5-1997

Evolutionary Genetics of Canyon Treefrogs (*Hyla arenicolor*)

Roy A. Murray
Utah State University

EVOLUTIONARY GENETICS OF CANYON TREEFROGS

(*HYLA ARENICOLOR*)

by

Roy A. Murray

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Biology Ecology

Approved:

Paul G. Wolf Major Professor	James W. Haefner Committee Member
Frank J. Messina Committee Member	Eugene W. Schupp Committee Member
D. Richard Cutler Committee Member	James P. Shaver Dean of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

1997

ABSTRACT

Evolutionary Genetics of Canyon Treefrogs (*Hyla arenicolor*)

by

Roy A. Murray, Doctor of Philosophy

Utah State University, 1997

Major Professor: Dr. Paul G. Wolf
Department: Biology

Population genetics is the study of the mechanisms that cause genetic change in populations over time. Genetic changes may lead to both adaptive evolution and speciation. While the former process is fairly well understood, many questions remain unanswered with regard to the process of speciation. How important is population isolation in the process of speciation? How long must populations be isolated before speciation is complete? Are the genetic changes that take place during speciation caused mainly by natural selection or does genetic drift play a substantial role? Can genetic drift alone lead to reproductive isolation? These types of questions have been debated since the time of Darwin, but only for the last 30 years have scientists had the tools to examine genes directly at the population level.

This study is a first look at the population genetics of Canyon Treefrogs. Both the physiology of these frogs and the region they inhabit have contributed

to a highly isolated population structure. Employing both laboratory analysis and computer modeling, I have examined the genetic changes occurring among these populations and also addressed biogeographical questions regarding how disjunct population structures arise in nature.

Isozyme and DNA sequencing results indicate an exceptional period of isolation among widely separated populations of Canyon Treefrogs. Eastern and western populations of frogs may have been isolated for as long as 1.5 million years. Divergence of the eastern-most population is at a level seen most often among congeneric amphibians based on results of previous studies. Sequencing analysis also points to the possibility of mtDNA introgression in two southern populations of Canyon Treefrogs.

Laboratory and computer modeling results suggest the importance of vicariant events in the establishment of the current disjunct population structure. Populations probably expanded their ranges along local drainage corridors during favorable periods. Subsequent shifts in drainage patterns severed avenues of gene flow, leading to genetic isolation.

(140 pages)

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Early evolutionary theory emphasized the importance of population isolation in the process of speciation. Both Darwin and Wallace were impressed by the abundance of endemic species on oceanic islands, and these observations contributed to their belief that a cessation of genetic exchange was important in the formation of new species. These perspectives remained central to evolutionary theory, establishing the course of research through much of the modern synthesis period and focusing attention on the biotic and physical factors that might lead to a cessation of gene flow among populations. The view of species as evolutionary units held together by the cohesive action of gene flow was not seriously questioned until the late 1960's when Ehrlich and Raven (1969) challenged the importance of gene flow based on three observations: (1) gene flow appeared to be more restricted in nature than previously thought; (2) there was often little differentiation among disjunct populations; (3) populations were shown to differentiate in the presence of gene flow under different selective pressures. Ehrlich and Raven published their paper at the dawn of the molecular age; hence much of their evidence was necessarily based on morphological measures of differentiation and direct estimates of gene flow (i.e., individual movements). They acknowledged the paucity of data available on gene flow in natural populations and emphasized

the importance of future study for a resolution of the issue.

Almost 30 years later, there are an abundance of published studies estimating levels of gene flow in nature, but the question of its relevance in the evolutionary process remains largely unresolved. The volumes of molecular data that have been accumulated in the intervening period suggest that gene flow is indeed often restricted among disjunct populations. The evidence for this lies in the observed differences in allele frequencies and DNA sequences among populations. It is interesting to note that whereas Ehrlich and Raven's supposition that gene flow is often restricted has been supported, the very evidence in support contradicts their idea that isolated populations are undifferentiated. However, much of the observed molecular variation is thought to be neutral with respect to fitness; thus its importance to the process of speciation remains unclear.

In support of observation (3) above, Endler (1973, 1977) and others have shown that allelic gradations in the presence of gene flow are fairly common in nature, including what appear to be clines arising under primary intergradation. Clines are not always due to selective gradients but may be artifacts of secondary contact or arise under isolation by distance (Berry and Kreitman, 1993; Endler, 1977). Hence the homogenizing effects of gene flow are not always sufficient to prevent population differentiation even in the absence of selection. Whether or not the speciation process can go to completion (i.e., complete reproductive isolation) in the presence of gene flow

is unclear but theory suggests that assortative mating is a necessary prerequisite.

Isolated populations have also attracted attention recently due to the questions they raise with regard to species conservation. Biologists have recognized that the same factors governing biotic processes on oceanic islands may now be at work in islands of habitat created by anthropogenic alterations to surrounding landscapes (Bierregaard et al., 1992). Foremost of these processes from a conservation perspective is extinction. According to island biogeography theory (MacArthur and Wilson, 1967), the rate of extinction on islands is influenced by island size; smaller islands provide fewer resources and will therefore support smaller populations. Both demographic and genetic factors presumably play an increasingly influential role in the extinction process as population sizes become small (Lande, 1988). The important genetic processes in small populations include inbreeding and loss of genetic variation due to random drift, both of which may reduce the mean fitness of a population (O'Brien et al., 1985; Ralls et al., 1988). In addition to the role of island size in determining extinction rates, the degree of insularization is likely to play a role as well. The extent to which an island or habitat patch is isolated from potential immigrants can have a profound effect on the longevity of a population occupying the particular patch (Harrison, S., 1991; Sjögren, 1991); immigrants potentially provide both breeding individuals and genetic variation to isolated demes. Thus, from a conservation

perspective, it is important to gain an understanding of the unique genetic processes occuring in small, isolated populations in order to implement well-informed conservation efforts.

One of the more recent developments that is allowing researchers to approach these questions is the ability to reconstruct intraspecific phylogenies (Harrison, R. G., 1991; Miles and Dunham, 1993). Prior to the molecular era, studies of intraspecific phylogeny were rarely attempted because of the limited amount of character variation within a species. Due to advances in molecular techniques such as restriction fragment length polymorphism (RFLP) analysis and direct DNA sequencing, intraspecific studies are now common in the literature (Avise et al., 1987). Intraspecific phylogenies provide an important piece of the puzzle, because they reveal the ancestral relationships among disjunct populations, which in turn provide clues to the biogeographic processes that may have contributed to population isolation. Additionally, spatial differentiation of populations is a similar process to temporal divergence, such that the manner in which populations differ spatially can be used to infer how populations change over time. In many instances molecular data can also provide an estimate of how much time has passed since disjunct populations shared a common ancestor, allowing an assessment of how much morphological or behavioral change can be expected over evolutionary time scales.

The study described in the following pages examines the genetics,

phylogeny, and biogeography of populations of Canyon Treefrogs (*Hyla arenicolor*), which exist in isolated pockets of moist habitat throughout the hot desert regions of the southwestern United States and south to central Mexico. The distribution of the species is remarkable in itself, covering more than 2000 kilometers north to south throughout an area characterized by extremes of aridity and temperature. This distribution is even more remarkable given the species' apparently limited adaptations to these extremes (Preest et al., 1992; Snyder and Hammerson, 1993; Wylie, 1981). Canyon Treefrog physiology and behavior constrain their existence to permanent and semi-permanent sources of water, which in the U.S. portion of their range is most often provided by runoff from the widely separated mountains of the Basin and Range Province and the high landforms of the Colorado Plateau. The current study examines populations in four major drainage basins: the Colorado River, the Gila River, the Rio Grande, and the Canadian River drainages. The Colorado and Gila Rivers share a confluence near Lake Mead; the remaining drainage basins currently share no connection. Thus populations of *H. arenicolor* are likely to be isolated genetically and as a result may show marked differentiation at neutral loci depending on how long their current disjunct distribution has been in place.

Levels of genetic variation within and among populations of *H. arenicolor* are of interest as an indication of the manner and extent of genetic change that can arise in natural systems under isolation. Within-population variation

provides an estimate of the effective size of the population, which in turn may give some indication of the recent demographic history of the population. Within-population variation may also provide some evidence of the genetic "health" of a population, in terms of levels of inbreeding and potential for adaptive evolution. Genetic variation among populations gives an indication of the extent and duration of their isolation and under certain assumptions can be used to provide indirect estimates of interdemic gene flow. The partitioning of variation among populations is also an important process from a conservation standpoint because each deme provides a separate storehouse for maintaining distinct alleles.

Phylogenetic relationships among populations of *H. arenicolor* are of interest to provide an historical framework from which to view information on genetic variation within the species. The phylogeny may also give some indication as to the biogeographic history of the species, such as clues as to whether the current disjunct population structure is a result of range contraction from a formerly widespread and continuous distribution, or whether current populations have been founded through a process of dispersal along drainage corridors.

Previous research on Canyon Treefrogs is limited. The systematic relationship of *H. arenicolor* within the genus has been examined using call structure (Blair, 1958, 1960), reproductive compatibility (Pierce, 1968, 1975), morphology (Jameson and Richmond, 1971), and isozymes (Hedges, 1986).

On the basis of call structure and morphology, *H. arenicolor* has traditionally been placed in the *versicolor* group along with *Hyla versicolor*, *Hyla avivoca*, *Hyla chrysoscelis*, and *Hyla andersoni*, all eastern U.S. in their distributions. However, Hedges' (1986) inclusive isozyme study grouped *H. arenicolor* with *Hyla eximia*, both of which occur in the southwestern U.S. and Mexico. This relationship would place *H. arenicolor* in the *eximia* group with *Hyla euphorbiacea*, *Hyla plicata*, and *Hyla walkeri*. Pierce's hybridization studies were of limited success due to the difficulty of rearing frogs in the laboratory through the metamorphosis period. They would not have provided useful systematic information in any case, because *H. eximia* was not included in the study. Thus, the phylogentic relationship of *H. arenicolor* to other members of the genus remains in question.

Duellman's (1970) volume contains notes on the morphology, natural history, and distribution of *H. arenicolor* in Mexico. Wylie (1981) researched the general biology of the species focusing on a single population in southern Arizona over several years. He described life-history and behavioral traits and looked at the physiological trade-offs associated with the frogs' basking behavior. Preest et al. (1992) measured the effects of body temperature and hydration state on metabolic rates in *H. arenicolor*, showing that metabolic rates are significantly affected by these parameters in active frogs but not in resting frogs. The ability of *H. arenicolor* to thermoregulate during diurnal basking is apparent from measurements of a four-fold increase in evaporative

water loss as body temperatures rise from 20°C to 30°C (Preest et al., 1992). Snyder and Hammerson (1993) compared evaporative water loss rates for *H. arenicolor* with that for other anurans, showing that *H. arenicolor* falls into the "moderately waterproof" category, which is typical for arboreal frogs. Canyon Treefrogs may lose up to 25% of their body mass to evaporative water loss in the course of a single day of basking. As with the Preest et al. (1992) study, Snyder and Hammerson (1993) found evidence that *H. arenicolor* actively regulate cutaneous water loss, which in turn allows them to regulate body temperature.

The goals of the present study are (1) to assess levels of genetic variation within and among populations, (2) to estimate levels of population divergence and gene flow, (3) to propose a population phylogeny based on molecular data, and (4) to examine the possible biogeographic scenarios that might have given rise to the species' current disjunct distribution.

LITERATURE CITED

AVISE, J. C., J. ARNOLD, R. M. BALL, E. BERMINGHAM, T. LAMB, J. E. NEIGEL, C. A. REEB, AND N. C. SAUNDERS. 1987. Intraspecific phylogeography: The mitochondrial bridge between population genetics and systematics. Ann. Rev. Ecol. Syst. 18:489-522.

BERRY, A., AND M. KREITMAN. 1993. Molecular analysis of an allozyme cline: Alcohol dehydrogenase in *Drosophila melanogaster* on the east coast of North

America. Genetics 134:869-893.

BIERREGAARD, R. O., JR., T. E. LOVEJOY, V. KAPOS, A. AUGUSTO DOS SANTOS, AND R. W. HUTCHINGS. 1992. The biological dynamics of tropical rainforest fragments. BioSci. 42(11):859-866.

BLAIR, W. F. 1958. Call structure and species groups in U.S. treefrogs (*Hyla*). Southwest. Nat. 3:77-89.

—. 1960. Mating call as evidence of relations in the *Hyla eximia* group. Southwest. Nat. 5:129-135.

DUELLMAN, W. E. 1970. The Hylid Frogs of Middle America. The University of Kansas, Lawrence.

EHRLICH, P. R., AND P. H. RAVEN. 1969. Differentiation of populations. Science 165:1228-1232.

ENDLER, J. A. 1973. Gene flow and population differentiation. Science 179:243-250.

—. 1977. Geographic Variation, Speciation, and Clines. Princeton University Press, Princeton, New Jersey.

HARRISON, R. G. 1991. Molecular changes at speciation. Ann. Rev. Ecol. Syst. 22:281-308.

HARRISON, S. 1991. Local extinction in a metapopulation context: An empirical evaluation. Biol. J. Linn. Soc. 42:73-88.

HEDGES, S. B. 1986. An electrophoretic analysis of holarctic hylid frog evolutions. Syst. Zool. 35:1-21.

JAMESON, D. L., AND R. C. RICHMOND. 1971. Parallelism and convergence in the evolution of size and shape in Holarctic *Hyla*. Evolution 25:497-508.

LANDE, R. 1988. Genetics and demography in biological conservation. Science 241:1455-1460.

MACARTHUR, R. H., AND E. O. WILSON. 1967. The Theory of Island Biogeography. Princeton University Press, Princeton, New Jersy.

MILES, D. B., AND A. E. DUNHAM. 1993. Historical perspectives in ecology and evolutionary biology: The use of phylogenetic comparative analyses. Ann. Rev. Ecol. Syst. 24:587-619.

O'BRIEN, S. J., M. E. ROELKE, L. MARKER, A. NEWMAN, C. A. WINKLER, D. MELTZER, L. COLLY, J. F. EVERMANN, M. BUSH, AND D. E. WILDT. 1985. Genetic basis for species vulnerability in the cheetah. Science 227:1428-1434.

PIERCE, J. R. 1968. Isolation and differentiation in the Canyon Treefrog, *Hyla arenicolor*. Ph.D. Dissertation, The University of Texas, Austin, Texas.

—. 1975. Genetic compatibility of *Hyla arenicolor* with other species in the family, Hylidae. Texas J. Sci. 16(3-4):431-441.

Preest, M. R., D. G. Brust, AND M. L. Wygoda. 1992. Cutaneous water loss and the effects of temperature and hydration state on aerobic metabolism of canyon treefrogs, *Hyla arenicolor*. Herpetologica 48:210-219.

Ralls, K., J. D. Ballou, AND A. Templeton. 1988. Estimates of lethal equivalents and the cost of inbreeding in mammals. Conserv. Biol. 2:185-193.

SJÖGREN, P. 1991. Extinction and isolation gradients in metapopulations: the case of the pool frog (*Rana lessonae*). Biol. J. Linn. Soc. 42:135-147.

SNYDER, G. K., AND G. A. HAMMERSON. 1993. Interrelationships between water economy and thermoregulation in the Canyon tree-frog *Hyla arenicolor*. J. Arid Envir. 25:321-329.

WYLIE, S. R. 1981. Effects of basking on the biology of the canyon treefrog, *Hyla arenicolor* Cope. Ph.D. Dissertation, Arizona State University, Phoenix, Arizona.

CHAPTER 2

AN ELECTROPHORETIC ANALYSIS OF GENETIC VARIATION AND POPULATION STRUCTURE IN *HYLA ARENICOLOR*

ABSTRACT: Canyon treefrogs (*Hyla arenicolor*) are wide ranging yet limited to isolated pockets of habitat, posing quandaries for both population genetics and biogeography. How do populations in isolation originate and what are the effects of isolation on the genetics of disjunct populations? In this study, protein electrophoresis was used to estimate levels of variation within populations of *H. arenicolor*, and genetic divergence and gene flow among populations. Levels of variation appear to reflect geographic location rather than present population size. Low levels of variation were detected in peripheral populations with sharply higher levels in more central populations. Genetic divergence among widely separated populations is high. Estimates of gene flow correlate well with geographic distance among populations, suggesting an isolation by distance effect in establishing patterns of divergence. The isolation by distance pattern may be a result of historical vicariant events that have introduced barriers to gene flow among populations.

INTRODUCTION

Canyon treefrogs (*Hyla arenicolor* Cope) are distributed in North America throughout the desert Southwest from southern Utah and Colorado to

central Mexico (Behler and King, 1979; Duellman, 1970). The distribution of this species in the U.S. spans both the Colorado Plateau Province and the eastern portion of the Basin and Range Province, areas characterized by low annual rainfall and extreme temperatures during the summer months. High landforms of the Colorado Plateau and widely spaced mountains of the Basin and Range Province serve as repositories for winter snow and summer rain, providing sources of permanent runoff in areas otherwise marked by xeric conditions. These mountain oases provide isolated patches of habitat for many non-desert adapted species in the Southwest region, including *H. arenicolor* (e.g., Brown, 1971). *Hyla arenicolor* is limited to canyons and rocky washes that support permanent or semi-permanent streams or ponds, and is thus restricted within the region to the comparatively mesic environments near sources of runoff.

These habitat restrictions, along with the inhospitable nature of intervening environments, are likely to place severe limitations on rates of genetic exchange among widely separated populations of *H. arenicolor*. Populations in isolation are expected to differentiate genetically over time as alleles drift in frequency, as alleles come under the influence of local selective factors, or as new alleles are introduced via mutation. Isolation may also lead to reduced levels of genetic variation within populations if effective population sizes are small. The characterization of these processes in natural populations is of interest from an evolutionary standpoint as local changes may ultimately

lead to the formation of new species. Furthermore, understanding genetic processes in isolated populations is important as human-induced alterations to the environment create islands of habitat, isolating populations into smaller and smaller areas, often with little opportunity for gene flow among habitat patches.

In addition to the external factors that play a role in restricting gene flow, Canyon Treefrogs are limited in their potential for dispersal due to physiological and behavioral characteristics. Physiological adaptations to extremes of temperature and aridity take on many forms in desert amphibians; in particular, the ability to regulate cutaneous water loss under desiccating conditions varies considerably among anuran taxa (Snyder & Hammerson, 1993). Within the genus *Hyla*, a few species appear to be moderately "waterproof" in that they have some ability to regulate evaporative water loss. Preest et al. (1992) and Snyder and Hammerson (1993) have demonstrated some ability of *H. arenicolor* to osmoregulate; however, measured rates of evaporative water loss by no means place *H. arenicolor* in the highly waterproof category among anurans. The ability to limit cutaneous water loss is most likely an adaptation coincident with the species' basking behavior. Basking occurs throughout much of the year and appears to increase growth rate in individuals by raising body temperatures (Preest et al.,1992; Wylie, 1981), but requires that the frogs remain close to water for daily rehydration.

Wylie (1981) examined adult dispersal in *H. arenicolor* using mark/recapture studies and found that between-year movements are generally

limited to a few meters within a drainage. Low dispersal is typical for anurans, which tend to have small home ranges. In the larval stage, *H. arenicolor* may experience forced dispersal downstream during heavy rains, but is restricted in its capacity for upstream dispersal by limited swimming ability. High water events probably play a role in adult dispersal as well. This form of forced dispersal may be important as a means of gene flow and as a mechanism for the establishment of new populations.

In general, however, the dispersal ability of *H. arenicolor* into the surrounding environment appears to be highly restricted by both physiology and behavior; movements are essentially constrained to short distances within drainage corridors. If this is indeed the case, how did the species attain such a wide-ranging, disjunct distribution throughout a region of exceptional aridity?

