population *x* and µ is the mutation rate per gene per generation), the genetic diversity parameter θ ($= 4N_e\mu$) and time of population divergence *t* (= tµ, where t is the time since splitting).

The sequence data were partitioned into two groups based on results from the NCA analysis: one including all sequences recovered from sites A-D and G (referred to hereafter as Population 1) and the other including sequences from sites E and F (Population 2). An initial run was performed with the upper limits of the prior distributions for θ_1, θ_2, m_1, m_2 and *t* all set to 10, with a burn-in of 1 million steps followed by a 12-hour run using an infinite-sites model. The inheritance scalar was set at 0.5 to make the results comparable to those from a diploid autosomal locus. The mitochondrial genome is haploid and maternally inherited, so the scalar should normally be 0.25, but *E. hubrichti* is hermaphroditic (Anderson & Smith 2005), so all adults can potentially reproduce as females. Three subsequent runs were performed with priors conditioned on the initial run, each with different starting seeds and

the following upper prior distribution bounds: $\theta_1 = 20$, $\theta_2 = 10$, $m_1 = 10$, $m_2 = 30$ and t = 100, with a run length of eight hours and burn-in as described above. Estimates of effective sample sizes for each parameter were used to assess mixing, and runs were assumed to be sampling from the stationary distribution for each parameter if posterior distributions generated by all runs were similar.

Additional parameters of interest can be extracted from IM output if the mutation rate and generation time are known. The COI mutation rate in *E. hubrichti* is unknown, but mutation rates for this gene have been estimated for several other mollusks, including geminate species across the Isthmus of Panama. For this study, a range of divergence rates (0.67 – 2.4% per million years) was used to estimate N_e and migration rates for *E. hubrichti*, following Donald *et al*. (2005). This range encompasses COI divergence rate estimates from trans-Panamanian arcid bivalves (Marko 2002), the marine gastropod *Tegula* (Hellberg & Vacquier 1999) and the freshwater gastropod *Hydrobia* (Wilke & Pfenninger 2002). This equates to a range of mutation rates of 0.335 – 1.2% per million years (i.e., one-half the divergence rate), or (after multiplying by 576, the number of COI nucleotides examined here) 1.92 – 6.912 mutations per gene fragment per million years = 1.92 – 6.912 x 10^{-6} mutations per year. The generation time of *E. hubrichti* is also unknown, but a generation time of one year is reasonable given findings with lab cultures of larger species of polygyrid snail (T. Pearce, pers. comm.), leading to a mutation rate range of 1.92 – 6.912 x 10^{-6} mutations per generation.

Results

Mark-recapture

Results of the mark-recapture surveys are shown in Table 1. An additional "high abundance" site (F) was found in the last year of the study; this site was not included in the mark-recapture study. Mature *E. hubrichti* can survive over at least two winters; several snails that were marked in the summer of 2001 were recovered alive in the summer of 2003. Across the three years of mark-capture work at the two sites, no individuals originally from site C were found at site D or vice-versa.

Sequence comparisons and statistics

A total of 53 COI sequences were generated, revealing eight unique haplotypes (Table 2). Data alignment was unambiguous, and translation of the nucleotide sequences into amino acids revealed no stop codons, suggesting that the sequences are functional mitochondrial protein-coding gene sequences. Across the 576 base pairs analyzed, uncorrected ("p") pairwise distances among the haplotypes ranged from a high of 0.01389 (between haplotypes I-VIII) to a low of 0.00174 (between haplotypes III-IV). Variation among the haplotypes was

low: only 10/576 sites (1.73%) were variable and, of these, only five were parsimony-informative. Nine substitutions were at the third codon position, one was at the first codon position, and all changes were transitions (3 A⇔G, 7 C⇔T). Nucleotide diversity (π) ranged from zero (sites A, B, G) to 0.005022 (site E).

Phylogenetic analysis

The best-fitting model for the data set was the HKY85 (Hasegawa *et al.* 1985) model with no among-site rate variation parameters. P-values for the LRT's were (null model vs. alternative model) as follows: F81 (Felsenstein 1981) vs. HKY = 0.000001, HKY vs. Tamura-Nei (Tamura & Nei 1993) = 0.325023, HKY vs. HKY+ G > 0.999999, HKY vs. HKY + I > 0.999999. The AIC value for the HKY model was 1645.2354, and the Akaike weight was 0.2361 (the model with the second highest Akaike weight was TrN = 0.1410). MP and ML analyses of the eight haplotypes resulted in one tree (Figure 2; TL = 10, -lnL = 818.62845). There is a well-supported separation of the haplotypes into two clades, one containing haplotypes I, II, III and IV and the other containing the remaining haplotypes.

Nested clade analysis and AMOVA results

The haplotype network produced by TCS for the *E. hubrichti* data is shown in Figure 3. Nested clade analysis was performed only on clades showing both geographic and genetic variation (clades 1-2, 1-3, 2-1, 2-2 and the total network), and significant associations between haplotype and geography were found for all five of these groups (Table 3). Clade distances, nested clade distances, $D_c\text{I-} D_c\text{T}$ and $D_n\text{I-} D_n\text{T}$ values for all clades are shown in Figure 4, and NCA inferences for all five groups are listed in Table 3. Contiguous range expansion was inferred for Clade 1-2. Two clades—Clade 2-2 and the total cladogram—showed evidence of restricted gene flow (with evidence for some long-distance dispersal for Clade 2-2 and isolation by distance for the total cladogram). Allopatric fragmentation was inferred for the remaining two clades (Clades 1-3 and 2-1). These inferred fragmentation events separate most individuals found in E and F from the rest of the network (Figure 3).

When all sites were considered members of one hierarchical group for AMOVA, significant genetic structure was detected among sites (Φ_{ST} = 0.52439, P < 0.00001; Table 4). NCA suggested a fragmentation event between E+F and all other sites, but when sites were segregated into two groups—population 1 (all sites but E and F) and population 2 (sites E and F)—to match this inference, no significant structure was detected (Φ_{CT} = 0.34662; P < 0.14218). Variation among sites within these two groups was significant (Φ_{SC} = 0.36458, P < 0.00001). Pairwise site Φ_{ST} values are shown in Table 5.

Coalescent analyses

The location of the highest values of m_1, m_2, θ_1 and θ_2 in the marginal posterior distributions computed by IM, averaged across three independent runs, were: θ_1 = 2.3899 (0.6983-6.0903), θ_2 = 0.8938 (0.1311-3.0866), m_1 = 0 (0.0050-2.8350), m_2 = 0.855 (0.0150-6.6450) (ranges in parentheses are the shortest intervals in each distribution that include 90% of the probability) (Figure 5). The marginal posterior distributions of θ_A (θ for the ancestral population) and t were flat, and will not be discussed further. Assuming a range of plausible gastropod COI mutation rates, N_e for population 1 = 86,440 - 311,185, N_e for population 2 = 32,328 - 116,380, m_1 (the rate of migration of gene copies from population 2 into population 1) = 0 and m_2 = 1.6414 - 5.9098 x 10^{-6} per year (or generation, assuming a generation time of one year).

Discussion

Mark-recapture experiments at two of the sites studied here suggest that migration rates among sampling sites are not substantial. Over three summers, no snails from C were found at D or vice-versa, despite relatively high numbers of snails at both sites and the small distance (< 100 m) between them (Table 1, Figure 1). This finding initially suggested that gene flow among *E. hubrichti* populations may be limited. Estimates of population sizes varied depending on the method used, but ranged from about 200 individuals to nearly 1000 individuals.