Disjunct distributions are usually explained by invoking either a dispersal hypothesis or a vicariance hypothesis. The dispersal hypothesis holds that species move from a center of origin, often through areas of unsuitable habitat, to colonize remote areas of suitable habitat. Alternatively, the vicariance hypothesis suggests that species initially expand to the edge of suitable habitat during periods of favorable conditions and subsequently achieve disjunct distributions of relict populations as conditions change, ranges contract, and barriers to individual movements arise. Which of these two possibilities has actually occurred can be difficult to demonstrate because patterns of divergence among populations can be similar in either instance. Given the

limited potential for long distance dispersal for *H. arenicolor*, it is unlikely that this has been a major factor in the establishment of the species' disjunct distribution. Therefore, I will focus on the likelihood that vicariant events have played an important role in isolating formerly connected populations, and establish whether or not patterns of genetic relatedness are consistent with this hypothesis.

The objectives of this study are (1) to estimate levels of genetic variation within and among populations of *H. arenicolor* ; (2) to use cluster analysis to examine phenetic similarity among populations based on allele frequencies; (3) to examine the relationship between estimates of gene flow among populations and geographic distances separating them; and (4) to use these data to determine the likelihood that vicariant events have played an important role in establishing the species' current distribution.

MATERIALS AND METHODS

Sampling

Hyla arenicolor was sampled over most of its range with the exception of sites in Mexico (Fig. 2.1; Table 2.1). The southwestern Utah region was heavily sampled to assess genetic structure among drainages within a watershed. In this area, samples were collected over several kilometers within a drainage, in adjacent drainages both upstream and downstream of their

confluence, and in non-adjacent drainages within the same watershed. In other regions, sampling was less intense and intended to detect between-region divergence only. Samples within a drainage were widely spaced (> 100 m) to insure that a representative sample was obtained. Notably absent from the present study are samples from southwestern Texas, where the status of the species is uncertain. A search of Big Bend National Park, Texas in June of 1993 yielded only two adult *H. arenicolor* (near Boot Spring), and no evidence of breeding activity. Weather patterns in southwest Texas in the spring of 1993 were unusually dry. However, communications with Park personnel in subsequent years have indicated that the status of *H. arenicolor* has not improved since that time.

Because this study is concerned with current and historical avenues of gene flow for *H. arenicolor*, present-day drainage patterns are of particular interest. The following is a brief description of each of the sample sites and corresponding drainage systems.

Southwestern Utah.--Zion National Park contains several distinct drainages, each of which is cut deeply into the Markagunt Plateau, imposing isolation by topography as well as by distance and desert conditions. These watercourses converge to form the Virgin River watershed, which flows southwest, joining the Colorado River system at Lake Mead in southern Nevada. Some of the drainages in Zion maintain continuous flow through the year and support large populations of *H. arenicolor*. North Creek is an

example; during the breeding season, hundreds of adult frogs can be seen basking in the course of a kilometer along the creek. In contrast, other drainages within the Park flow only after rainfall. Frog habitat in these ephemeral washes consists of large potholes cut into the sandstone, often more than a meter in diameter and depth. While these potholes provide a year-round source of moisture, frog populations they support are small in comparison to the sources of permanent flow. Five to 10 adults can be found near any given pothole during the breeding season (R.M., personal observation).

Two additional populations outside of Zion National Park but within the Virgin River watershed were sampled. These are the Leeds Canyon population and the Red Cliffs population. Both are west of Zion and are maintained by runoff from Pine Valley Mountain.

Central Arizona.--Oak Creek is separated from the Virgin River drainage by the Colorado River and the Grand Canyon. Water from Oak Creek flows south into the Verde River, continues on to the Salt River, and eventually reaches the Gila River west of Phoenix. The Gila River then flows west to its confluence with the Colorado River about 80 kilometers north of the Gulf of California.

Southern Arizona.--Three distinct populations were sampled: Madera Canyon, near the Mexican border, Marijilda Creek near Mt. Graham, and Cave Creek in the Chiricahua Mountains. Moisture from these drainages rarely

reaches an outlet as surface water. During wet periods, the east slopes of Mt. Graham and the east slopes of the Chiricahuas flow to the San Simon River and then to the Gila River in eastern Arizona. However, the San Simon riverbed is dry during much of the year. A similar situation exists for Madera Canyon, which drains into the ephemeral Santa Cruz River and on to the Gila River west of Phoenix.

Southwestern New Mexico.--Two areas were sampled: the Pine Creek drainage on the west side of the Mogollon Mountains and the Garcia Falls area in the San Mateo Mountains. Pine Creek flows south into the eastern reaches of the Gila River. The Garcia Falls runoff drains west into the Alamosa River, then south and east to the Rio Grande.

Eastern New Mexico.--The Canadian River drainage in northeastern New Mexico is an area outside of the Basin and Range Province and is distinctly different from the other sites in terms of its surrounding ecotypes. This part of New Mexico represents the western edge of the Great Plains; the terrain is rolling grassland with few trees. Morphologically, adult *H. arenicolor* in this region appear distinct from those at other sample areas, but more analysis is needed to confirm this observation. The Canadian River flows east to the Arkansas River and finally into the Mississippi.

Laboratory Techniques

Eggs and larvae were collected in the field and allowed to develop in the

laboratory to a pre-metamorphosis stage, at which time they were positively identified to species. Larvae were then euthanized by immersion in a 0.2% aqueous solution of 3-aminobenzoic ethyl ester (MS-222, Sigma, Inc.) immediately prior to extractions. Whole-animal extracts were prepared by homogenization in distilled deionized water and centrifugation to separate solids from soluble proteins. Samples were stored at -80°C prior to electrophoresis. Electromorphs were detected using 12% starch gels following the buffer and stain recipes of Hedges (1986). In Table 2.2 is a list of the enzymes used in the final data analysis.

Data Analysis

Resulting electrophoretic data were analyzed using BIOSYS-1 (Swofford and Selander, 1981). BIOSYS-1 provides estimates of genetic variation in terms of average heterozygosity, percent polymorphic loci, and mean number of alleles per locus. Nei's gene diversity indices (Nei, 1973) were used to assess genetic differentiation among subdivided populations. H_S measures gene diversity within populations, D_{ST} measures gene diversity among populations, and H_T is the total gene diversity, such that $H_S + D_{ST} = H_T$. The quantity G_{ST} is then defined as $G_{ST} = D_{ST}/H_T$, providing an estimate of genetic differentiation among populations similar to Wright's F_{ST} (Wright, 1951).

Nei's genetic identities (Nei, 1978) were calculated to assess the degree of phenetic similarity among populations. Cluster analysis was then applied to

genetic identity estimates using the UPGMA algorithm to provide a hierarchical depiction of population similarities.

Pairwise F_{ST} values were calculated and used to estimate gene flow (Nm) between populations based on the equation $F_{ST} = 1/(4Nm + 1)$ (Wright, 1931, 1951), where N is the effective population size and m is the migration rate among populations. I then examined the correlation of gene flow estimates with geographic distance between populations, using regression analysis to test for isolation by distance (Slatkin, 1993). Gene flow was plotted against geographic distance using two different criteria: great circle distance between populations and distance between populations along current drainage corridors. If gene flow along drainage corridors is currently affecting the correlation of alleles among populations, the latter regression should show a higher significance level. Conversly, a high degree of correlation between gene flow and great circle distances may be an indication of a more continuous distribution of the species in the past.

RESULTS

Banding patterns detected on starch gels were assumed to represent the products of gene loci inherited in Mendelian fashion. Nine of the 18 loci examined were polymorphic. Measures of genetic variation for each population are given in Table 2.3. Percent polymorphic loci using the 95% criterion ranged from 5.6% at Leeds Canyon and Red Cliffs, Utah to 33% in the Garcia

Falls, New Mexico population. Mean heterozygosity ranged from 0.006 at Canadian River to 0.071 in both the Garcia Falls, New Mexico and Oak Creek, Arizona populations. Mean number of alleles per locus was also highest in the Garcia Falls population at 1.4. Pooled estimates for all populations were 44% polymorphic loci, 0.037 mean heterozygosity, and 1.8 alleles per locus. Interestingly, the lowest levels of variation were found in populations at the periphery of the species range (i.e., southwest Utah and the Canadian River drainage) with higher levels of variation in more central populations. Low levels of variation within the heavily sampled region in Zion National Park precluded an analysis to detect unidirectional gene flow or additional hierarchical levels of genetic structure among drainages and watersheds.

Cluster analysis results (Fig. 2.2) show that the phenetic relationship among populations based on the nine polymorphic loci is as expected based on geographic location. The three most closely related populations, Zion National Park, Leeds Canyon, and Red Cliffs are all contained within the Virgin River watershed and exhibit very little differentiation based on the isozyme data. The Oak Creek population joins the Utah group at the 0.94 similarity level. The next major cluster contains the Garcia Falls, Pine Creek, Marijilda Creek, Cave Creek, and Madera Canyon populations, which constitute a southeastern Arizona/southwestern New Mexico complex. This cluster joins the Utah/central Arizona group at the 0.82 similarity level. Finally, the Canadian River population, one of only two sampled populations on the east

side of the Continental Divide, joins the cluster at the 0.72 similarity level, representing a highly divergent group.

Seven of the nine polymorphic loci exhibit high values of G_{ST} (Table 2.4), for a mean across loci of 0.72. These values indicate that 72% of the genetic variation for *H. arenicolor* exists among populations, suggesting that gene flow is limited at this level. Of the two loci with lower G_{ST} values, a single allele for *Pgm* occurred at high frequency across all populations. Only one individual was polymorphic for *Mdh*, limiting the usefulness of this locus in the gene diversity analysis.

In contrast to the pooled G_{ST} values (Table 2.4), several of the pairwise F_{ST}'s (Table 2.5) indicate high levels of gene flow between populations (e.g., Zion National Park and Leeds Canyon at F_{ST} = 0.003, Nm = 80). Regression analysis of *log(Nm)* vs. *log(distance)* plots (Fig. 2.3) reveals a higher degree of correlation for great circle distances (r^2 = 0.84) than for drainage distances (r^2 = 0.77) among populations. (It should be noted that high r^2 values are expected for these regressions since *Nm* estimates represent averages across loci.) The number of points in Fig. 2.3b is fewer due to the omission of Canadian River and Garcia Falls populations, which have no current connection to the other drainage systems. In each of these plots, the three points with the highest gene flow and smallest between-population distances result from comparisons among the three southwestern Utah populations. These data points have a large influence on the regression due to their positions near the endpoint of the

regression line. Since both genetic similarity and gene flow results suggest that the southwest Utah populations represent a single panmictic deme, these populations were combined and the regressions were recomputed (Fig. 2.4). The regressions in Fig. 2.4 further emphasize the much stronger correlation for gene flow vs. great circle distance ($r^2 = 0.72$) than for gene flow versus drainage distance ($r^2 = 0.27$), although when the obvious outliers (solid dots, Fig. 2.4) are eliminated, the difference is not as pronounced ($r^2 = 0.80$ for Fig. 2.4a and $r^2 = 0.61$ for Fig. 2.4b).

Discussion

Genetic Variation Within Populations

Measures of genetic variation for *H. arenicolor* fall within the range of estimates for other Hylidae. Hedges (1986) studied protein variation in 30 species of hylid frogs with percentage of polymorphic loci ranging from zero to 33 with a mean of 11.5, and average heterozygosity ranging from zero to 0.13 with a mean of 0.053. Case et al. (1975) studied protein variation in four species of *Hyla* with percentage of polymorphic loci ranging from 7.2 to 50 and average heterozygosity ranging from 0.007 to 0.093. Nevo (1978) compiled results from 243 electrophoretic studies and showed levels of variation for vertebrates ranging from 0 to 51 for percent polymorphic loci and 0 to 0.24 for mean heterozygosity.

Theoretically, levels of genetic variation for neutral alleles within populations depend on the effective sizes of the populations. However, because several factors contribute to effective population size, the relationship between levels of variation and actual population size is not always straightforward. The present data set for *H. arenicolor* provides an example of the discrepancies that can arise. Personal observations (R.M.) suggest that the populations in and around Zion National Park are quite large, yet these populations exhibit exceptionally low levels of genetic variation. In contrast, populations at Garcia Falls and Cave Creek appear to be small in comparison yet show much higher levels of variation. No rigorous effort was made to estimate actual population sizes; not only are such efforts problematic, but preliminary results suggested an attempt to correlate genetic variation with population size would lack significance. This is not an unusual finding; measures of genetic variation often fail to correlate with estimates of population size (Hamrick et al., 1991; Hartl and Hell, 1994; Nei and Graur, 1984; Nickrent and Wiens, 1989; Pemberton, 1985). There are several possible explanations for this lack of association. (1) Genetic variation depends heavily on historical population size. A population that experiences a period of reduced size, whether through fluctuations in birth and death rates or as a result of a founding event, will show reduced variation for many generations following the bottleneck (Nei et al., 1975). (2) Low variation can result from isolation; populations that regularly exchange migrants will maintain higher levels of

variation than populations that do not. (3) While the degree of isolation is important, so also is the time of isolation. Large populations that have been isolated for long periods of time may be closer to gene flow/genetic drift equilibrium levels than smaller populations that have exchanged migrants in the recent past.

Populations of *H. arenicolor* with the lowest estimates of variability (southwest Utah and Canadian River) are also on the northwestern and northeastern periphery of the species' range. In more centrally located populations, there are increased levels of variation. This trend has been predicted on theoretical grounds (Rohlf and Schnell, 1971) and has been noted empirically as well (Mayr, 1963; Prakash et al., 1969). Mayr (1963) suggested that peripheral populations exist in marginal environments, at the limits of a species' tolerances, and thus fewer genotypes will be able to persist in these areas. Carson (1955) postulated that low variation in peripheral populations is due to selection for increased amounts of recombination. However, using a computer model of isolation by distance, Rohlf and Schnell (1971) showed that the phenomenon may simply be a consequence of spatial arrangements. Since peripheral populations will have fewer adjacent populations with which to exchange migrants, immigration rates will be necessarily reduced leading to lower equilibrium levels of variation. Although this is a plausible explanation for the trend in *H. arenicolor*, the historical effects discussed earlier cannot be ruled out. If the peripheral populations are relicts from a process of range

contraction, they may have experienced a longer period of isolation than more centrally located populations, placing them closer to equilibrium levels of variation. Alternatively, if peripheral populations were founded via dispersal, they may exhibit reduced variation as a result of genetic bottlenecks during founding events.

Genetic Divergence among Populations

Gene diversities among populations of *H. arenicolor* (Table 2.4) reveal large deviations from Hardy-Weinberg proportions overall (H_T = 0.35), with 72% of this due to population subdivision (G_{ST} = 0.72). These numbers indicate that allele frequencies are quite different among populations, suggesting that gene flow is exceptionally limited at the regional scale. Further evidence for this is provided by the cluster analysis results (Fig. 2.2). Thorpe (1983) examined the relationship between measures of genetic similarity and level of taxonomic separation by compiling the results of a large number of genetic and biochemical studies. In his results for amphibians, he shows that most conspecifics fall above a genetic identity value of 0.8. Furthermore, the modal value of genetic identity for congeneric species is near 0.65 with few congeneric species above the 0.85 similarity level. Considering the genetic identity values for *H. arenicolor* in this context places the Canadian River population (I = 0.72) on a boundary between the conspecific and congeneric levels. The link between the southern Utah/northern Arizona group and the

southern Arizona/southern New Mexico group (I = 0.82) also represents a high degree of divergence for conspecific populations.

There are two conspicous incongruities between the phenogram shown in Fig. 2.2 and phylogenetic results for *H. arenicolor* based on mitochondrial DNA sequences (Murray and Wolf, submitted): (1) The phylogeny groups the Marijilda and Madera populations into a clade that is highly distinct from the remaining populations; (2) the phylogeny places the Canadian River population in a clade with other New Mexico and southern Arizona populations. It has been suggested that the exceptionally high level of sequence divergence for Marijilda and Madera populations (10-11%) points to a hybrid origin for the mitochondrial DNA haplotypes found in these populations (Murray and Wolf, submitted). Given the close association of these two populations with other populations in the region based on isozymes results, this may be the only plausible explanation for this level of sequence divergence within a species. The lack of congruence in results for the Canadian River population is a likely consequence of the difference between the phenetic and phylogenetic methods of analysis. Phylogenetic analysis in not sensitive to unique character states within a clade and thus may closely group two or more populations even though one of these is distinct phenetically.

Gene Flow and Vicariance

Wright's (1931) method for estimating gene flow from F_{ST} values is

based on an infinite island model of population structure, and assumes both neutrality of alleles and an equilibrium between the introduction of alleles through migration and the loss of alleles through drift. Slatkin and Barton (1989) demonstrated that F_{ST} provides a robust estimate of gene flow under a variety of population models and selection regimes. It is important to keep in mind, however, that gene flow estimates derived from allele frequency data are long-term averages of genetic exchange among populations, and can therefore reflect either current gene flow patterns or similarity due to ancestral relationships among populations that presently have no contact.

The highest levels of similarity for *H. arenicolor* appear among the three populations within the Virgin River drainage (Table 2.5). High levels of gene flow are not unexpected for these populations as their respective drainages share a confluence a few kilometers outside of the Red Cliffs area. The numbers suggest that these populations represent a single panmictic deme. Likewise, low gene flow between Canadian River and all other populations is not unexpected. More surprising are the levels of gene flow among the southern Arizona populations since their drainages are ephemeral and their confluences are hundreds of kilometers to the northwest. Also, Garcia Falls and Pine Creek have an *Nm* of 2.9, yet these drainages share no confluence; they are in fact on opposite sides of the Continental Divide.

Using a computer model of gene flow among geographically structured populations, Slatkin (1993) showed that the effect of restricted gene flow

among adjacent demes is to establish a linear relationship between the logarithm of gene flow and the logarithm of geographic distance for pairs of populations. Applying this analysis to the gene flow estimates for *H. arenicolor* revealed a stronger correlation for great circle distances among populations than for drainage distances (Figs. 2.3 & 2.4), suggesting that patterns of similarity among populations are not related to dispersal events along current drainage corridors. Thus, genetic relatedness among disjunct populations of *H. arenicolor* is likely to reflect historical connections that have since been altered by climatic or geologic changes. There is ample geologic evidence for changes in major flow patterns throughout the desert Southwest (Axelrod, 1983; Hunt, 1974; Kottlowski et al., 1965; Lucchitta, 1972). Most notably, uplift of the Colorado Plateau, which occurred throughout the Tertiary period, redefined the drainage patterns of the southern Utah and northern Arizona regions. Additionally, there is evidence for major shifts in flow patterns for both the upper Gila River region and the entire Rio Grande channel. Shifting drainage patterns are a likely explanation for the establishment of isolated populations of *H. arenicolor* given the species' limited vagility. The present genetic relationship among populations would then reflect their time since last contact. Under this scenario, low levels of variation are expected in peripheral populations due to an equilibrium level of heterozygosity determined by mutation rates and a long period of isolation. Central populations may show higher levels of genetic variation due to more recent contact with adjacent

populations. A recent period of migration events would cause these central populations to be displaced from an equilibrium level of variation determined by mutation rates alone.

Summary

Isozyme anlysis of genetic variation within populations of *Hyla arenicolor* revealed low levels of variation in peripheral populations with higher levels in central populations. Genetic divergence among widely separated populations was high. Estimates of gene flow among populations are highly correlated with great circle distances separating them, but not well correlated with between-population distances as measured along current drainage corridors. This study suggests that isolation among populations of *H. arenicolor* is a result of vicariant events caused by shifting regional drainage patterns. Geologic evidence for major shifts in the Colorado River, Gila River, and Rio Grande drainage systems lends support to the hypothesis that shifting drainage patterns have played a significant role in establishing the current distribution of *H. arenicolor*.

LITERATURE CITED

AXELROD, D. I. 1983. Paleobotanical history of the western deserts. Pp. 113-129. *In* S. G. Wells and D. R. Haragan (Eds.), Origin and Evolution of Deserts. University of New Mexico Press, Albuquerque, New Mexico.

BEHLER, J. L., AND F. W. KING. 1979. The Audubon Society Field Guide to North American Reptiles and Amphibians. Alfred A. Knopf, Inc., New York.

BROWN, J. H. 1971. Mammals on mountaintops: Nonequilibrium insular biogeography. Am. Nat. 105:467-478.