COI variation in *E. hubrichti* is relatively low compared with levels of variation in this gene region in other gastropod species (e.g., Holland & Hadfield 2002; Liu *et al.* 2003), but those studies focused on species with much broader geographic ranges than *E. hubrichti*. The relatively low level of sequence variation within the species could be due to a number of factors. The species could be relatively young (i.e., thousands to tens of thousands of years old, in which case substantial sequence variation has not had time to accumulate), or it could be hundreds of thousands or millions of years old and simply have a low substitution rate. Although the Illinoian glaciation stopped just north of the Larue-Pine Hills region (and certainly must have had a profound impact on the local biota), the Larue-Pine Hills region itself was not glaciated during the Pleistocene (Killey 1998). Thus, it is possible that *E. hubrichti* persisted through the glacial maximum. *Euchemotrema hubrichti* fossils have been found from Pleistocene talus at the base of the main bluff (Pilsbry 1940), but precise dating of these fossils has not been attempted. If the origin of *E. hubrichti* substantially predates the Illinoian glaciation, there are at least two explanations for the low level of sequence variation: 1) the rate of mtDNA evolution in the species has been low, or 2) *E. hubrichti* has gone through one or more bottlenecks that have reduced genetic variation.

Despite the low level of genetic variation, the AMOVA and NCA results demonstrate that there are significant associations between geography and genetic variation in this species. Of particular interest is the apparent separation of E and F (i.e., Population 2) from the remaining sites (Population 1). E and F share a haplotype (VI) found nowhere else. In addition, two other haplotypes (I and II) are found only at E and another (VIII) is found only at F. Given

the proximity of E and F (they are ~200 m apart), the sharing of a haplotype is unsurprising. However, the failure to find other haplotypes at E and F is somewhat peculiar. E and F lie almost directly on a north-south line between D and G (Figure 1), which share a haplotype (IV) that was also found at the northernmost sites sampled (A and B; Figure 3).

NCA infers an allopatric fragmentation between populations 1 and 2, but the separation between these populations does not match all of the possible patterns resulting from allopatric fragmentation as described by Templeton *et al*. (1995). For example, the clades/haplotypes isolated by the presumed fragmentation event are not connected to the rest of the cladogram by a "larger-than-average" number of mutational steps. However, as also noted by Templeton *et al*. (1995), fragmentation is presumably an event of some sort, and as such different subsets of the haplotype network should coincide with geographically congruent populations. This is the case for *E. hubrichti*. Two groups of haplotypes at opposite ends of the haplotype network—Clade 1-1 (comprising haplotypes I and II) and haplotypes VI and VIII—are separated from the rest of the network by an inferred fragmentation event, and these haplotypes are only found at E and F. The separation between Population 1 and Population 2 is not complete, however—although four haplotypes are found only at either E or F, haplotype III is found at C, D and E.

There is some discrepancy between the NCA and AMOVA results. As noted previously, NCA infers an allopatric fragmentation event separating E and F from the remaining sites. When the data were partitioned into two groups—one including E and F (Population 2), the other including all remaining sites (Population 1)—no significant genetic structure was detected. All pairwise Φ_{ST} comparisons involving F were significant, supporting the inference that F is genetically differentiated from most of the other populations.

Although the presence of a phylogeographic discontinuity in the COI gene tree for *E. hubrichti* may be due to a fragmentation event at some point in the past, as suggested by NCA, such a discontinuity could also arise at random. Such discontinuities, which are not associated with any historical barrier to gene flow, may be especially common in species with low population sizes or dispersal rates (Irwin 2002; Kuo & Avise 2005). Furthermore, variation in habitat quality throughout the range of a species can exacerbate the tendency for random phylogeographic breaks to arise within a single gene genealogy (Irwin 2002). In the absence of additional evidence (e.g., data from multiple loci or other species), such random phylogeographic breaks are difficult to distinguish from discontinuities produced by historic barriers to gene flow. The mark-recapture experiment described here strongly suggests that *E. hubrichti* has both a low dispersal rate and small population sizes (but see below). *E. hubrichti* is also patchily distributed—all sites where *E. hubrichti* has been found in abundance are exposed limestone. The Larue-Pine Hills bluffs are composed of a lower Devonian cherty (Bailey) limestone that is covered in many places by a layer of loess (Harris *et al*. 1977; Killey 1998; Wiggers 1997). This loess layer (now covered with grasses or hardwood forest and leaf litter) has produced a series of exposed limestone "islands" that appear to be the preferred

habitat for *E. hubrichti*. Loess and vegetation may act as a partial dispersal barrier for *E. hubrichti*, limiting movement among exposed outcrops. Thus, for several reasons, *E. hubrichti* (and many other low-dispersal species with patchy distributions and small population sizes) may be particularly prone to displaying phylogeographic breaks in the absence of historic, long-term barriers to gene flow.

Coalescent methods offer a way to incorporate the stochasticity of the coalescent process into estimation of population genetic parameters such as θ and *m* (Knowles 2004; Pearse & Crandall 2004). The results from the IM analyses are largely concordant with aspects of the biology of *E. hubrichti*, although estimates of N_e are substantially larger than census population size estimates at two sites would suggest. The asymmetry in estimated θ (and thus N_e) values for populations 1 and 2 are reasonable—population 1 includes two known "high-abundance" sites (C and D, the sites evaluated in the mark-recapture study) and covers a much larger spatial range than population 2—even though the absolute values are higher than expected. The migration rate estimates suggest that migration from population 1 to population 2 has occurred (although it is very infrequent), but migration from population 2 to population 1 has effectively been zero. This is analogous to expectations from island biogeography theory (MacArthur & Wilson 1967). If population 1 is analogous to a continental population (relatively large in number of species/individuals and area), and population 2 (relatively small by both measures) is analogous to an island population, asymmetric migration rates are unsurprising. Site F in particular stands out in this respect, as it is a small limestone outcrop that is separated by several meters from the other limestone bluffs that constitute preferred *E. hubrichti* habitat. As such, site F may act as an island for this species, with zero dispersal back to the "mainland" population 1. The inference from NCA of restricted gene flow via isolation by distance for the total cladogram also seems to support this explanation.

Low-dispersal species—species that should be ideal candidates for phylogeographic research—may be particularly prone to displaying phylogeographic discontinuities in the absence of long-term barriers to gene flow (Irwin 2002), but this may be largely irrelevant for conservation considerations. These results demonstrate that 1) the distribution of COI haplotypes in *E. hubrichti* is geographically structured (AMOVA and NCA) and 2) *E. hubrichti* migration rates among some sites are extremely low over short (mark-recapture) and long (IM) time spans. This study shows that factors that limit gene flow (particularly low dispersal rates coupled with small population sizes and patchy habitat) can be important even over small spatial scales. The two populations revealed here may not merit special protection as separate conservation units (Moritz 1995), but they do suggest that natural recolonization of sites from which snails have been extirpated (either naturally or by human activities such as controlled burns) may be a very slow process, given the low migration rates detected by mark-recapture experiments and coalescent-based population genetic analyses. Given these observations, protecting both populations may be prudent.

Acknowledgements

References

Anderson FE (2005) *Conservation assessment for the carinate pillsnail, Euchemotrema hubrichti.* USDA Forest Service (Eastern Region), Milwaukee, Wisconsin.

Anderson FE, Smith DA (2005) A redescription of the carinate pillsnail, *Euchemotrema hubrichti* (Pilsbry, 1940) (Pulmonata: Polygyridae), with notes on habitat and genetics. *Zootaxa*, **807**, 1-11.

Arnaud JF, Madec L, Bellido A, Guiller A (1999) Microspatial genetic structure in the land snail *Helix aspersa* (Gastropoda : Helicidae). *Heredity*, **83**, 110-119.

Bohonak AJ (1999) Dispersal, gene flow, and population structure. *Q. Rev. Biol.*, **74**, 21-45.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Mol. Ecol.*, **9**, 1657-1660.

Donald KM, Kennedy M, Spencer HG (2005) Cladogenesis as the result of long-distance rafting events in South Pacific topshells (Gastropoda, Trochidae). *Evolution*, **59**, 1701-1711.