CARSON, H. L. 1955. The genetic characteristics of marginal populations of *Drosophila*. Cold Springs Harbor Symposium on Quantitative Biology 20:276-289.

CASE, S. M., P. G. HANELINE, AND M. F. SMITH. 1975. Protein variation in several species of *Hyla*. Syst. Zool. 24:281-295.

DUELLMAN, W. E. 1970. The Hylid Frogs of Middle America. The University of Kansas, Lawrence, Kansas.

HAMRICK, J. L., M. J. W. GODT, D. A. MURAWSKI, AND M. D. LOVELESS. 1991. Correlation between species traits and allozyme diversity: Implications for conservation biology. Pp. 75-86. *In* D. A. Falk and K. E. Holsinger (Eds.), Genetics and Conservation of Rare Plants. Oxford University Press, Oxford, England.

HARTL, G. B., AND P. HELL. 1994. Maintenance of high levels of allelic variation in spite of a severe bottleneck in population size: The brown bear (*Ursus arctos*) in the Western Carpathians. Biodiv. Cons. 3:546-554.

HEDGES, S. B. 1986. An electrophoretic analysis of holarctic hylid frog evolutions. Syst. Zool. 35:1-21.

HUNT, C. B. 1974. Natural Regions of the United States and Canada. W. H. Freeman, San Francisco, California.

KOTTLOWSKI, F. E., M. E. COOLEY, AND R. V. RUHE. 1965. Quaternary geology of the Southwest. pp. 287-298. *In* H. E. Wright Jr. and D. G. Frey (Eds.), The Quaternary of the United States. Princeton University Press, Princeton, New Jersey.

LUCCHITTA, I. 1972. Early history of the Colorado River in the basin and range province. Geol. Soc. Am. Bull. 83:1933-1948.

MAYR, E. 1963. Animal Species and Evolution. Belknap Press, Cambridge, Massachusetts.

NEI, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Nat. Acad. Sci. USA 70(12):3321-3323.

—. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.

NEI, M., AND D. GRAUR. 1984. Extent of protein polymorphism and the neutral mutation theory. Evol. Biol. 3:418-426.

NEI, M., T. MARUYAMA, AND R. CHAKRABORTY. 1975. The bottleneck effect and genetic variability in populations. Evolution 29:1-10.

NEVO, E. 1978. Genetic variation in natural populations: patterns and theory. Theor. Pop. Biol. 13:121-177.

NICKRENT, D. L., AND D. WIENS. 1989. Genetic diversity in the rare California shrub *Dedeckera eurekensis* (Polygonaceae). Syst. Bot. 14(2):245-253.

PEMBERTON, J. M. 1985. Lack of biochemical variation in British fallow deer. Heredity 55:199-204.

PRAKASH, S., R. C. LEWONTIN, AND J. L. HUBBY. 1969. A molecular approach to the study of genic heterozygosity in natural populations IV. Patterns of genic variation in central, marginal and isolated populations of *Drosophila pseudoobscura*. Genetics 61:841-858.

PREEST, M. R., D. G. BRUST, AND M. L. WYGODA. 1992. Cutaneous water loss and the effects of temperature and hydration state on aerobic metabolism of canyon treefrogs, *Hyla arenicolor*. Herpetologica 48:210-219.

ROHLF, F. J., AND G. D. SCHNELL. 1971. An investigation of the isolation-by-distance model. Am. Nat. 105:295-324.

SLATKIN, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47:264-279.

SLATKIN, M., AND N. H. BARTON. 1989. A comparison of three methods for estimating average levels of gene flow. Evolution 43:1349-1368.

SNYDER, G. K., AND G. A. HAMMERSON. 1993. Interrelationships between water economy and thermoregulation in the Canyon tree-frog *Hyla arenicolor*. J. Arid Envir. 25:321-329.

SWOFFORD, D. L., AND R. K. SELANDER. 1981. BIOSYS-1. A computer program for the comprehensive analysis of electrophoretic data in population genetics and systematics. J. Hered. 72:281-283.

THORPE, J. P. 1983. Enzyme variation, genetic distance and evolutionary divergence in relation to levels of taxonomic separation. Pp. 131-152. *In* G. S. Oxford and D. Rollinson (Eds.), Protein Polymorphisms: Adaptive and Taxonomic Significance. Academic Press, New York, New York.

WRIGHT, S. 1931. Evolution in Mendelian populations. Genetics 16:97-159.

—. 1951. The genetical structure of populations. Ann. Eugen. 15:323-354.

WYLIE, S. R. 1981. Effects of basking on the biology of the canyon treefrog, *Hyla arenicolor* Cope. Ph.D. Dissertation, Arizona State University, Phoenix, Arizona.

TABLE 2.1.--Collection sites for *Hyla arenicolor*.

Population	Sample size	Locality	
Zion NP, UT	107	37° 16'N	112° 58'W
Leeds Canyon, UT	17	37° 17'N	113° 24'W
Red Cliffs, UT	15	37° 14'N	113° 25'W
Oak Creek, AZ	18	34° 52'N	111° 46'W
Marijilda Creek, AZ	28	32° 42'N	109° 48'W
Cave Creek, AZ	21	31° 52'N	109° 13'W
Madera Canyon, AZ	20	31° 44'N	110° 53'W
Pine Creek, NM	18	33° 12'N	108° 44'W
Garcia Falls, NM	21	33° 30'N	107° 27'W
Canadian River, NM	27	36° 01'N	104° 17'W

TABLE 2.2.--Enzyme systems used for analysis of genetic variation in *Hyla arenicolor*.

Protein	Locus	Enzyme commission number	Buffer system[a]
Acid phosphatase	*Acp*	3.1.3.2	5
Adenylate kinase	*Ak*	2.7.4.3	1
Aminopeptidase	*Apep*	3.4.11.1	4
Aspartate aminotransferase	*Aat*	2.6.1.1	1
Carbonate dehydratase	*Cd*	4.2.1.1	6
Creatine kinase	*Ck*	2.7.3.2	6
Dipeptidase	*Dpep*	3.4.13.11	4
Phosphoglucose isomerase	*Pgi*	5.3.1.9	5
Isocitrate dehydrogenase	*Idh*-1&2	1.1.1.42	1
Lactate dehydrogenase	*Ldh*-1&2	1.1.1.27	3
Malate dehydrogenase	*Mdh*	1.1.1.37	3
Malic enzyme	*Me*-1&2	1.1.1.40	3
Mannosephosphate isomerase	*Mpi*	5.3.1.8	5
Phosphoglucomutase	*Pgm*	2.7.5.1	5
Xanthine dehydrogenase	*Xdh*	1.2.1.37	6

[a] See Hedges (1986)

TABLE 2.3.--Genetic variability at 18 loci in all populations (standard errors in parentheses).

Population	Mean no. of alleles per locus	% of loci polymphc[a]	Mean heterozygosity	
			Direct count	HdyWbg expected[b]
1. Zion NP	1.1 (0.1)	11.1	0.018 (0.013)	0.017 (0.012)
2. Leeds Cn	1.1 (0.1)	5.6	0.016 (0.013)	0.015 (0.012)
3. Red Cliffs	1.1 (0.1)	5.6	0.019 (0.016)	0.016 (0.016)
4. Oak Cr	1.2 (0.1)	22.2	0.071 (0.038)	0.062 (0.033)
5. Marijilda	1.3 (0.1)	16.7	0.058 (0.035)	0.076 (0.041)
6. Cave Cr	1.2 (0.1)	22.2	0.056 (0.040)	0.041 (0.027)
7. Madera Cn	1.3 (0.2)	16.7	0.053 (0.029)	0.048 (0.027)
8. Pine Cr	1.3 (0.1)	27.8	0.068 (0.036)	0.077 (0.035)
9. Garcia Fls	1.4 (0.2)	33.3	0.071 (0.034)	0.126 (0.046)
10. Cana. R	1.1 (0.1)	5.6	0.006 (0.004)	0.010 (0.008)

[a] 95 percent criterion
[b] Unbiased estimate (see Nei, 1978)

TABLE 2.4.--Gene diversity measures at variable loci over all populations.

Locus	H_S	D_{ST}	H_T	G_{ST}
Ak	0.15	0.35	0.50	0.70
Idh-1	0.14	0.18	0.32	0.56
Ldh-1	0.10	0.40	0.50	0.80
Ldh-2	0.031	0.18	0.21	0.86
Me-1	0.093	0.23	0.32	0.72
Mpi	0.035	0.39	0.43	0.92
Pgi	0.12	0.54	0.66	0.82
Pgm	0.19	0.020	0.21	0.095
Mdh	0.011	0.001	0.012	0.083
Mean	0.096	0.26	0.35	0.72

TABLE 2.5.--Pairwise F_{ST}'s and gene flow measures (F_{ST} above diagonal, *Nm* below diagonal).

Population	1	2	3	4	5	6	7	8	9	10
1 Zion NP	*****	0.003	0.004	0.38	0.59	0.73	0.70	0.61	0.52	0.86
2 Leeds Cn	80	*****	0.005	0.45	0.55	0.79	0.76	0.65	0.49	0.93
3 Red Cliffs	58	44	*****	0.42	0.51	0.77	0.73	0.62	0.44	0.93
4 Oak Cr	0.42	0.30	0.34	*****	0.42	0.65	0.61	0.50	0.36	0.81
5 Marijilda	0.17	0.20	0.24	0.34	*****	0.17	0.13	0.21	0.24	0.74
6 Cave Cr	0.091	0.066	0.075	0.13	1.2	*****	0.044	0.34	0.36	0.83
7 Madera Cn	0.11	0.079	0.091	0.16	1.6	5.4	*****	0.29	0.32	0.82
8 Pine Cr	0.16	0.13	0.15	0.25	0.92	0.49	0.62	*****	0.08	0.75
9 Garcia Fls	0.23	0.26	0.31	0.44	0.81	0.46	0.54	2.9	*****	0.58
10 Cana. R	0.039	0.019	0.019	0.06	0.089	0.051	0.053	0.083	0.18	*****

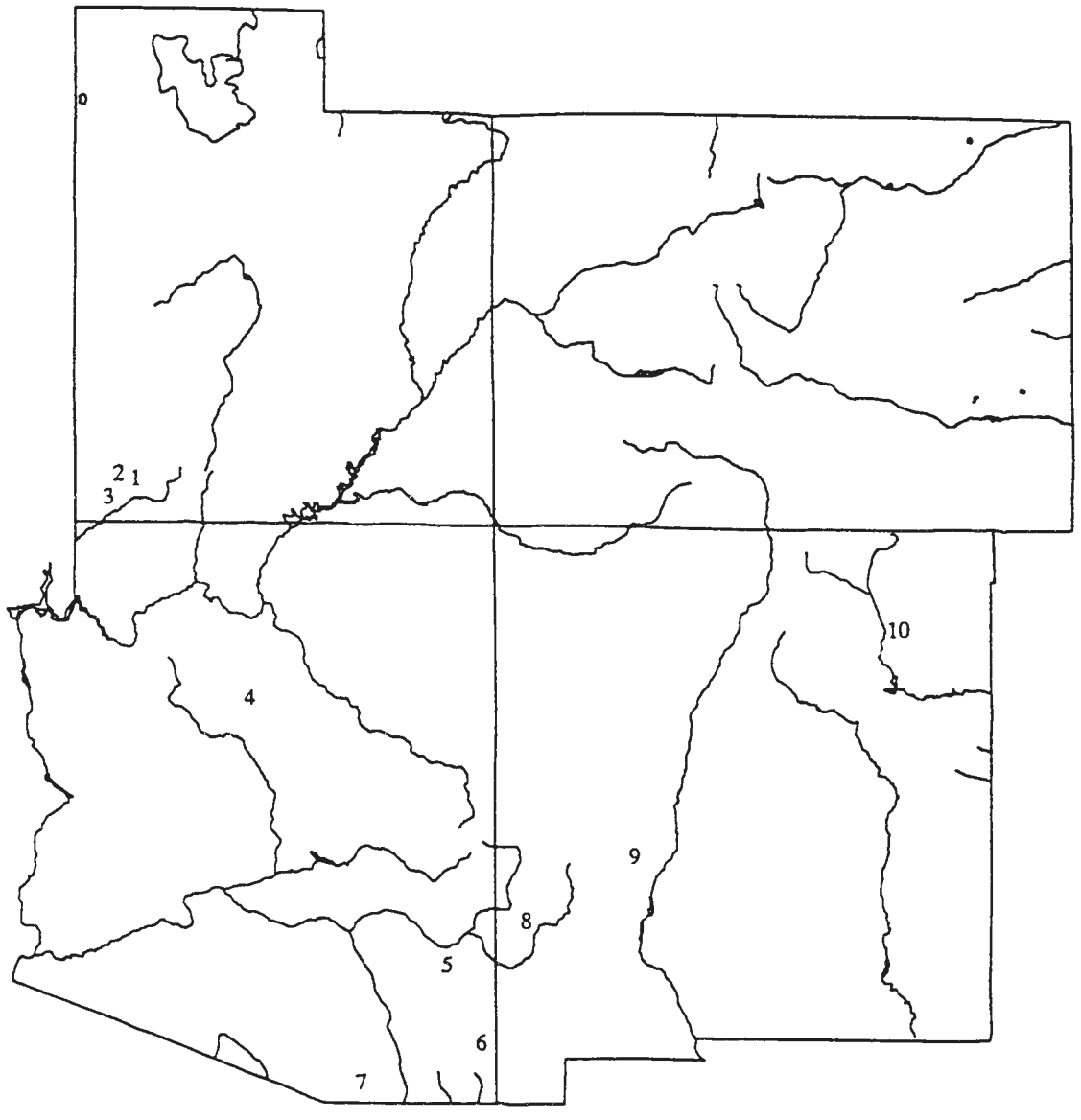

FIG. 2.1.–Sampled locations for *Hyla arenicolor*. Seven locations were sampled in Zion National Park (population 1) to assess genetic differentiation at the local scale.

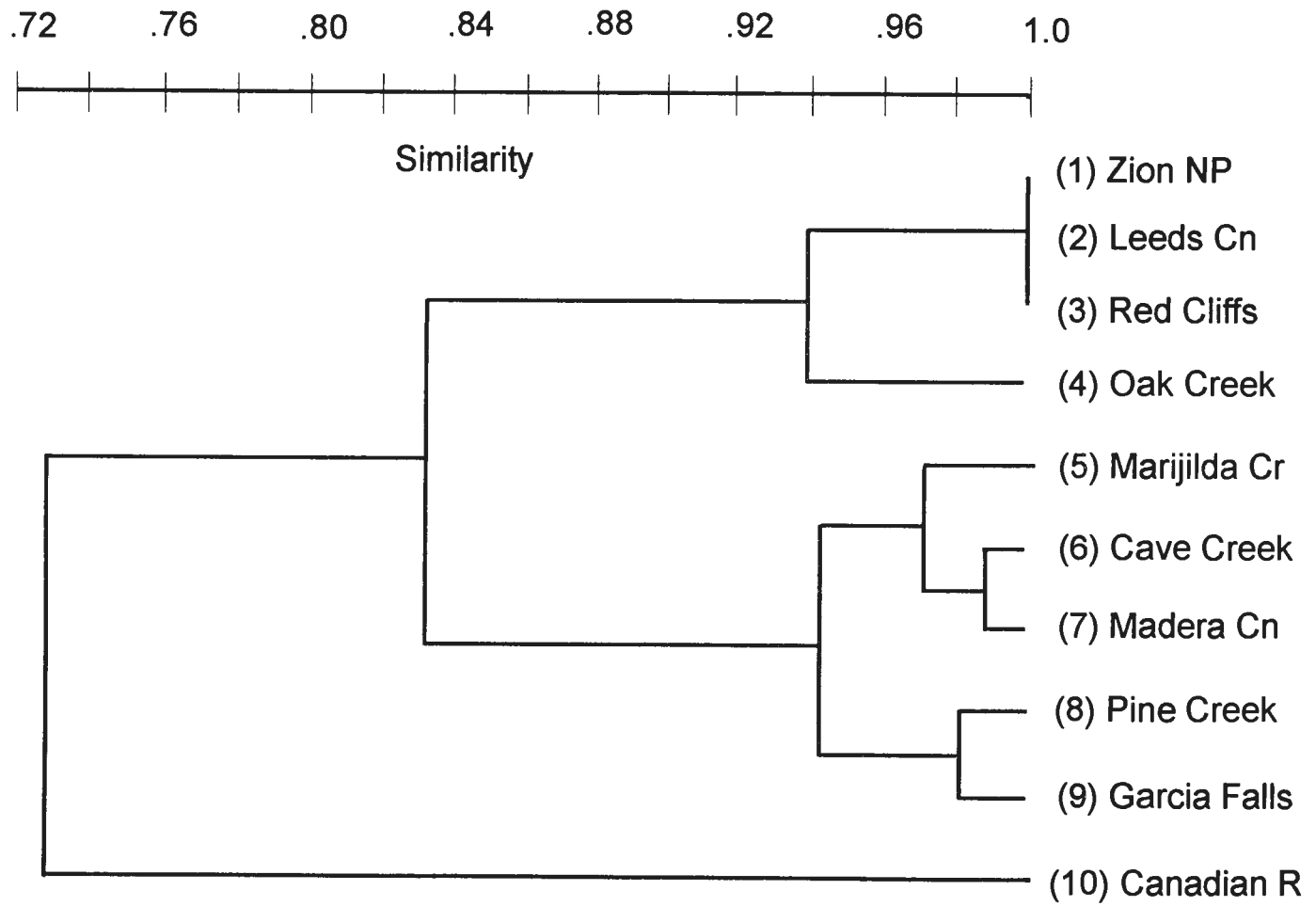

FIG. 2.2.--Phenogram based on cluster analysis using UPGMA. Coefficients used are Nei's (1978) genetic identity.

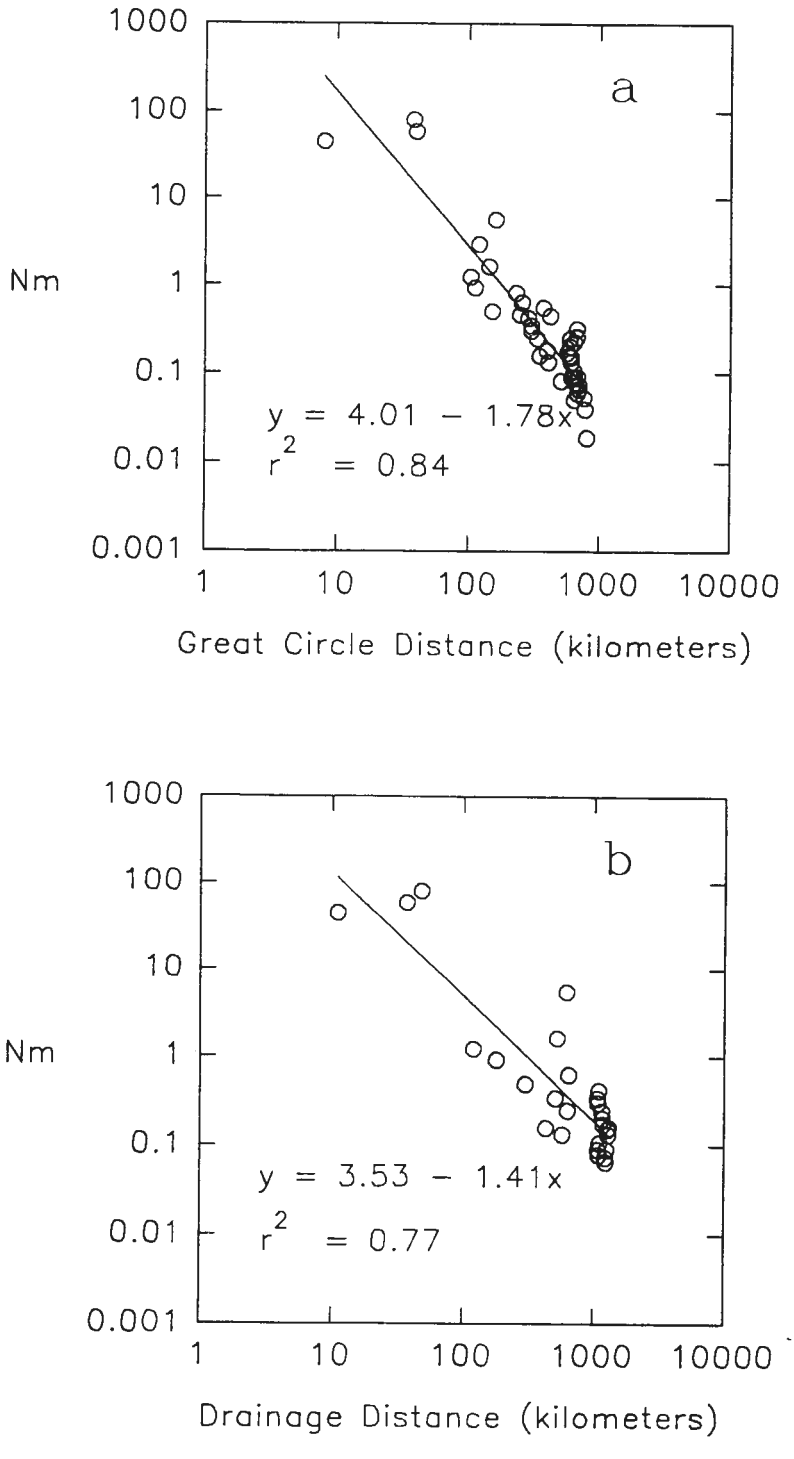

FIG. 2.3.--Regression of log(*Nm*) vs. log(distance). Top graph shows great circle distance between populations while bottom graph uses distance along current drainage corridors.