Dundee DS, Phillips PH, Newsom JD (1967) Snails on migratory birds. *The Nautilus*, **80**, 89-91.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479-491.

Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.*, **17**, 368-376.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783-791.

Folmer O, Black M, Heoh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Mol. Mar. Biol. Biotechnol.*, **3**, 294-299.

Gittenberger E, Groenenberg DSJ, Kokshoorn B, Preece RC (2006) Molecular trails from hitch-hiking snails. *Nature*, **439**, 409-409.

Harris SE, Horrell CW, Irwin D (1977) *Exploring the land and rocks of southern Illinois : a geological guide*. Southern Illinois University Press, Carbondale.

Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.*, **21**, 32-38.

Hellberg ME, Vacquier VD (1999) Rapid evolution of fertilization selectivity and lysin cDNA sequences in teguline gastropods. *Mol. Biol. Evol.*, **16**, 839-848.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747-760.

Holland BS, Hadfield MG (2002) Islands within an island: phylogeography and conservation genetics of the endangered Hawaiian tree snail *Achatinella mustelina*. *Mol. Ecol.*, **11**, 365-375.

Hubricht L (1943) Hunting *Stenotrema hubrichti*. *The Nautilus*, **56**, 73-75.

Hubricht L (1985) The distributions of the native land mollusks of the eastern United States. *Fieldiana, Zoology, N.S.*, **24**, 1-191.

Irwin DE (2002) Phylogeographic breaks without geographic barriers to gene flow. *Evolution*, **56**, 2383-2394.

Killey MM (1998) *Illinois' ice age legacy*. Illinois State Geological Survey, Champaign, Illinois.

Knowles LL (2004) The burgeoning field of statistical phylogeography. *J. Evol. Biol.*, **17**, 1-10.

Krebs CJ (1999) *Ecological methodology*. Benjamin/Cummings, Menlo Park, California.

Kuo CH, Avise J (2005) Phylogeographic breaks in low-dispersal species: the emergence of concordance across gene trees. *Genetica*, **124**, 179-186.

Liu HP, Hershler R, Clift K (2003) Mitochondrial DNA sequences reveal extensive cryptic diversity within a western American springsnail. *Mol. Ecol.*, **12**, 2771-2782.

MacArthur RH, Wilson EO (1967) *The theory of island biogeography*. Princeton University Press, Princeton, N.J.

Marko PB (2002) Fossil calibration of molecular clocks and the divergence times of geminate species pairs separated by the Isthmus of Panama. *Mol. Biol. Evol.*, **19**, 2005-2021.

Mollusc Specialist Group (1996) *Stenotrema hubrichti*. 2004 IUCN Red List of Threatened Species. Web site address: http://www.redlist.org (accessed December 27, 2004).

Moritz C (1995) Uses of molecular phylogenies for conservation. *Philos. Trans. Roy. Soc. B*, **349**, 113-118.

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: A Markov chain Monte Carlo approach. *Genetics*, **158**, 885-896.

Paulo OS, Jordan WC, Bruford MW, Nichols RA (2002) Using nested clade analysis to assess the history of colonization and the persistence of populations of an Iberian Lizard. *Mol. Ecol.*, **11**, 809-819.

Pearse DE, Crandall KA (2004) Beyond F-ST: Analysis of population genetic data for conservation. *Conserv. Genet.*, **5**, 585-602.

Pilsbry HA (1940) Land Mollusca of North America (north of Mexico), Volume 1, Part 2. *Monographs of the Academy of Natural Sciences, Philadelphia*, **3**, I-VI, 2 unnumbered, 575-994, I-IX.

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**, 817-818.

Posada D, Crandall KA, Templeton AR (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Mol. Ecol.*, **9**, 487-488.

Rees WJ (1965) The aerial dispersal of Mollusca. *Proc. Malac. Soc. Lond.*, **36**, 269-282.

Schneider S, Roessli D, Excoffier L (2000) Arelquin ver. 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland,

Swofford DL (2002) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Associates, Sunderland, Massachusetts.

Tamura K (1992) Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G+C content biases. *Mol. Biol. Evol.*, **10**, 678-687.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.*, **10**, 512-526.

Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Mol Ecol*, **7**, 381-97.

Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Mol. Ecol.*, **13**, 789-809.

Templeton AR, Crandall KA, Sing CF (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619-633.

Templeton AR, Routman E, Phillips CA (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.

Templeton AR, Sing CF (1993) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. IV. Nested analyses with cladogram uncertainty and recombination. *Genetics*, **134**, 659-669.

Thomaz D, Guiller A, Clarke B (1996) Extreme divergence of mitochondrial DNA within species of pulmonate land snails. *Proc. R. Soc. Lond. B*, **263**, 363-368.

Turner TF, Trexler JC, Harris JL, Haynes JL (2000) Nested cladistic analysis indicates population fragmentation shapes genetic diversity in a freshwater mussel. *Genetics*, **154**, 777-785.

Verovnik R, Sket B, Trontelj P (2004) Phylogeography of subterranean and surface

    populations of water lice *Asellus aquaticus* (Crustacea : Isopoda). *Mol. Ecol.*, **13**, 1519-

    1532.

Watanabe Y, Chiba S (2001) High within-population mitochondrial DNA variation due to

    microvicariance and population mixing in the land snail *Euhadra quaesita* (Pulmonata :

    Bradybaenidae). *Mol. Ecol.*, **10**, 2635-2645.

Wiggers R (1997) *Geology underfoot in Illinois*. Mountain Press Publishing Company, Missoula,

    Montana.

Wilke T, Pfenninger M (2002) Separating historic events from recurrent processes in cryptic

    species: phylogeography of mud snails (*Hydrobia* spp.). *Mol. Ecol.*, **11**, 1439-1451.

Figure Legends

Figure 1. Map showing the positions of *E. hubrichti* collection sites along the limestone bluff line in the Larue-Pine Hills Research Natural Area, Union County, southwestern Illinois, USA. The approximate location of the Larue-Pine Hills region is highlighted on the inset map. Population 1 comprises sites A-D and G; population 2 comprises sites E and F.

Figure 2. Maximum likelihood phylogram for the eight unique *E. hubrichti* COI haplotypes found in this study. Numbers on branches are MP (above) and ML (below) bootstrap support values. The scale bar indicates 0.001 substitutions per site. The tree is arbitrarily rooted.

Figure 3. Haplotype network estimated from the *E. hubrichti* COI data. Haplotype frequencies are proportional to the area of the associated circle (e.g., haplotype I was found in a single snail). Small open circles indicate missing haplotypes that were not observed. Each haplotype is coded by locations where the haplotype was found (see Table 2 for further information). The figure also shows the nesting design used for nested clade analysis. Haplotype III has the highest outgroup probability (0.274). Thick bars between clades 1-1 and 1-2 and haplotypes V and VI denote inferred allopatric fragmentation events.

Figure 4. Results of the nested geographical analysis of the *E. hubrichti* COI haplotypes. The nested design, haplotypes and clade designations are shown in Figure 3. Clade (D_c) and nesting clade (D_n) distances are reported for each level of the nesting design. Superscripts refer to significantly small (S) and large (L) distances. The average differences between interior (I) and tip (T) clade and nested clade distances, where available, are presented in rows labeled "D_xI-D_xT".

Figure 5. Marginal posterior distributions for θ (A, B) and *m* (C, D) for populations 1 and 2, averaged over three independent runs.

N

Big Muddy
River

to Mississippi RIver
(~ 4 km from bluff)

Government
Rock

Shawnee
National
Forest
(Larue-Pine Hills
National Natural Landmark)

A

B

C

D

E

F

McGee
Hill

G

Illinois

1 km

Site A ▤ E ▦
B ■ F ▥
C ▦ G ▦
D ☐