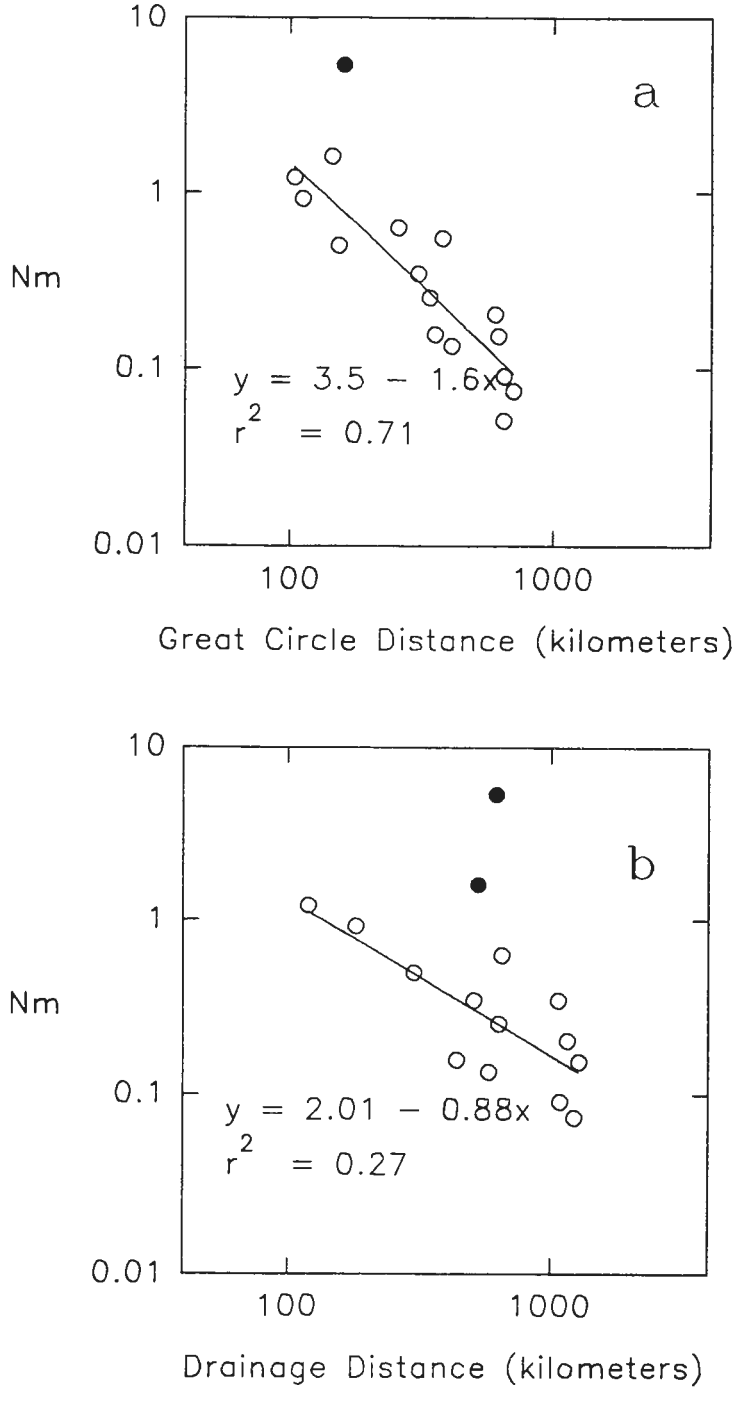

FIG. 2.4.--Regression of log(*Nm*) vs. log(distance). The three southwestern Utah populations were combined into a single population for these regressions. As in Fig. 2.3, top graph uses great circle distance and bottom graph uses drainage corridor distance. With outliers (solid dots) removed, data in (a) give $r^2 = 0.80$ and data in (b) give $r^2 = 0.61$.

CHAPTER 3

PHYLOGENY AND BIOGEOGRAPHY OF CANYON TREEFROGS (*HYLA ARENICOLOR*) BASED ON MITOCHONDRIAL DNA SEQUENCES

ABSTRACT

Intraspecific phylogenies play an important role in our understanding of evolutionary processes occurring among disjunct populations. Phylogenies also provide valuable information for constructing biogeographic hypotheses. Mitochondrial DNA sequences from portions of the *cyt*b, *NADH*-2, and *CO*-I genes were used to estimate the phylogeny of populations of Canyon Treefrogs (*Hyla arenicolor*) in the southwestern United States. The species is wide ranging yet limited to isolated pockets of moisture in this desert region. Disjunct populations have probably been isolated since the close of the Tertiary period. DNA results suggest the influence of historical drainage patterns in establishing population structure; genetic distances among populations appear to reflect late Tertiary and Pleistocene drainage connections and are not highly correlated with geographic distance. This result is in contrast to an earlier isozyme study that showed a strong correlation between genetic distance and geographic distance for *H. arenicolor*.

INTRODUCTION

The geographic subdivision of species into isolated populations has important evolutionary consequences. Until recently, however, investigating relationships among isolated populations has been difficult due to the limited amount of morphological variation typically observable at the intraspecific level. Isozyme electrophoresis techniques provided the first biochemical look at genetic differentiation among populations, but the nature of isozyme data limits their utility for cladistic analysis of intraspecific relationships. More recently, direct DNA sequencing and RFLP techniques have become the tools of choice for examining phylogenetic relationships among subdivided populations (Avise *et al.*, 1987). The mitochondrial genome in animals has proven to be particularly useful in this regard due to its maternal inheritance, non-recombining transmission, and rapid rate of evolution (Brown, 1983; Avise *et al.*, 1987; Avise, 1989).

Intraspecific phylogenies are valuable because they provide a genealogical context in which to view the genetic changes that have occurred among disjunct populations (Harrison, 1991). However, thus far they have been used more extensively for investigating the biogeographic factors that give rise to allopatric distributions (e.g., Bermingham and Avise, 1986; Lamb *et al.*, 1989; Hedges *et al.*, 1991). As with earlier studies of biotic distributions at higher taxonomic levels, intraspecific phylogenies are providing insights into

the geologic processes that have contributed to divergence and speciation. Given sufficient historical limitations on dispersal and gene flow, the spatial distribution of mtDNA haplotypes is expected to be restricted, resulting in a correspondence between hypothesized clades and geographic areas of endemism (Avise *et al.*, 1987). Concordant areas of endemism across several taxonomic groups are taken as evidence of extrinsic geologic factors restricting the movement of individuals and gametes. Intraspecific studies necessarily focus on a shorter time scale than investigations above the species level and are thus pertinent for investigating more recent biogeographical events. In particular, many intraspecific studies have focused on geologic events in North America during the Quaternary, an influential period in establishing present-day species distributions (e.g., Blair, 1965; Davis, 1983).

In explaining disjunct distributions, biogeographers subscribe to one of two hypotheses: vicariance or dispersal. Under a vicariance scenario, a formerly widespread and continuous distribution is subsequently fragmented by the formation of physical barriers to individual movements. Under a dispersal hypothesis, the barriers to dispersal exist prior to the disjunct distribution but are circumvented by a founding group that establishes a population in a previously uncolonized area. Discerning between these two explanations based on biological data alone can be problematic (Platnick and Nelson, 1978) because patterns of divergence are often consistent with either scenario. An assessment of both geologic evidence and biotic patterns is usually necessary

in order to postulate a coherent picture of historical events.

In this study I use mtDNA sequences to examine the phylogeny of populations of Canyon Treefrogs (*Hyla arenicolor* Cope). This species is found in isolated pockets of riparian habitat throughout the desert southwest region of the United States and south to central Mexico. Physiological (Preest *et al.*, 1992; Snyder and Hammerson, 1993) and behavioral (Wylie, 1981) aspects of *H. arenicolor* biology suggest that the species has low potential for dispersal (except possibly passive dispersal), which is typical for anurans. Additional limits to dispersal by *H. arenicolor* are due to surrounding desert environments, the often extreme topography of their habitat, and the large distances separating pockets of suitable habitat. Many of the drainages that support populations have no present-day connections or are ephemeral in their connections. These circumstances present an interesting problem in biogeography: How does a moisture dependent, poorly dispersing species attain a wide distribution throughout a region dominated by desert environments?

My goal in the present study is to (1) provide data on mtDNA variation in *H. arenicolor*, (2) estimate the phylogenetic relationship of mtDNA haplotypes for the species, (3) relate patterns of divergence to theories of geologic change in the region, and (4) postulate a biogeographic scenario to explain the species' current distribution.

MATERIALS AND METHODS

Hyla arenicolor was sampled in the larval stage at 10 sites in Utah, Arizona, and New Mexico (Fig. 3.1; Table 3.1). Samples of *Hyla versicolor* from northern Alabama served as an outgroup for rooting the mtDNA phylogenies. Samples were transferred to the laboratory where they were raised in aquaria to a pre-metamorphosis stage. Individuals were then identified to species and euthanized in a 2% solution of MS-222. Total genomic DNA extractions were performed using phenol-chloroform, ethanol precipitation procedures (Hillis *et al.*, 1990). Isolated DNA was resuspended in TE buffer for subsequent analysis.

The polymerase chain reaction (PCR) was used to amplify two regions of the mitochondrial genome, a ~600 bp region of the cytochrome b (*cyt*b) gene and a ~2000 bp region encompassing genes encoding the NADH-2 (*ND*-2) subunit, the cytochrome C oxidase-I (*CO*-I) subunit, and four tRNAs. Primers for PCR reactions were: *cyt*b 5'-TGAGGACAAATATCATTCTGAG-GGGCTGCAG-3' and 5'-TCTTCTACTGGTTGTCCTCCGATTCA -3' (modified from Gerhardt *et al.*, 1994); *ND*-2/*CO*-I 5'-TAAGCTATCGGGCCCATACC-3' (L-3880) and 5'-ACTTCAGGGTGCCCA-AAGAATCA-3' (H-6033; Riddle *et al.*, 1993). Amplified products were purified using a polyethylene glycol precipitation procedure (Morgan and Soltis, 1993) and sequenced using the dideoxy chain termination method (Sanger *et al.*, 1977). The amplified region

of the *cyt*b gene was sequenced using the PCR primers specified above. Primer L-3880 was used in sequencing reactions to obtain an ~400 bp region of the *ND*-2 subunit. Primer H-6033 and an additional internal primer (5'-TAAACTGTTCATCCC-GTCCC-3'; H-1850) were used to obtain sequences for a ~700 bp region of the *CO*-I subunit. Primer H-6033 was ineffective on samples taken from Madera Canyon and Marijilda Creek. Sequence information for the *CO*-I region at these localities was obtained for the H-1850 primer only (336 bp). Similarly, primer H-1850 was ineffective on *H. versicolor* samples. Primer H-6033 was used to obtain sequences for these individuals (533 bp). Three fragments from each of two frogs per population were sequenced.

Sequences were aligned using Sequencher™ software; the sequences were of sufficient similarity (including the outgroup) that no difficulties were encountered in the alignment procedure. Each of the three fragments were analyzed separately and then again as a single combined sequence. Combining sequence data from different regions of the genome is appropriate if trees generated in separate analyses do not differ significantly from one another (Huelsenbeck *et al.*, 1996). Phylogenetic trees were generated by the method of maximum parismony (PAUP v3.0; Swofford, 1990) using an exhaustive search to find the shortest trees, and a bootstrap search to assess levels of confidence for branches. All sites were weighted equally. g_1 statistics were computed for the distribution of tree lengths to assess phylogenetic signal

(Hillis and Huelsenbeck, 1992). The combined sequence data were also analyzed using a maximum likelihood tree generating algorithm (fastDNAml; Felsenstein, 1981; Olsen *et al.*, 1994) for comparison with the maximum parsimony phylogeny. Pairwise differences, transition/transversion ratios, and number of variable sites were determined using custom software.

RESULTS

Aligned sequences for each region of DNA are shown in Figs 3.2-3.4. For the *cyt*b region there were 569 aligned sites, 91 of which were variable within *H. arenicolor*, with 72 of these informative under the rules of parsimony. The *ND*-2 region was aligned at 393 sites, 55 of which were variable with 36 informative sites. The *CO*-I region was aligned at 672 sites, 65 of which were polymorphic with 42 informative sites. No within-population variation was detected for any region.

Pairwise differences between populations for the combined sequence data (Table 3.2) and percent transitions versus percent sequence divergence plots (Fig. 3.5) suggest three main geographic clades of *H. arenicolor* within the U.S. (Fig. 3.1; Fig. 3.6): a southwestern Utah/central Arizona group; a New Mexico/ southeastern Arizona group; and an exceptionally divergent southern Arizona group. With the exception of the southern Arizona group, pairwise distances between populations range from 0.4% to 3.2% (combined data). The southern Arizona group consists of two populations, Marijilda Creek in the Mt.

Graham area and Madera Canyon near Nogales, Arizona. Levels of sequence divergence between these two populations and the other eight populations are high, ranging from 10-12%.

Phylogenetic trees produced by the method of maximum parsimony are shown in Fig. 3.6. Trees for each segment are in general agreement, with slightly higher resolution provided by the longer *cyt*b and *CO*-I fragments. g_1 statistics for the distribution of tree lengths (Fig. 3.6) show a left skew that is significantly different from that expected for random data ($p < 0.01$ for each region separately and for combined data; Hillis & Huelsenbeck, 1992), an indication of phylogenetically informative differences in the sequence data. The maximum likelihood tree generated from the combined data (Fig. 3.7) is in close agreement with maximum parsimony trees as well. Some additional resolution for New Mexico/southeastern Arizona populations is suggested by the maximum likelihood phylogeny, which appears to show evidence for a south-to-north range expansion in this geographic region. However, following the approach of Kishino and Hasegawa (1989), I tested the significance of this pattern against an alternative north-to-south arrangement of populations and found no significant difference between the trees. Thus, the consensus maximum likelihood tree is identical to that for maximum parsimony with a polytomy for five populations in Arizona and New Mexico.

DISCUSSION

The accuracy of phylogenetic analysis based on maximum-parsimony is contingent on limiting the number of homoplasies (convergent or parallel changes) in the data set. For sequence data, this means that regions of DNA included in the analysis must be evolving at an appropriate rate for the taxa involved. The rapid rate of evolution in animal mtDNA, estimated at 2% sequence divergence per million years for several vertebrate groups (Brown *et al.*, 1979; Wilson *et al.*, 1985; Shields and Wilson, 1987), can be ideal for intraspecific comparisons. However, the linear relationship between percent sequence divergence and divergence time breaks down beyond ~15-20% sequence divergence (Brown *et al.*, 1979), presumably as sites available for silent mutational changes become saturated (Brown *et al.*, 1982). It has also been noted that transitions occur more frequently than transversions during the initial divergence period (Brown *et al.*, 1982; DeSalle *et al.*, 1987). Thus, the percent sequence divergence and the transition-to-transversion ratio can both be used to assess the appropriateness of a given segment of DNA for phylogenetic analysis.

In the current data, a transition bias is evident for most pairwise comparisons with the exception of a few at very low sequence divergence (Table 3.2; Fig. 3.5). This pattern of base substitution, with transversions sometimes exceeding transitions at low sequence divergence but rarely at

higher levels of divergence, is evident in other published sequence data as well (Brower, 1994). To determine if this pattern is likely to arise by chance, I constructed a computer model of sequence evolution based on multinomial probabilities, with transitions twice as likely as transversions. The 99% probability limits of the computer model were overlain on the combined sequence data (Fig. 3.5). A few data points fell slightly outside these limits but the data conformed to the model remarkably well, an indication that a more complicated explanation for the observed pattern of base substitution may be unnecessary. Since comparisons of *H. arenicolor* with the outgroup continue to show a strong transition bias (> 70%), multiple substitutions are probably not a concern for the current data set. Also, estimates of sequence divergence fall within the linear region of the divergence curve, with a maximum intraspecific divergence of ~12%. The *H. versicolor* outgroup exhibits a 16-17% sequence divergence from *H. arenicolor*, an indication that these regions of mtDNA could also be applied to interspecific studies within *Hyla*.

How might *H. arenicolor* have attained its present geographic distribution and what evidence is provided by the sequence data for distinguishing among the possibilities? Under a vicariance hypothesis, *H. arenicolor* may have been more or less continuously distributed throughout the Southwest region during a more mesic period (e.g., during the Pleistocene) and subsequently undergone range contraction as climatic conditions became hotter and drier. Alternatively, individuals may have dispersed along drainage

corridors, arriving at patches of favorable habitat in the process. Under the former scenario, if sufficient gene flow occurred among adjacent populations prior to and during range contraction, divergence at the molecular level during this period would be a function of distance between populations. Currently, levels of divergence among populations might still be a function of distance because isolation by distance would be the most recent factor creating a non-random (geographically structured) pattern of divergence. Under the dispersal model, observed molecular differentiation could reflect current drainage patterns or historical drainage patterns, but in either case a correlation based strictly on straight-line distance would not be expected. Populations that lie within a geographic region defined by current or historical drainage connections should be more similar at the molecular level than populations in separate and distinct drainage regions regardless of the distances between populations. Both hypotheses depend on some understanding of the climatic and geologic history of the Southwest region.

Early Tertiary sediments and fossils provide evidence for a western continent characterized by inland seas and warm climates (Axelrod, 1983). Mountain-building events later in the Tertiary uplifted both the Sierra Nevada Mountains and the Rocky Mountains, introducing a double rainshadow to the southwestern interior region that substantially reduced annual rainfall amounts. More recently, pluvial conditions that existed during the Pleistocene continued to support pine and spruce woodlands in the southern extent of the Basin and

Range province (Martin and Mehringer, 1965). Many intermountain basins in this region contained pluvial lakes as recently as 20,000 years b.p. (Kottlowski *et al.*, 1965). Thus, the extremes of temperature and aridity in the present-day desert Southwest are a relatively recent phenomenon. However, whether or not the region received enough moisture near the end of the Tertiary to support a continuous distribution of *H. arenicolor* remains unclear.

Other major geologic events influencing biotas in the region occurred in the Tertiary as well. During the Paleocene, the Colorado plateau region was lower than surrounding landforms, forming a basin for runoff from northern Arizona and southern Nevada (Lucchitta, 1979). Uplifting of the Colorado plateau began in the middle Tertiary and continued through the Pliocene period (Hunt, 1974). The current westward flow of the Colorado River may not have been initiated until the middle Pliocene (McKee and McKee, 1972; Young and Brennan, 1974). This period was also characterized by extensive volcanism and tectonism farther south in the Basin and Range province, causing major shifts in drainage patterns throughout the region (Hunt, 1974). There is evidence that the upper Gila River area in southeastern Arizona drained eastward toward the Rio Grande as recently as the Pliestocene (Kottlowski *et al.*, 1965). Also, the present-day course of the Rio Grande itself may not have been established until the mid-Pliestocene (Ruhe, 1962; Kottlowski et al., 1965). The pertinent conclusions from these lines of evidence are that drainage patterns throughout the region have been in a state of flux over

several million years.

Divergence data (Table 3.2) and phylogenetic trees (Figs. 3.6 & 3.7) are evidence for three main geographic groups of *H. arenicolor* in the United States: a Colorado Plateau group (three populations), an eastern Basin and Range/Rocky Mountain piedmont group (five populations), and a highly divergent southern Basin and Range group (two populations). Geographic distance between populations does not correlate well with genetic distance. A regression analysis of genetic distance (based on sequence divergence) against geographic distance gave an r^2 of 0.6 (excluding the highly divergent southern Arizona clade). Perhaps more revealing than this statistic, however, is the relationship between the Red Cliffs, Oak Creek, and Pine Creek populations. Oak Creek is ~302 km from Red Cliffs and ~335 km from Pine Creek, a difference of only 33 km. Under a model of isolation by distance and range contraction, these populations should show similar molecular distances as well. However, from Table 3.2 it can be seen that Oak Creek frogs are considerably more divergent from Pine Creek than from Red Cliffs. While present-day drainage patterns do not provide a close link between Oak Creek and Red Cliffs, historical drainage patterns do since both of these populations would have been associated with the Tertiary Colorado basin.

Other aspects of the proposed phylogeny also point to the influence of ancient drainage patterns. The historical eastward flow of the upper Gila River region would have linked the Pine Creek population with others in the Rio

Grande drainage (Garcia Falls, Tierra Blanca, and possibly Cave Creek), a connection that is supported by the current data. This eastern clade also points to the possibility of a historical connection between the Rio Grande, the Canadian River, and the San Simon River valley. Additionally, some geographically adjacent populations are more differentiated than other widely separated populations. Madera Canyon and Marijilda Creek populations in particular are highly divergent from all other groups yet are geographically close to Cave Creek, Pine Creek, and Tierra Blanca.

These results contrast a population-level study of *H. arenicolor* based on isozyme variation (Murray and Wolf, submitted). Gene flow estimates derived from variation at nine protein-coding loci were strongly correlated with great circle distances and less well correlated with drainage distances separating populations. Madera Canyon and Marijilda Creek populations are particulary notable in this regard; a UPGMA phenogram based on the isozyme data grouped these populations into a cluster with the Cave Creek population at a similarity level of 0.97 (Nei's [1978] unbiased genetic identity). Discordance between mitochondrial and chromosomal DNA variation can arise due to stochastic lineage sorting, unequal rates of evolution, or introgressive hybridization (Avise, 1989). Lineage sorting is an unlikely explanation for the current study because of the large genetic distances between the major clades. Ignoring for a moment the Madera Canyon and Marijilda Creek populations, differences between the isozyme data and the mtDNA data with regard to

isolation by distance versus drainage pattern associations could be a result of unequal rates of evolution between nuclear and mitochondrial genes. The establishment of new nuclear alleles in a population occurs at a slower rate than the corresponding process in mtDNA due to the accelerated rate of nucleotide substitution in mtDNA (Brown *et al.*, 1979) and the smaller effective population size of the mtDNA genome (Moore, 1995). Analyses of variation in nuclear and mtDNA genes are therefore focused on different time scales and may lead to differing biogeographical conclusions.

Given the level of mtDNA divergence in the Madera Canyon and Marijilda Creek populations, hybridization with other *Hyla* species is a more likely explanation for the origin of these haplotypes. The ability of *Hyla* species to hybridize has been demonstrated in the laboratory (Pierce, 1968, 1975) as well as in natural settings where they occur sympatrically (Lamb and Avise, 1986; Gerhardt *et al.*, 1994). Also, the capture of interspecific organellar DNA through hybridization while retaining a single species' nuclear genome has been reported in a number of taxa (Ferris *et al.*, 1983; Spolsky and Uzzell, 1984; Tegelstrom, 1987; Dowling and Hoeh, 1991; Lehman *et al.*, 1991; Soltis and Kuzoff, 1995). The range of *Hyla arenicolor* overlaps that of *Hyla eximia* in a few areas in southern Arizona and extensively in Mexico, providing opportunity for the transfer of mtDNA between these two species.

Can an estimate be placed on the times since these mtDNAs diverged from a common ancestral haplotype? Although the overall rate of 2%

sequence divergence per million years for mtDNA may be inappropriate for some taxa (see Avise, 1989; Lamb *et al.*, 1989; Hedges *et al.*, 1991), its application to the present study provides evidence that the Colorado Plateau group has been isolated from the eastern Basin and Range group for ~1.5 million years (combined data). This seems to be a reasonable estimate, especially in light of similar studies on other taxa within the region (Lamb *et al.*, 1989; Riddle and Honeycutt, 1990). Lamb et al. (1989) estimated a divergence time of 2-3 million years for populations of desert tortoise (*Xerobates agassizi*) on either side of the lower Colorado River. While the Colorado River is not implicated as a major zoogeographic barrier by the present study, the tortoise data demonstrate the levels of interpopulation divergence that can be expected for species with low vagility.

Comparing the highly divergent southern Arizona populations with other groups yields a divergence time of ~5 million years in the past. This estimate seems unlikely given the close proximity of neighboring populations, again pointing to the likelihood of an alternative process such as hybridization in the establishment of these haplotypes.

Tree branch lengths (Fig. 3.6) provide evidence that the Colorado Plateau group and the eastern Basin and Range group are equally divergent from a common ancestor. Unfortunately, the geographic location of an ancestral population remains uncertain given the likelihood of a hybrid origin for the southern Arizona haplotypes. Some evidence for a northward range

expansion is provided by the maximum likelihood phylogeny in the New Mexico clade; southern populations are ancestral to northern populations according to the tree in Fig. 3.7. The branching order shown in Fig. 3.7 is not strongly supported by the current data, however. Lack of resolution for the five populations making up the eastern clade suggests that these populations have had recent contact or that they were founded over a relatively short period of time from ancestral stock. A determination of whether or not northern populations are derived from southern ancestral stock will require additional data from populations south of the U.S. border.

In summary, the sequence data provide some evidence that populations associated with major drainages have been founded via dispersal. Subsequent vicariant events such as the uplift of the Colorado Plateau, separating southwestern Utah populations from central Arizona populations, also appear to have played a role in establishing present-day patterns of divergence. Pleistocene drainage patterns in the Rio Grande region may have led to rapid dispersal of *H. arenicolor* northward from southeastern Arizona to north/central New Mexico. There is evidence that vicariant events separating these populations into disjunct drainages (Gila River, Rio Grande, and Canadian River) occurred recently enough that significant divergence has not yet occurred among haplotypes. Support for these hypotheses could be provided through similar phylogenetic studies on species with ranges that overlap the range of *H. arenicolor*. Phylogenies for other species in the Basin and Range

and Great Basin Provinces are available (e.g., Lamb et al., 1989; Riddle and Honeycutt, 1990), but the ranges of these species are not sufficiently coincident with the range of *H. arenicolor* to assess concordant patterns of endemism. An informative comparison with the *H. arenicolor* phylogeny might be afforded by a study of sympatric anurans (*Bufo punctata*) or fish species inhabiting the same major drainages (e.g., *Rhinichthys osculus*).

REFERENCES

Avise, J. C. (1989). Gene trees and organismal histories: A phylogenetic approach to population biology. *Evolution* **43(6):** 1192-1208.

Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., and Saunders, N. C. (1987). Intraspecific phylogeography: The mitochondrial bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst.* **18:** 489-522.

Axelrod, D. I. (1983). Paleobotanical history of the western deserts. *In* "Origin and Evolution of Deserts" (S. G. Wells and D. R. Haragan, Eds.), pp. 113-129, University of New Mexico Press, Albuquerque.

Bermingham, E., and Avise, J. C. (1986). Molecular zoogeography of freshwater fishes in the southeastern United States. *Genetics* **113:** 939-965.

Blair, W. F. (1965). Amphibian speciation. *In* "The Quaternary of the United States" (H. E. Wright Jr. and D. G. Frey, Eds.), pp. 543-556, Princeton University Press, Princeton, NJ.

Brower, A. V. Z. (1994). Phylogeny of *Heliconius* butterflies inferred from mitochondrial DNA sequences (Lepidotera: Nymphalidae). *Mol. Phy. Evo.* **3(2):** 159-174.

Brown, W. M. (1983). Evolution of animal mitochondrial DNA. *In* "Evolution of Genes and Proteins" (M. Nei and R. K. Koehn, Eds.), pp. 62-88, Sinauer Associates, Inc., Sunderland, MA.

Brown, W. M., George, M., Jr., and Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *Proc. Nat. Acad. Sci.* **76:** 1967-1971.

Brown, W. M., Prager, E. M., Wang, A., and Wilson, A. C. (1982). Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *J. Mol. Evol.* **18:** 225-239.

Davis, M. B. (1983). Holocene vegetational history of the eastern United States. *In* "Late-Quaternary Environments of the United States" (H. E. Wright, Jr., Ed.), pp. 166-181, University of Minnesota Press, Minneapolis.

DeSalle, R., Freedman, T., Prager, E. M., and Wilson, A. C. (1987). Tempo and mode of sequence evolution in mitochondrial DNA of Hawaiian *Drosophila*. *J. Mol. Evol.* **26:** 157-164.

Dowling, T. E., and Hoeh, W. R. (1991). The extent of introgression outside the contact zone between *Notropis cornutus* and *Notropis chrysocephalus* (Teleostei: Cryprinidae). *Evolution* **45:** 944-956.

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

Ferris, S. D., Sage, R. D., Huang, C.-M., Nielson, J. T., Ritte, U., and Wilson, A. C. (1983). Flow of mitochondrial DNA across a species boundary. *Proc. Nat. Acad. Sci.* **80:** 2290-2294.

Gerhardt, H. C., Ptacek, M. B., Barnett, L., and Torke, K. G. (1994). Hybridization in the diploid-tetraploid treefrogs *Hyla chrysoscelis* and *Hyla versicolor*. *Copeia* **1994(1):** 51-59.

Harrison, R. G. (1991). Molecular changes at speciation. *Ann. Rev. Ecol. Syst.* **22:** 281-308.

Hedges, S. B., Bezy, R. L., and Maxson, L. (1991). Phylogenetic relationships and biogeography of Xantusiid lizards inferred from mitochondrial DNA sequences. *Mol. Biol. Evol.* **8:** 767-780.

Hillis, D. M., Larson, A., Davis, S. K., and Zimmer, E. A. (1990). Nucleic acids III: Sequencing. In "Molecular Systematics" (D. M. Hillis and C. Moritz, Eds.), pp. 318-370, Sinauer Associates, Inc., Sunderland, MA.

Hillis, D. M., and Huelsenbeck, J. P. (1992). Signal, noise, and reliability in molecular phylogenetic analyses. *J. Hered.* **83(3):** 189-195.

Huelsenbeck, J. P., Bull, J. J., and Cunningham, C. W. (1996). Combining data in phylogenetic analysis. *Trend. Ecol. Evol.* **11(4):** 152-158.

Hunt, C. B. (1974). "Natural Regions of the United States and Canada," W. H. Freeman, San Francisco, CA.

Kishino, H., and Hasegawa, M. (1989). Evaluation of the maximum likelihood estimate of evolutionary tree topologies from DNA sequence data, and the

branching order in Hominoidea. *J. Mol. Evol.* **29:** 170-179.

Kottlowski, F. E., Cooley, M. E., and Ruhe, R. V. (1965). Quaternary geology of the Southwest. *In* "The Quaternary of the United States" (H. E. Wright Jr. and D. G. Frey, Eds.), pp. 287-298, Princeton University Press, Princeton, NJ.

Lamb, T., and Avise, J. C. (1986). Directional introgression of mitochondrial DNA in a hybrid population of tree frogs: The influence of mating behavior. *Proc. Nat. Acad. Sci.* **83:** 2526-2530.

Lamb, T., Avise, J. C., and Gibbons, J. W. (1989). Phylogeographic patterns in mitochodrial DNA of the desert tortoise (*Xerobates agassizi*), and evolutionary relationships among the North American gopher tortoises. *Evolution* **43:** 76-87.

Lehman, N., Eisenhawer, A., Hansen, K., Mech, L. D., Peterson, R. O., Gogan, P. J. P., Wayne, R. K. (1991). Introgression of coyote mitochondrial DNA into sympatric North American gray wolf populations. *Evolution* **45:** 104-119.

Lucchitta, I. (1979). Late Cenozoic uplift of the southwestern Colorado Plateau and adjacent lower Colorado River region. *Tectonophys.* **61:** 63-95.

Martin, P. S., and Mehringer, P. J., Jr. (1965). Pleistocene pollen analysis and biogeography of the Southwest. *In* "The Quaternary of the United States" (H. E. Wright Jr. and D. G. Frey, Eds.), pp. 433-451, Princeton University Press, Princeton, NJ.

McKee, E. D., and McKee, E. H. (1972). Pliocene uplift of the Grand Canyon

region--time of drainage adjustment. *Geol. Soc. Am. Bull.* **83:** 1923-1932.

Moore, W. S. (1995). Inferring phylogenies from mtDNA variation: Mitochondrial-gene trees versus nuclear-gene trees. *Evolution* **49:** 718-726.

Morgan, D. R., and Soltis, D. E. (1993). Phylogenetic relationships among members of Saxifragaceae sensu lato based on *rbcL* sequence data. *Ann. Missouri Bot. Gard.* **80:** 631-660.

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89:** 583-590.

Olsen, G. J., Matsuda, H., Hagstrom, R., and Overbeek, R. (1994). fastDNAml: A tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. *Comp. Appl. Biosci.* **10:** 41-48.

Pierce, J. R., 1968. "Isolation and Differentiation in the Canyon Treefrog, *Hyla arenicolor,*" Ph.D. thesis, The University of Texas.

Pierce, J. R. (1975). Genetic compatibility of *Hyla arenicolor* with other species in the family, Hylidae. *Tex. J. Sci.* **16(3-4):** 431-441.

Platnick, N. I., and Nelson, G. (1978). A method of analysis for historical biogeography. *Syst. Zool.* **27:** 1-16.

Preest, M. R., Brust, D. G., and Wygoda, M. L. (1992). Cutaneous water loss and the effects of temperature and hydration state on aerobic metabolism of canyon treefrogs, *Hyla arenicolor. Herpetologica* **48:** 210-219.

Riddle, B. R., and Honeycutt, R. L. (1990). Historical biogeography in North American arid regions: An approach using mitochondrial-DNA phylogeny in

grasshopper mice (genus *Onychomys*). *Evolution* **44:** 1-14.

Riddle, B. R., Honeycutt, R. L., and Lee, P. L. (1993). Mitochondrial DNA phylogeography in northern grasshopper mice (*Onychomys leucogaster*)–the influence of Quaternary climatic oscillations on population dispersion and divergence. *Mol. Ecol.* **2:** 183-193.

Ruhe, R. V. (1962). Age of the Rio Grande valley in southern New Mexico. *J. Geol.* **70:** 151-167.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Nat. Acad. Sci.* **74:** 5463-5467.

Shields, G. F., and Wilson, A. C. (1987). Calibration of mitochondrial DNA evolution in geese. *J. Mol. Evol.* **24:** 212-217.

Snyder, G. K., and Hammerson, G. A. (1993). Interrelationships between water economy and thermoregulation in the Canyon tree-frog *Hyla arenicolor*. *J. Arid Environ.* **25:** 321-329.

Soltis, D. E., and Kuzoff, R. K. (1995). Discordance between nuclear and chloroplast phylogenies in the *Heuchera* group (Saxifragaceae). *Evolution* **49:** 727-742.

Spolsky, C., and Uzzell, T. (1984). Natural interspecies transfer of mitochondrial DNA in amphibians. *Proc. Nat. Acad. Sci.* **81:** 5802-5805.

Swofford, D. L. (1990). "PAUP: Phylogenetic Analysis Using Parsimony," Version 3.0, Illinois Natural History Survey, Champaign, IL.

Tegelstrom, H. (1987). Transfer of mitochondrial DNA from the northern red-

backed vole (*Clethrionomys rutilus*) to the bank vole (*C. glareolus*). J. Mol. Evol. **24:** 218-227.

Wilson, A. C., Cann, L. C., Carr, S. M., George, M., Gyllensten, U. B., Helm-Bychowski, K. M., Higuchi, R. G., Palumbi, S. R., Prage, E. M., Sage, R. D., and Stoneking, M. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc.* **26:** 375-400.

Wylie, S. R., 1981. "Effects of Basking on the Biology of the Canyon Treefrog, *Hyla arenicolor* Cope," Ph.D. thesis, Arizona State University.

Young, R. A., and Brennan, W. J. (1974). Peach Springs Tuff: Its bearing on structural evolution of the Colorado Plateau and development of Cenozoic drainage in Mohave County, Arizona. *Geol. Soc. Am. Bull.* **85:** 83-90.

TABLE 3.1

Collection Sites for *Hyla Arenicolor*

Population	Locality	
1. Zion NP, UT	37° 16'N	112° 58'W
2. Red Cliffs, UT	37° 14'N	113° 25'W
3. Oak Creek, AZ	34° 52'N	111° 46'W
4. Canadian River, NM	36° 01'N	104° 17'W
5. Garcia Falls, NM	33° 30'N	107° 27'W
6. Tierra Blanca, NM	32° 52'N	107° 35'W
7. Pine Creek, NM	33° 12'N	108° 44'W
8. Marijilda Creek, AZ	32° 42'N	109° 48'W
9. Cave Creek, AZ	31° 52'N	109° 13'W
10. Madera Canyon, AZ	31° 44'N	110° 53'W

TABLE 3.2

Percent Sequence Divergence (Above Diagonal) and Transitions/Transversions (Below Diagonal) for Combined Fragments

Pop.	1	2	3	4	5	6	7	8	9	10	11
1 Zion	*****	0.8	1.3	2.9	2.6	2.8	3.1	11	2.9	11	17
2 RedC	8/5	*****	1.0	3.2	2.8	2.7	3.0	12	2.8	11	17
3 OakC	14/7	15/2	*****	3.2	2.8	2.8	3.1	11	2.8	10	16
4 Cana	36/11	39/14	39/14	*****	0.7	1.0	1.2	12	0.9	11	17
5 Garc	34/9	37/8	38/8	4/7	*****	0.4	0.5	11	0.4	10	16
6 Tier	36/10	37/7	27/7	8/9	4/2	*****	0.6	11	0.5	10	16
7 Pine	39/11	40/8	41/8	9/10	5/2	7/4	*****	11	0.6	11	16
8 Mari	111/37	114/36	106/35	110/41	109/34	112/33	112/33	*****	11	2.2	17
9 Cave	35/12	36/9	37/9	5/9	3/3	5/3	6/3	109/33	*****	10	16
10 Made	111/32	113/29	105/28	109/34	108/27	110/26	110/26	16/12	107/26	*****	16
11 H.v.	184/65	185/62	182/62	183/66	186/59	186/57	186/59	148/54	186/57	141/48	*****

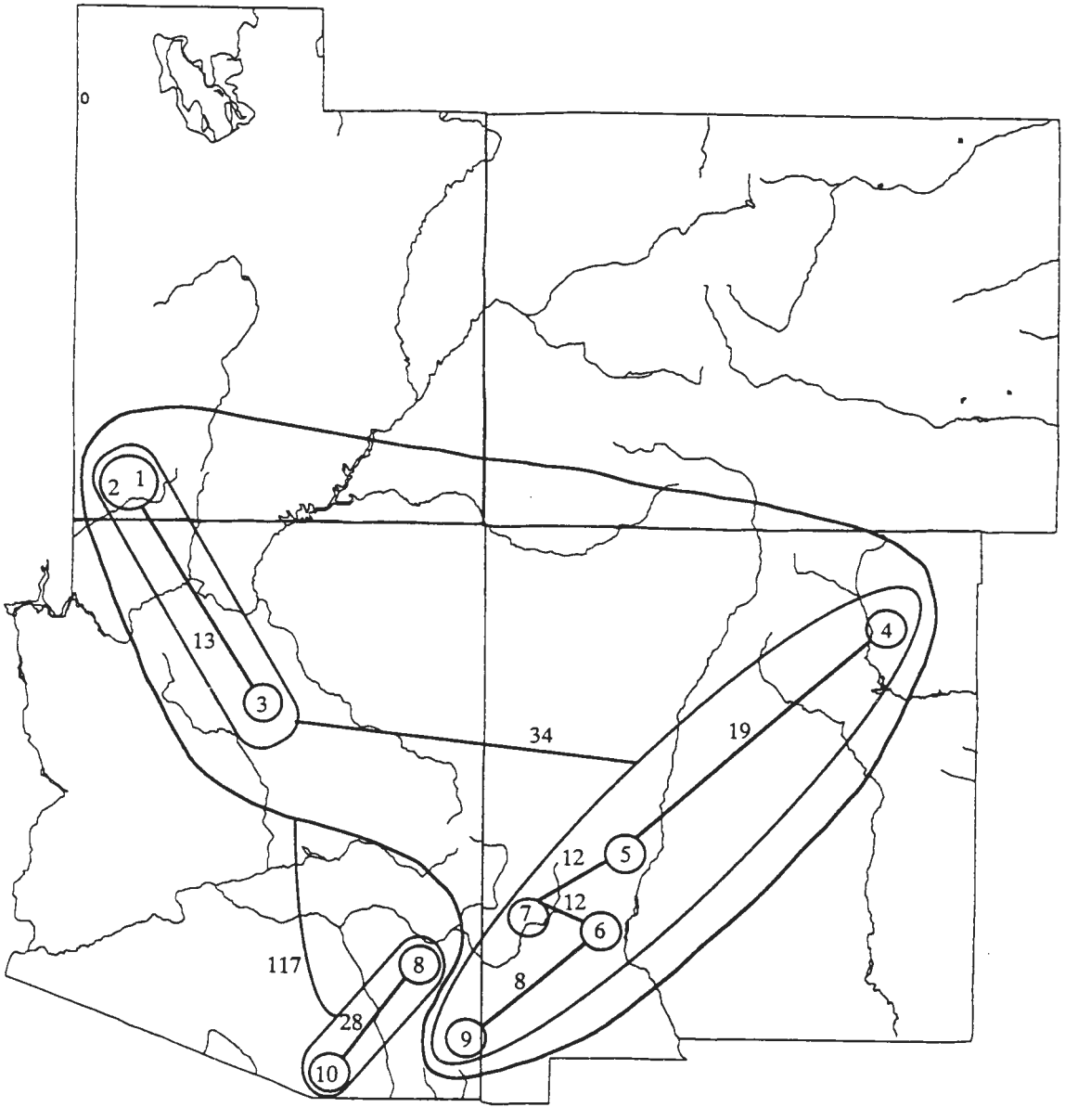

FIG. 3.1. Collection sites and geographic phylogeny for *Hyla arenicolor*. Lines connecting populations are labeled with the number of nucleotide substitutions along respective branches in the phylogeny based on combined sequence data.

```
Zion    CT AAC CTC CTC TCA GCC GGC CCC TAT ATC GGG ACC GAA CTA GGG GAA TGA ATT
RedC    .. ... ... ... ... ... .C. ... ... ... ... ... ... ... .TA C.. ... ...
OakC    .. ... ... ... ... ... .C. ... ... ... ... ... ... ... .TA C.. ... ...
Cana    .C ... ... ... ... ... .C. ... ... ... ... ... ... ... .T. ... ... ...
Garc    .C ... ... ... ... ... .C. ... ... ... ... ... ... ... .T. ... ... ...
Tier    .C ... ... ... ... ... .C. ... ... ... ... ... ... ... .TA C.. ... ...
Pine    .C ... ... ... ... ... .C. ... ... ... ... ... ... ... .TA C.. ... ...
Mari    .C ..G ... .G. ... ..G .C. ..G ... G.T ..A ..T ... ... ... C.G ... G..
Cave    .C ... ... ... ... ... .C. ... ... ... ... ... ... ... .TA C.. ... ...
Made    TC ... ... ... ... ... .C. ..A ... G.T ..A ..T ... ... .TT C.G ... ...
H.v.    TC ... ..T ..T ... ..T .C. ..A ... ..T ..A ..T ... ..T .TT C.. ... ...

Zion    TGA GGG GGC TTT TCA GTA GAT AAC GCT ACA CTA ACC CGA TTC TTC ACA TTT CAC
RedC    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC    ... ... ... ... ... ... ... ... ..C ... ... ... ... ... ... ... ... ...
Cana    ... ... ... ... ... C.. ... ... ..C ... ... ... ... ..T ... ... ... ...
Garc    ... ... ... ... ... ... ... ... ..C ... ... ... ... ..T ... ... ... ...
Tier    ... ... ... ... ... ... ... ... ..C ... ... ... ... ..T ... ... ... ...
Pine    ... ... ... ... ... ... ... ... ..C ... ... ... ... ..T ... ... ... ...
Mari    ... ... ... ... ... ... ..C ... ..C ... T.G ... ... ..T ..T ... ... ...
Cave    ... ... ... ... ... ... ... ... ..C ... ... ... ... ..T ... ... ... ...
Made    ... ... ... ... ... ... ..C ... ..C ... T.. ... ... ..T ..T ... ... ...
H.v.    ... ... ... ... ... ..T ..C ..T ... ... T.. ... ... ..T CA. ... ..C ...

Zion    TTC ATT CTC CCG TTC ATC ATC GCA GGC GCT TCA ATG ATT CAC CTT CTG TTC CTT
RedC    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Garc    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Tier    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Pine    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Mari    ..T ... ... ... ... ..T ... ... ... ... ... ... ... ... ..A ..T ...
Cave    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Made    ..T ... ... ... ... ..T ... ... ... ... ... ... ... ... ..A ..T ...
H.v.    ..T ... T.A ..T ... ..T ..T ... ..A ..A ... ..A G.C ... ..C ..C ... ...

Zion    CAT CAA ACA GGG TCA TCA AAC CCA ACA GGA TTA AAC TCT AAC CCA GAC AAA ATC
RedC    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Garc    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Tier    ... ... ... ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ...
Pine    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Mari    ..C ... ... ... ... ... ... ... ... ... ... ... ..C ... ... ..T ... ...
Cave    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Made    ..C ... ... ..A ... ... ... ... ... ... ... ... ..C ... ... ..T ... ...
H.v.    ... ..G ... ..A ..G ..T ?.. ... ... ..? ... ..T ... ..T ..T ... ... ...

Zion    CCC TTC CAC GCC TAT TAT TCG TAT AAA GAC GCA TTC GGC TTT GCG CTT CTT CTA
RedC    ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC    ... ... ... ... ... ... ... ... ..C ... ... ... ... ... ... ... ... ...
Cana    ... ... ... ... ... ..C ... ... ... ... ... ... ... ..T ... ... ... ...
Garc    ... ... ... ... ... ..C ... ... ... ... ... ... ... ..T ... ... ... ...
Tier    ... ... ... ... ... ..C ... ... ... ... ... ... ... ..T ... ... ... ...
Pine    ... ... ... ... ... ..C ... ... ... ... ... ... ... ..T ... ... ... ...
Mari    ... ... ... ... ..T ... ..C ..A ..C ... ... ... ... ..T ... ..C ..A ... ..C ...
Cave    ... ... ... ... ... ..C ... ... ... ... ... ... ... ..T ... ... ... ...
Made    ... ... ... ... ..T ... ..C ..A ..C ... ... ... ... ..T ... ..C ..A ... ..C ...
H.v.    ... ... ... ... ... .GC ..C ... ... ... ... ... ... ..T ... ..C ..A ... ..C ...
```

FIG. 3.2. mtDNA sequence of portion of *cyt*b gene in populations of *Hyla arenicolor* (corresponding to positions 16691-17256 in *Xenopus laevis*; ref). Dot (.) denotes identity with the first sequence; dash (-) denotes a gap; ? denotes an ambiguity.

```
Zion   GCC CTA CTA GCT GCC CTA TCC ACC TTC GCC CCC AAC ATC CTA GGA GAC CCT GAT
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..C ...
Cana   ... ... ... ... ... ... ... ... ..T ... ..T ... ... ... ..G ... ... ...
Garc   ... ... ... ... ... ... ... ... ..T ... ..T ... ... ... ..G ... ... ...
Tier   ... ... ... ... ... ... ... ... ..T ... ..T ... ... ... ..G ... ... ...
Pine   ... ... ... ... ... ... ... ... ..T ... ..T ... ... ... ..G ... ... ...
Mari   ... ..G ... ... ... T.. ... ... ..T ... ... ... ..T ..G ... ... ..C ...
Cave   ... ... ... ... ... ... ... ... ..T ... ..T ... ... ... ..G ... ... ...
Made   ... ..G ... ... ... T.. ... ... ..T ..T ... ... ..T ..G ... ... ..C ...
H.v.   ..A ..G ... ..A ..T ... ... ... ... ... ... ?.. ... ..G ... ... .?A ...

Zion   AAC TTC ACC CCA GCT AAC CCG CTA GTA ACC CCC CCG CAT ATT AAA CCT GAG TGA
RedC   ... ... ... ... ... ... ... G.. ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... G.. ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Garc   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Tier   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Pine   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Mari   ..T ..T ... ... ..C ... ..A ... ... ... ..T ... ... ... ..G ... ..A ...
Cave   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Made   ..T ... ... ... ..C ... ..A ... ... ... ..T ... ... ... ..G ... ..A ...
H.v.   ... ... ... ..C ..C ... ..A T.. ... ..T ... ..C ..C .C. ... ... ..A ...

Zion   TAC TTT CTA TTC GCA TAC GGT ATC CTT CGC TCC ATC CCA AAT AAA CTA GGG GGG
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ..T ... ..G ..T ... ... ... ... ... ... ... ... ..G ... ... G.. ... ...
Garc   ..T ... ..G ..T ... ... ... ... ... ... ... ... ..G ... ... ?.. ... ...
Tier   ..T ... ..G ..T ... ... ... ... ... ... ... ... ..G ... ... ... ... ...
Pine   ..T ... ..G ..T ... ... ... ... ... ... ... ... ..G ... ... ... ... ...
Mari   ... ... ... ..T ... ... .CA ..T ... ... ..A ... ..G ... ... ... ... ...
Cave   ..T ... ..G ..T ... ... .C. ... ... ... ... ... ..G ... ... ... ... ...
Made   ... ... ... ..T ... ... .CA ..T ... ... ..A ... ..G ... ... ... ... ...
H.v.   ... ... ... ..T ..T ... .CA ..T ... ... ... ..T ... ... ... ..G ..A ..A

Zion   GTC CTC GCC CTT CTT TTC TCG ATC ATG ATT GTA TTT CTC ATG CCC ATC CTT CAC
RedC   ... ... ... ... ..C ... ... ... ... ... ..G ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ..T ... T.. ... ... ... ... ...
Cana   ... ... ... ... ... ... ... ... ... ... C.T ..C ... ..A ... ... ... ...
Garc   ... ... ... ... ... ... ... ... ... ... C.T ..C ... ..A ... ... ... ...
Tier   ... ... ... ... ... ... ... ... ... ... C.T ... ... ..A ... ... ... ...
Pine   ... ... ... ... ... ... ... ... ... ... C.T ..C ... ..A ... ... ... ...
Mari   ... ... ... ..C ... ... ..C ... ..A ... C.T ..C ..T ... ... ..T ... ...
Cave   ... ... ... ... ... ... ... ... ... ... C.T ..C ... ..A ... ... ... ...
Made   ... ... ... ..C ... ... ..C ... ..A ... C.T ..C ..T ... ... ..T ... ...
H.v.   ... G.. ... ... ..A ..T ..C ..T ..A ..C C.C ..C ..A ..A ... ... ... ...

Zion   ACG TCT AGC CAA CGA ACT ACT GCC TTT CGC CCC CTA GCT AAA CTA TTA TTT TGG
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ..A ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ..A ... ... ... ... ... ... ... ... ... ... ... ..C ... ... ... ... ...
Garc   ..A ... ... ... ... ... ... ... ... ... ... ... ..C ... ... ... ... ...
Tier   ..A ... ... ... ... ... ... ... ... ... ... ... ..C ... ... ... ... ...
Pine   ..A ... ... ... ... ... ... ... ... ... ... ... ..C ... ... ... ... ...
Mari   ..A ..C ... ... ... ..C ... ... ... ... ... ..G ..C ... ... ... ... ..A
Cave   ..A ... ... ... ... ... ... ... ... ... ... ... ..C ... ... ... ... ...
Made   ..A ..C ... ... ... ..C ... ... ... ... ... ..G ..C ... ... ... ... ..A
H.v.   ..A ... .A. ... ... ..C ..C ... ... ..A ..T T.. ..C ... T.. ..G ... ...
```

FIG. 3.2--*Continued*

```
Zion   ACC CTA GTA GCC AAT ACA ATA ATC CTA ACA
RedC   G.. ... ... ... G.. ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ... ... ... ... ... ... ... ...
Garc   ... ... ... ... ... ... ... G.. ... ...
Tier   ... ... ... ... ... ... ... G.. ... ...
Pine   ... ... ... ... ... ... ... G.. ... ...
Mari   ... ..G ..G ... ..C ... G.. ... ... ...
Cave   ... ... ... ... ... ... ... ... ... ...
Made   ... ..G ..G ... ..C ... G.. ... ... ...
H.v.   ... T.. ... ... ..C ?.. ... ..T ..? ...
```

FIG. 3.2--*Continued*

```
Zion                AAACATGTTGGTTAAAACCCCTCCTTTACTA ATT AAC CCC TTT GCC CTA TTT
RedC                ............................... ... ... ... ... ... ... ...
OakC                ............................... ... ... ... ... ... ... ...
Cana                ............................... ... ... ... ... ... ... ...
Garc                ............................... ... ... ... ... ... ... ..C
Tier                ............................... ... ... ... ... ... ... ..C
Pine                ............................... ... ... ... ... ... ... ..C
Mari                ............................... ... ... ... ... ... ... ..C
Cave                ............................... ... ..T ... ... ... ... ..C
Made                ............................... ... ... ... ... ... ... ..C
H.v.                ...............C............... ... ... ... ... ... ... ..C

Zion   ATT TTA CTT CTA GCC TAG CAC TAG GCA CTG TTA CAA CAC TGT CAA GCT TCC ACT
RedC   ... ... T.. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... T.. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ... T.. ... ... ... ... ... ... ... C.. ... ... ... ... ... ... ...
Garc   ... ... T.. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Tier   ... ... T.. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Pine   ... ... T.. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Mari   ... ... TA. ... ... ... ... ... ... ... .A. .C. G.. .A. ... ... A.. ...
Cave   ... ... T.. ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Made   ... ... TA. ... ... ... ... ... ... ... .A. .C. G.. .A. ... ... A.. ...
H.v.   ... .C. TC. ... ... ... ... .T. ... ..A ... .T. .CT .A. ... ... AT. ...

Zion   TGA TCC TAG CCT GGA TCG GGC TTG AAA TCA ACA CAC TAG CTA TTA TCC CAT TAA
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Garc   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Tier   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Pine   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Mari   ... ... ... .T. .A. .T. .T. .?. ... ... ... .G. ... ... ... ... ..C .G.
Cave   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Made   ... ... ... .T. .A. .T. .T. ... ... ... ... .G. ... ... ... ... ..C .G.
H.v.   ... ... ... ... .A. ... .C. ... ... .T. .T. .GT ... ..G ... .T. .CC ...

Zion   ATA CTA AAA CTC CCC ACC CAC GGG CTA TTG AAG CCG CAA CAA AAT ACT TCT TAA
RedC   ... ... ... ... .T. ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... .C. ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ... ... ... ... ... ... .C. ... ... ... ... ... ... ... ... ...
Garc   ... ... ... ... ... ... ... ... .C. ... ... ... ... ... ... ... ... ...
Tier   ... ... ... ... ... ... ... ... .C. ... ... ... ... ... ... ... ... ...
Pine   ... ... ... ... ... ... ... ... .C. ... ... ... ... ... ... ... ... ...
Mari   ..G ?C. ... .C. ... ... .G. ... ... ... ... ... ... ... ... ... ... ...
Cave   ... ... ... ... ... ... ... ... .C. ... ... ... ... ... ... ... ... .G.
Made   ..G .C. ... .C. ... ... ... ... ... ... ... ... ... ... ... ... ... .G.
H.v.   ... .C. ... .C. ... ... ... .A. ... ... ... ... ... ... ... ... ... ...

Zion   ACC AAG CAG CAG CAT CAG CCC TAA TTT TGT TTG CCA GCA CAA TTA ATG CCT GAC
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Garc   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Tier   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Pine   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Mari   ... ... ... ... ... ... ... ... ... ... .A. .C. ... .G. ... ... ... ...
Cave   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Made   ..T ... ... ... ... ... ... ... ... ... .A. .CA ... .G. ... ... ... ...
H.v.   ..T ... ... ... ... ... ... ... ... ... .C. .A. .T. ... ... ... ... ...
```

FIG. 3.3. mtDNA sequence of portion of *ND*-2 gene in populations of *Hyla arenicolor* (corresponding to positions 5948-6340 in *Xenopus laevis*). Dot (.) denotes identity with the first sequence; dash (-) denotes a gap; ? denotes an ambiguity.

```
Zion   CTA CTG GTG AGT GAG CC- ATT AAT ACC CAA ATT AGC GCC GCC CCC TCT ATT CTG
RedC   ... ... ... ... ... ..- ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ..- ... ... ... ... ... ... ... A.. ... ... ..A
Cana   ... ... ... ... ... ..- ... ... ... ... ... ... T.. C.. ... ... ...
Garc   ... ... ... ... ... ..- ... ... ... ... ... ... T.. ... ... ... ...
Tier   ... ... ... ... ... ..- ... ... ... ... ... ... T.. ... ... ... ...
Pine   ... ... ... ... ... ..- ... ... ... ... ... ... T.. ... ... ... ...
Mari   .G. ... ... ... .G. ..- ... ..C ... ... ..C ... T.. ..G G.. A.G ..C ..A
Cave   ... ... ... ... ... ..- ... ... ... ... ... ... T.. ... ... ... ...
Made   ... ... ... ... .G. ..- ... ... ... ... ..C ... T.. ... ... A.. ..C ..A
H.v.   ..C .G. ... ... .?A G.T ..C ... ..T ... ... .A. ..T A.. ..? ..A ... T..

Zion   GCT TCA ATC GCT T-G TGC ATA AAA CTA GGC ATT GCC CCG TTC CAC TTC TGA CTT
RedC   ... ... ... ... .-. ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   A.. ... ... ... .-. ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ... ... .-. ... ... ... ... ... ... ... ... ... ... ... ... ...
Garc   ... ... ... ... .-. ... ... ... ... ... ... ... ... ... ... ... ... ...
Tier   ... ... ... ... .-. ... ... ... ... ... ... ... ... ... ... ... ... ...
Pine   ... ... ... ... .-. ... ... ... ... ... ... ... ... ... ... ... ... ...
Mari   A?. ... TC? ??. .TA ... ... ... ... ..T ... ... ... ... ..T ..T ... ...
Cave   ... ... ... ... .-. ... ... ... ... ... ... ... ... ... ... ... ... ...
Made   A.. ... TCG C.. .TA ... ... ... ... ..T ... ... ... ... ..T ..T ... ...
H.v.   .T. ... ..T ..C C-A ... ..G ... T.. ..A ... ..T ..A ..T ... ... ... ...

Zion   CCA GAA GTC CTC CAA GG
RedC   ... ... ... ... ... ..
OakC   ... ... ... ... ... ..
Cana   ... ... ... ... ... ..
Garc   ... ... ... ... ... ..
Tier   ... ... ... ... ... ..
Pine   ... ... ... ... ... ..
Mari   ... ... ... ... ... ..
Cave   ... ... ... ... ... ..
Made   ... ... ... ... ... ..
H.v.   ..G ... ... ... ... ..
```

FIG. 3.3--*Continued*

```
Zion   AAT CAT AAA GAT ATT GGT ACC CTA TAC CTT GTA TTC GGG GCC TGA GCC GGG ATG
RedC   ... ... ... ... ..C ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ..C ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ... ... ..C ... ... ... ... ... ... ... ... ... ... T.. ..T ...
Garc   ... ... ... ... ..C ... ... ... ... ... ... ... ... ... ... ... ..T ...
Tier   ... ... ... ... ..C ... ... ... ... ... ... ... ... ... ... ... ..T ...
Pine   ... ... ... ... ..C ... ... ... ... ... ... ... ... ... ... ... ..T ...
Mari   ..A .?. .T. ... ..C ..C ... ... ... ... ... ... ... ..T ..G ..? ..A ...
Cave   ... ... ... ... ..C ... ... ... ... ... ... ... ... ... ... T.. ..T ...
Made   ..C .CA .T. ... ..C ..C ... ... ... ... ... ... ..A ..T ..G ... ..A ...
H.v.   ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???

Zion   GTC GGC ACT GCC CTT AGC CTC TTA ATC CGA GCA GAG CTC AGC CAA CCT GGC TCC
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ... ... ... ... ... ... ... ... ... ..A ... ... ... ... ... ...
Garc   ... ... ... ... ... ... ... ... ... ... ... ..A ... ... ... ... ... ...
Tier   ... ... ... ... ... ... ... ... ... ... ... ..A ... ... ... ... ... ...
Pine   ... ... ... ... ... ... ... ... ... ... ... ..A ... ... ... ... ... ...
Mari   ..A ... ... ... ... ... ..A ... ... ... ... ..A ... ... ... ... ... ...
Cave   ... ... ... ... ... ... ... ... ... ... ... ..A ... ... ... ... ... ...
Made   ..A ... ... ... ... ... ..A ... ... ... ... ..A ... ... ... ... ... ...
H.v.   ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???

Zion   CTT CTA GGC GAT GAC CAA ATC TAT AAT GTT ATC GTC ACA GCC CAT GCC TTT GTT
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ... ... ... ... ... ... ... ..C ... ... ... ... ..C ... ... ...
Garc   ... ... ... ... ... ... ... ... ... ..C ... ... ... ... ..C ... ... ...
Tier   ... ... ... ... ... ... ... ... ... ..C ... ... ... ... ..C ... ... ...
Pine   ... ... ... ... ... ... ... ..C ... ..C ... ... ... ... ..C ... ... ...
Mari   ... G.. ... ... ... ..G ..? ... ..A ... ... ..T ... ... ... ..? ... ...
Cave   ... ... ... ... ... ... ... ... ... ..C ... ... ... ... ..C ... ... ...
Made   ... T.. ... .?. ... ..G ..? ... ..? ... ... ..T ... ... ... ... ... ...
H.v.   ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ?.. ... C.. ... ... ... ... ..?

Zion   ATG ATT TTC TTT ATA GTA ATG CCC ATC CTA ATC GGG GGA TTC GGC AAC TGA TTG
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ... ... ... ... ... ... ... T.. ... ... ... ... ... ... ... ...
Garc   ... ... ... ... ... ... ... ... ... T.. ... ... ... ... ... ... ... ...
Tier   ... ... ... ... ... ... ... ... ... T.. ... ... ... ... ... ... ... ...
Pine   ... ... ... ... ... ... ... ... ... T.. ... ... ... ... ... ... ... ...
Mari   ..A ... ..T ... ... ... ... ... ..T ... ..T ..C ..G ... ... ..? ?.. ...
Cave   ... ... ... ... ... ... ... ... ... T.. ... ... ... ... ... ... ... ...
Made   ..A ... ..T ... ... ... ... ... ..T ... ..T ..C ..G ... ... ... ... ...
H.v.   ..A .CC ... C.C ... ..G ..A ... ... T.. ... ... ..G ... ..T ..T ... ..A

Zion   ATT CCC CTG ATG ATT GGG GCA CCC GAC ATA GCT TTC CCG CGA ATA AAT AAT ATA
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Garc   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Tier   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Pine   ... ... ... ... ... ... ... ... ... ... ... ... ..A ... ... ... ... ...
Mari   ... ..A ..A ..A ... ... ... ... ... ?.. ..C ... ... ..G ..C ... ..G
Cave   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Made   ... ..A ..A ..A ... ... ... ..? ... ... ..C ... ... ... ... ..C ... ..G
H.v.   G.C ... T.A ..A ... ..A ... ... ..T ... ..C ?.. ..T ..G ... ... ..C ...
```

FIG. 3.4. mtDNA sequence of portion of *CO*-I gene in populations of *Hyla arenicolor* (corresponding to positions 7427-8098 in *Xenopus laevis*). Dot (.) denotes identity with the first sequence; dash (-) denotes a gap; ? denotes an ambiguity.

```
Zion   AGC TTT TGA CTT CTT CCC CCC TCT TTT CTT CTT CTT TTA GCA TCG GCT GGG GTC
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Garc   ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Tier   ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Pine   ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Mari   ... ... ... ... ... ... ..T ... ..C ..A ... ... ... ... ..A ... ... ...
Cave   ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Made   ... ... ... ... ... ... ..T ... ..C ..A ... ... ... ... ..A ... ... ...
H.v.   ... ... ... ..C .C. ... ... ... ... ... ..C ... ... ..G ..A ..C ... ...

Zion   GAA GCT GGG GCC GGG ACG GGG TGA ACA GTT TAC CCA CCT CTT GCT GGA AAC TTA
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ..G ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ... ?.. ... ... ..A ... ... ... ... ... ..A ... ... ... ... ...
Garc   ... ... ... ... ... ... ..A ... ... ... ... ... ..A ... ... ... ... ...
Tier   ... ... ... ... ... ... ..A ... ... ... ... ..T ..A ... ... ... ..T ...
Pine   ... ... ... ... ... ... ..A ... ... ... ... ... ..A ... ... ... ... ...
Mari   ... .T. ..A .GT ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???
Cave   ... ... ... ?.. ... ... ..A ... ... ... ... ... ..A ... ... ... ... ...
Made   ... ... ... .?T ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???
H.v.   ... ..A ..A ..T ..A ..A ..A ... ..T ..A ... ..T ..G ... ..C ... ..T ...

Zion   GCC CAC GCC GGA CCA TCC GTT GAT CTA ACT ATT TTC TCC CTC CAC TTG GCG GGG
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Garc   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Tier   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Pine   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Mari   ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???
Cave   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Made   ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???
H.v.   ... ... ..T ... ... ..A ..C ... T.G ... ..C ... ..A ... ... C.A ..A ..T

Zion   GTT TCT TCC ATC TTA GGT GCT ATT AAT TTT ATT ACT ACA ATT CTA AAT ATA AAA
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ... ... ..T ... ... ... ... ...
Cana   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Garc   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Tier   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Pine   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Mari   ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???
Cave   ... ... ... ... ..G ... ... ... ... ... ... ... ... ... ... ... ... ...
Made   ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???
H.v.   ... ..C ..A ..T ... ..G ... ... ... ... ... ... ..G ... ..G ... ... ...

Zion   CCC CCG TCA ATA ACA CAA TAC CAA ACA CCC CTG TTT GTG TGA TCT GTT CTC ATT
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Garc   ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Tier   ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Pine   ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Mari   ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???
Cave   ... ..A ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
Made   ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???
H.v.   ..T ... ..G ..G ... ... ... ... ... ... ..A ... ... ... ... ... ... ...
```

FIG. 3.4--*Continued*

```
Zion   ACC GCT GTG TTG TTA CTC CTG TCT CTC CCA GTT TTA GCA GCA GGC ATC ACC ATG
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... C.. ... ... ... ... ... ... ... ... ... ... ... ... ...
Cana   ... ... ..A ... ... ... ..A ... ... ..G ... ... ... ... ... ... ... ...
Garc   ... ... ..A ... ... ... ..A ... ... ..G ... ... ... ... ... ... ... ...
Tier   ... ... ..A ... ... ... ..A ... ... ..G ... ... ... ... ... ... ... ...
Pine   ... ..? ..A .?. ... ... ..A ... ... ..G ... ... ... ... ... ... ... .?.
Mari   ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???
Cave   ... ... ..A ... ... ... ..A ... ... ..G ... ... ... ... ... ... ... ...
Made   ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???

H.v.   ..T ... ... C.T C.T ..T T.A ... ..T ... ... C.G ... ... ..A ..T ... ..A
Zion   TTA CTG ACT GAC CGA AAC CTA AAT ACA ACA TTC TTT GAT CCG GCA GGA GGG GGG
RedC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ... ..A
Cana   ... ..A ... ... ... ... T.. ... ... ... ... ... ... ... ... ..G ... ..A
Garc   ... ..A ... ... ... ... T.. ... ... ... ... ..A ... ... ... ... ... ..A
Tier   ... ..A ... ... ... ... T.. ... ... ... ... ..? ... ... ... ... ... ..A
Pine   ... ..A ..? ... ... ... T.. ... ... ... ... ..A ... ... ... ... ... ..A
Mari   ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???
Cave   ... ..A ... ... ... ... T.. ... ... ... ... ..A ... ... ... ... ... ..A
Made   ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ??? ???
H.v.   C.G ..T ... ..T ..G ..T ... ..C ... ... ..T ..A ..C ... ... ..G ..A ..A

Zion   GAC CCC GTC CTT TAC CAA CAC CTG
RedC   ... ... ... ... ... ... ... ...
OakC   ... ... ... ... ... ... ... ...
Cana   ... ... ... ... ... ... ... ...
Garc   ... ... ... ... ... ... ... ...
Tier   ... ... ... ... ... ... ... ...
Pine   ... T.. ... G.. ... ... ..T ...
Mari   ??? ??? ??? ??? ??? ??? ??? ???
Cave   ... ... ... ... ... ... ... ...
Made   ??? ??? ??? ??? ??? ??? ??? ???
H.v.   ... T.. ..T .GG ... ... ... ...
```

FIG. 3.4--*Continued*

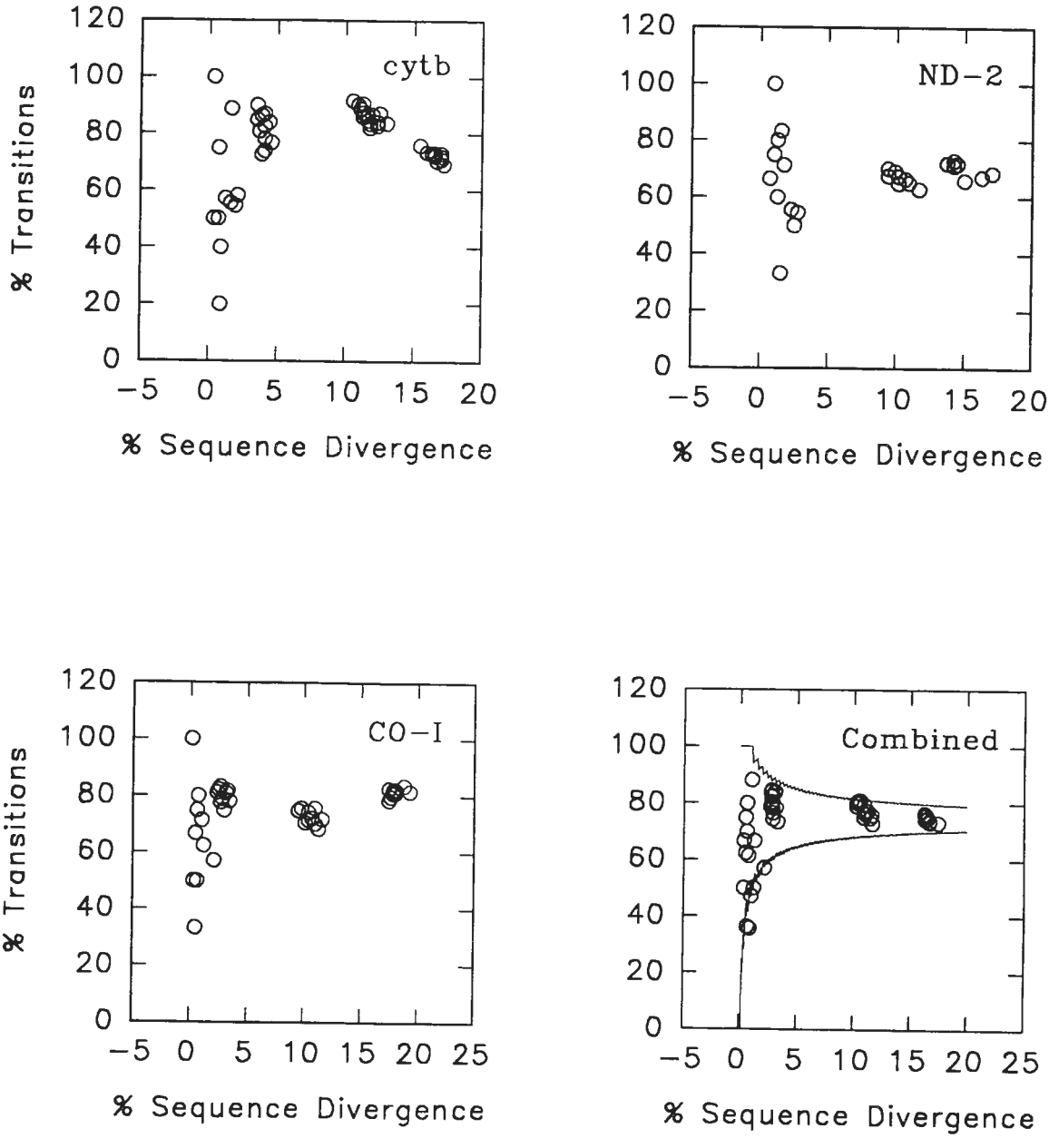

FIG. 3.5. Percent transitions vs. percent sequence divergence for populations of *Hyla arenicolor* and for outgroup *Hyla versicolor*.

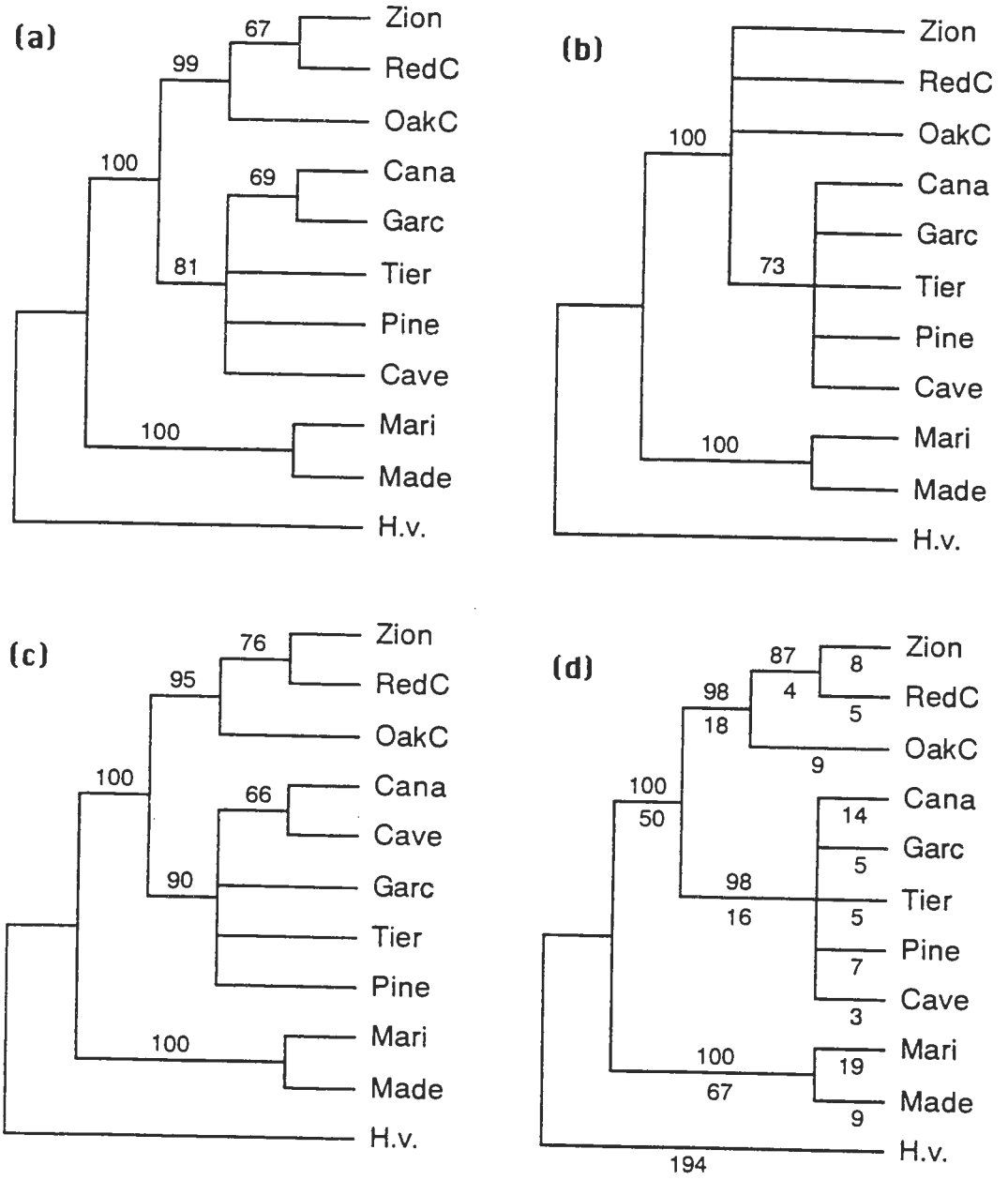

FIG. 3.6. Relationships among populations of *Hyla arenicolor* obtained by maximum parsimony. Numbers on branches indicate percentage of bootstrap replicates supporting each node (combined data also shows branch lengths below line). (a) Strict consensus of 3 trees obtained from *cyt*b sequence data, tree length = 169, CI = 0.90, RI = 0.88, g_1 = -1.84; (b) strict consensus of 5 trees obtained from *ND*-2 sequence data, tree length = 104, CI = 0.93, RI = 0.88, g_1 = -2.84; (c) strict consensus of 3 trees obtained from *CO*-II sequence data, tree length = 152, CI = 0.95, RI = 0.90, g_1 = -1.26; (d) strict consensus of 5 trees obtained from combined data, tree length = 427, CI = 0.92, RI = 0.88, g_1 = -1.94.

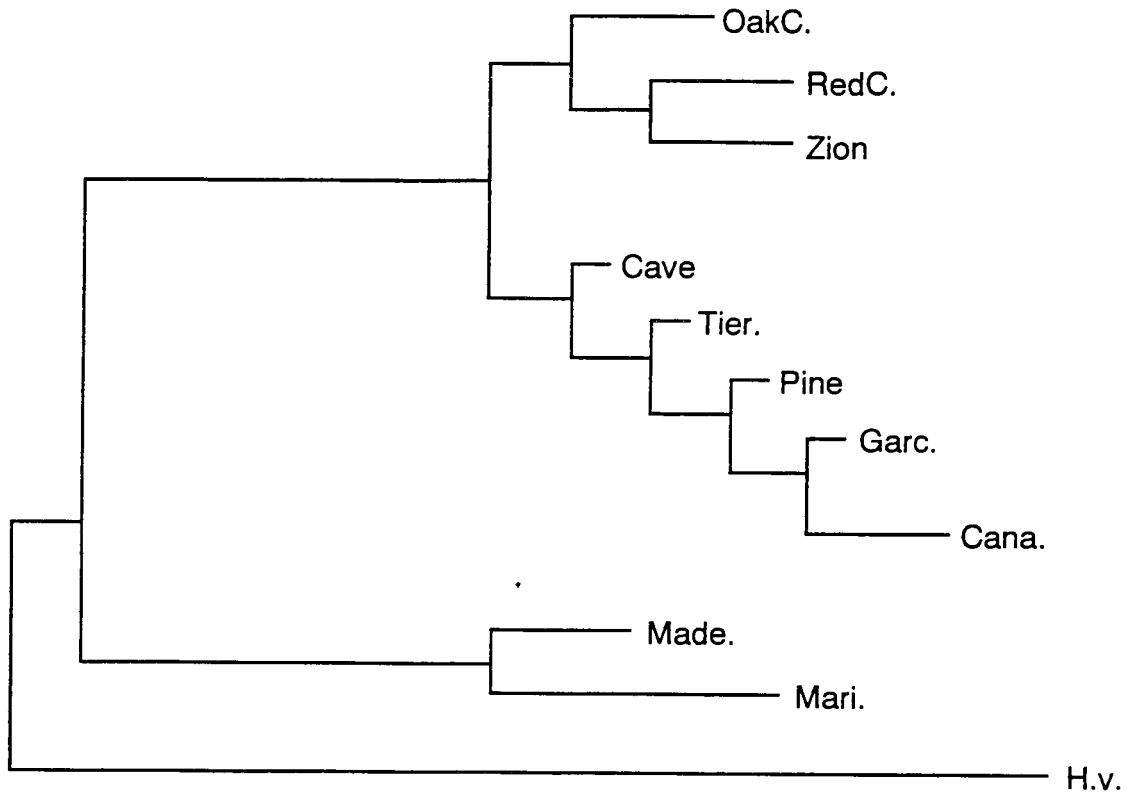

FIG. 3.7. Relationships among populations of *Hyla arenicolor* obtained by maximum likelihood method.

CHAPTER 4

GENES AND BIOGEOGRAPHY: MODELING GENETIC DIVERGENCE OF POPULATIONS UNDER ISOLATION BY DISTANCE, VICARIANCE, AND DISPERSAL

Abstract.--Populations in isolation can be established via dispersal of individuals from established populations or through a series of vicariant events that fragment a formerly continuous distribution. Theoretical results based on gene flow and coalescent models suggest that each of these processes leaves a distinct signal on the pattern of genetic variation among populations if gene flow is absent among populations, and if vicariant events are separated in time. A computer model of allelic variation and population structure was used to test these predictions and to determine the conditions under which the different patterns of genetic variation are to be expected. Results from the computer model indicate that dispersal and range contraction (vicariant) processes can be distinguished from one another, as well as from ongoing gene flow among populations, under a wide range of conditions. The application of these results to empirical studies of isozyme variation is discussed.

INTRODUCTION

Geographic structuring of genetic variation can be the result of population processes such as localized selection or genetic drift, or

alternatively, it can be due to historical biogeographic processes such as dispersal or vicariance. Because evolutionary change, and ultimately speciation, is a result of population differentiation, it is important to know which of these processes has been most influential in establishing geographic patterns of variation. However, distinguishing among population processes and historical events can be difficult because each can leave similar signatures on the patterns of genetic variation within and among populations. Researchers have recently begun to address the different causes of population structure, developing new analysis techniques for molecular data that attempt to distinguish among the various population processes (Excoffier and Smouse 1994; Templeton et al. 1995). Templeton et al. (1995) described a method of phylogenetic analysis that examines the spatial distribution of haplotypes using a hierarchical distance approach. The authors proposed that the effects of localized gene flow and dispersal can be distinguished in this manner because dispersal events will cause certain haplotypes to have broader spatial distributions. However, the analysis requires character state data (RFLP or sequence data) from a large number of individuals such that haplotypes separated by a single character state change can be arranged into "one-step clades." As such, the approach is not applicable to isozyme data because allele frequencies do not represent discrete character states. Thus, there remains a discouraging lack of analysis tools for isozyme data that attempt to address the causes of geographic structure.

Wright's *F*-statistics (Wright 1943, 1946) represent some of the earliest analytic theory dealing with the geographic partitioning of variation. *F*-statistics are commonly applied to isozyme data because they provide a useful measure of the degree of population substructuring. In addition, *F*-statistics are often used to provide estimates of the rate of genetic exchange among subpopulations. Equations relating F_{ST} to gene flow have been developed under a variety of gene flow models, and in many published reports F_{ST} estimates are converted to measures of gene flow irrespective of whether correlations in allele frequencies are likely to be due to current gene flow among populations. While this is often a convenient transformation (converting F_{ST}, which ranges from zero to one, into a measure of population similarity ranging from zero to infinity), it is not always appropriate to refer to the resulting similarity value as gene flow because the models relating F_{ST} to gene flow assume an equilibrium between gene flow and genetic drift. Gene flow estimates based on F_{ST}'s may seriously overestimate or underestimate actual gene flow amounts if populations are not in equilibrium. For example, two populations that were fragmented in the past will continue to show F_{ST}'s less than one due to shared alleles even if there is currently no gene flow among them. Additionally, gene flow may obscure any ancestral relationships among populations that might otherwise be detectable from molecular data. Thus, while *F*-statistics are valuable for assessing population subdivision, they do not by themselves reveal anything about the causal relationships among

subdivided populations and may provide inaccurate measures of gene flow under non-equilibrium conditions.

These points notwithstanding, *F*-statistics are proving to be useful in combination with geographical information. Slatkin (1993) presented analytic theory based on a stepping-stone model of gene flow (Kimura and Weiss 1964) that predicts a linear relationship between the logarithm of gene flow (estimated from pairwise F_{ST}'s) and the logarithm of distance between populations in both one and two dimensions. This relationship can be used to test for isolation by distance resulting from limited gene flow among populations. Slatkin (1993) also developed analytical results for a model of range expansion in which new populations are founded sequentially in time and space, but with no gene flow among populations after each founding event. The predicted relationship of gene flow to geographic distance is distinctly different for these two models, suggesting that it may be possible to distinguish between the effects of historical associations and current population genetic phenomena in establishing patterns of population structure.

In this chapter, I use a computer model to build on the results of Slatkin (1993), introducing additional biogeographic scenarios of vicariance and dispersal. My goal in this effort is to provide expected patterns of genetic differentiation for each scenario that can be applied in the analysis of isozyme data. Both equilibrium and non-equilibrium patterns will be examined. Published isozyme studies are numerous, going back to the mid-1960's when

protein electrophoresis was first applied to population level studies. However, few of these efforts have attempted to address the causes of geographic structuring of variation. Additionally, despite the popularity of more recent molecular techniques, protein electrophoresis remains a relatively rapid and inexpensive method for assessing genetic variation within and among populations. The approach presented here is intended to provide an additional tool in the analysis of isozyme data.

REVIEW OF THEORY

Early models of genetic differentiation and gene flow are based on probabilities of identity by descent (e.g., Nei 1972, 1973). For samples from two subpopulations, if f_0 is the probability of identity by descent of two genes sampled from the same subpopulation and f_1 is the probability of identity by descent for two genes samples from different subpopulations, then Wright's (1951) F_{ST} can be written as

$$F_{ST} = \frac{f_0 - f}{1 - f} \qquad \textbf{(1)}$$

where $f = (f_1 + f_0)/2$ (Nei 1973). Using a coalescent approach, Slatkin (1991) showed that identity by descent is related to expected coalescent time ($\bar{t}$) by

$$f \approx 1 - 2\mu\bar{t} \qquad \textbf{(2)}$$

where μ is the mutation rate and time is measured in generations. This is an approximate formula that holds when $\bar{t}$ is much less than $1/\mu$. Substituting (2) into (1) gives

$$F_{ST} \approx \frac{\bar{t}_1 - \bar{t}_0}{\bar{t}_1 + \bar{t}_0} \qquad \textbf{(3)}$$

if $\mu \ll 1$, where $\bar{t}_0$ is the expected coalescent time of two genes sampled from the same subpopulation and $\bar{t}_1$ is the expected coalescent time of two genes sampled from different populations. Using an island model of population structure, Wright (1951) showed that F_{ST} is related to gene flow by

$$Nm = \frac{1}{4}\left(\frac{1}{F_{ST}} - 1\right) \qquad \textbf{(4)}$$

where N is the population size and m is the migration rate. Slatkin (1991) denoted the estimate of gene flow between pairs of populations obtained from Equation (4) as $\hat{M}$. Substituting (3) into (4) gives

$$\hat{M} \approx \frac{\bar{t}_0}{2(\bar{t}_1 - \bar{t}_0)}. \qquad \textbf{(5)}$$

Expected coalescent times in these formulas depend on parameters of the population model. It has been shown for both an island model and a stepping-stone model that $\bar{t}_0$ depends on the total population size but is independent of the migration rate (Slatkin 1987a; Strobeck 1987; Hey 1991). Thus for the situation with *d* subpopulations each containing *N* diploid individuals, $\bar{t}_0 = 2Nd$ (Slatkin 1987a). The expected coalescent time for two genes sampled from different subpopulations is the sum of two components: the expected time until the genes are in the same population and the expected coalescent time for two genes in the same population. The latter term is again *2Nd*. The former term has been worked out for a circular array of subpopulations and is given by $\bar{t}_s = (d-i)i/2m$ where i is the number of steps separating two subpopulations (Slatkin 1991). (An equivalent formula for a linear array has not been derived due to the difficulty of obtaining an analytical solution [e.g., see Maruyama 1970, 1971]). Combining these two terms gives $\bar{t}_i = 2Nd + ((d - i)i)/2m$ for between population coalescent times (Slatkin 1991). Therefore, from (5)

$$\hat{M} \approx \frac{2Nm}{i} \qquad \textbf{(6)}$$

when $i \ll d$ (Slatkin 1991). Equation (6) provides an expected relationship between gene flow and geographic distance for a one-dimensional stepping-stone model at equilibrium. Slatkin (1993) used this result to predict the

outcome from his "radiation" model in which at time τ in the past a single population gives rise to *d* subpopulations, representing a rapid radiation from a center of origin. Between time τ and the present, a specified amount of migration occurs among adjacent subpopulations. Thus, if enough time has passed since the radiation event, an isolation by distance pattern as predicted by Equation (6) is expected. Equation (6) will subsequently be referred to as the isolation by distance model.

Slatkin (1993) also developed a model of range expansion in which at time τ in the past there exists a single population that gives rise to a neighboring subpopulation. Then, every $\Delta\tau$ generations the most recently founded subpopulation gives rise to a new subpopulation until there are a specified number of subpopulations (Fig. 4.1). No gene flow occurs among established subpopulations. This model corresponds to what might be expected under a biogeographic scenario of disjunct populations founded via dispersal and will be referred to subsequently as the dispersal model. Numbering the subpopulations from oldest to youngest, the expected coalescent time for two genes sampled from the same subpopulation is $\bar{t}_{ii} = 2N$ (Hudson 1990); the expected coalescent time for two genes, one sampled from population *i* and one from population *j* (*i* < *j*), is given by $\bar{t}_{ij} = 2N + \tau - (i-1)\Delta\tau$ (Slatkin 1991). Thus, substituting into Equation (5) gives

$$\hat{M} = \frac{N}{\tau - (j-1)\Delta\tau} \qquad (7)$$

which does not depend on the distance separating two subpopulations, but rather on which subpopulation of the two was founded earlier in time.

A third biogeographic scenario to be considered here will be referred to as the range contraction model. This scenario models a process of gradual regional change beginning at the boundary (or boundaries) of a species' range and proceeding in a specific direction over time, but with relict populations left behind in isolated refugia (Fig. 4.2). Thus, the species is subject to a series of vicariant events over time that fragment its original range. As with the isolation by distance model, each subpopulation is founded from the same original population, but unlike the isolation by distance model, each founding event is separated in time by $\Delta\tau$ and no gene flow occurs among established subpopulations. Numbering the populations from 1 to *d*, with 1 being the founding population and *d* the earliest subpopulation, the expected coalescent time for two genes sampled from the same subpopulation is again $\bar{t}_{ii} = 2N$. The expected coalescent time for two genes, one sampled from subpopulation *i* and one from subpopulation *j* ($i < j$), is given by $\bar{t}_{ij} = 2N + (j-1)\Delta\tau$. Substituting these equations into (5) gives

$$\hat{M} = \frac{N}{(j-1)\Delta\tau}. \qquad \textbf{(8)}$$

Thus, as with the dispersal model, $\hat{M}$ is a function of the subpopulation that was founded earlier in time. Note that $\hat{M}$ in Equations (7) and (8) is not actually a measure of gene flow because no gene flow is occurring among established subpopulations; rather, $\hat{M}$ in these equations is a measure of population similarity due to common ancestry. Equations (7) and (8) are also distinct from Equation (6) in that they do not represent equilibrium conditions but rather temporary patterns due to historical associations that will degrade over time.

It is useful to linearize equations (6)-(8) such that regression analysis can be applied to empirical data sets. Equation (6) can be transformed logarithmically to give

$$\log(\hat{M}) = \log(2Nm) - \log(i). \qquad \textbf{(9)}$$

Hence, under isolation by distance, plots of $\log(\hat{M})$ versus $\log(i)$ are expected to have a slope of negative one and a *y*-intercept of $\log(2Nm)$ (Slatkin 1993). Equations (7) and (8) are linearized as

$$\frac{1}{\hat{M}} = \frac{\tau+\Delta\tau}{N} - \frac{\Delta\tau}{N} i \qquad \textbf{(10)}$$

for the dispersal model, and

$$\frac{1}{\hat{M}} = -\frac{\Delta\tau}{N} + \frac{\Delta\tau}{N} j \qquad \textbf{(11)}$$

for the range contraction model. Equations (9)-(11) suggest that estimates of Nm and $\Delta\tau/N$ may be obtained for empirical data sets from patterns of similarity by examining slope and intercept values of the corresponding regressions.

Equations (6)-(11) provide expected patterns of population differentiation for three scenarios of population structure: isolation by distance, dispersal, and range contraction (or vicariance). In order to test whether or not the patterns predicted from these analytical models are likely to be found in empirical data sets, such as from isozyme studies, I used a numerical computer model to simulate each scenario. By taking a numerical approach, it was possible to test the sensitivity of the analytical models to the relevant assumptions, and in doing so, to introduce variations in the models that added biological realism. If genetic patterns resulting from population and biogeographic processes can be expected to match the analytical models under certain circumstances, it may be possible to use gene frequency data to establish which scenario has played the determining role in the formation of disjunct populations. It may also be possible to estimate the relevant population parameters such as Nm and $\Delta\tau$ from empirical data.

COMPUTER MODELS

Random Walk Model

As noted above, equation (6) is based on a circular array of subpopulations. The circular paradigm has been adopted in previous computer models of one-dimensional stepping-stone population structure because random walk probabilities for circular arrays have been worked out previously (Feller 1957). I have chosen not to use this approach since linear arrays are probably more common in nature. In order to determine the behavior of $\bar{t}_s$ for a linear array of subpopulations, I used a Monte Carlo approach to generate the distribution of times for genes separated by *i* steps to arrive in the same subpopulation for a given subpopulation size and migration rate. The computer simulation selects one gene from each subpopulation, keeping track of each selected gene's progress as all genes are exchanged at random among adjacent subpopulations. After each selected gene has been "paired" with every other selected gene by their arrival in the same subpopulation, the process repeats to generate mean times for each step size. The entire procedure then repeats 100 times to generate the expected values of the means. Results of this preliminary model are used to examine the characteristics of random walks in linear arrays and to determine the applicability of equation (6) to linear arrays.

Population Structure Models

The main simulation effort to be discussed below models a pair of alleles (0 or 1) at a single gene locus in subpopulations of constant size, *N*. Each individual in generation *t* is generated by picking two individuals at random from generation *t* - 1 and selecting one gene at random from each of these individuals. More specifically, a new individual is created from four random numbers; two random numbers between 1 and *N* determine the new individual's parentage, and two random numbers between 0 and 1determine which gene is drawn from each parent. Thus generations are discrete and non-overlapping, and diploid individuals are replaced by a Poisson-distributed number of offspring. New subpopulations are founded in a similar manner, taking a random sample of *2N* genes from the founding subpopulation. Migration is accomplished by generating *Nm* random numbers between 1 and *N*, each of which specifies an individual to be moved to an adjacent subpopulation. After a specified number of generations, allele frequencies, genotype frequencies, and *F*-statistics are calculated. *F*-statistics are calculated using the entire population in each case. Thus, there is no statistical sampling error; calculated F_{ST}'s are the "true" F_{ST}'s for each pairwise comparison. However, since a finite number of populations are modeled, there will be randomness in the results due to genetic sampling effects (see Weir 1996, for a discussion of genetic and statistical sampling). The entire process

is then repeated to generate a sample of independent gene loci. Except where noted, all cases presented below used a linear array of 10 subpopulations with $N = 500$.

The numerical model was run in three different configurations, corresponding to the three analytical models discussed previously. Additionally, each of these configurations was subject to modifications intended to add biological realism, and also to test the sensitivity of the model to various assumptions.

Isolation by Distance Model.--The isolation by distance model is similar conceptually to Slatkin's (1993) radiation model. A specified number of subpopulations are founded from a single gene pool at τ generations in the past. Between τ and the present, individuals migrate to adjacent subpopulations at a rate specified by m.

Modifications were implemented to the numerical isolation by distance model to create a scenario of disjunct population structure. In nature, closely grouped subpopulations may experience high rates of gene flow while more widely separated subpopulations are subject to disproportionately low gene flow due to various barriers to migration (e.g., physical barriers, behavioral barriers, etc.). I modeled this scenario using a linear array of nine subpopulations associated in groups of three (i.e., subpopulations 1, 2, and 3 formed a group as did 4, 5, and 6, etc.). Within a group, subpopulations are separated by a single step and are connected by high levels of gene flow.

Groups are separated by more than one step; gene flow among groups is disproportionately low (Fig. 4.3).

Dispersal Model.--A depiction of the dispersal model is given in Figure 4.2. Every $\Delta\tau$ generations a new subpopulation is founded from the most recently established subpopulation. This process continues until there are a specified number of subpopulations. No gene flow occurs among established subpopulations.

In nature, populations founded by dispersal will often experience bottlenecks as a result of the founding event. Bottlenecks can have a profound influence on the number of alleles maintained in a population (Nei et al. 1975). In the current context, bottlenecks should reduce the correlation of gene frequencies between founding and founder populations, possibly disrupting the expected pattern of similarity under the dispersal model. Bottlenecks were implemented in the numerical dispersal model by generating random numbers between 1 and N_b to select individuals whose genes are represented in the new subpopulation, where $N_b < N$.

Range Contraction Model.--A depiction of the range contraction model is given in Figure 4.3. Every $\Delta\tau$ generations, the range of the species is reduced, but with a relict subpopulation remaining behind (similar to a founding event). The contraction process continues until there are a specified number of subpopulations. No gene flow occurs among established subpopulations.

Range contraction due to climate change, glaciation, or some other

regional event may begin at the edges of a species' range and proceed toward the center of the range (or alternatively it may begin at one end of the range and proceed toward the other end). During this process, it is possible that the central population will remain large and panmictic. A large central population should tend to increase the correlation of gene frequencies among fragmented populations since the rate of genetic drift in the central population will be low. Vicariant events separated in time will all sample genes from this stable central population. Thus, in the numerical range contraction model, a new subpopulation (size N) is founded every $\Delta\tau$ generations from a central population (size N_v, where $N_v > N$) by generating N random numbers between 1 and N_v, thereby specifying the genes that will be represented in the relict subpopulation.

Monte Carlo Replicates.--Finally, to determine whether there is statistical power for estimating Nm and $\Delta\tau/N$ from regression analysis, I generated Monte Carlo replicates of each of the implementations described above. Each loop of the baseline algorithms generated a set of points from which a linear least-squares regression was calculated. Distributions of slope and y-intercept values for each biogeographic scenario were generated in this manner.

MODEL RESULTS

Random Walk Model

In arriving at equation (6), it was assumed that $i \ll d$ such that $(d - i)i \approx di$ and

$$\bar{t}_s \approx di/2m, \tag{12}$$

based on a circular array of subpopulations. For most biological situations, where d is likely to be on the order of 10 to 20 subpopulations, this is a fairly poor approximation. Results from the random walk model (Fig. 4.4) indicate that equation (12) is actually a better approximation for linear arrays of subpopulations than it is for circular arrays when d is small. Equation (12) gives a reasonable approximation for mean arrival times along a linear array for small d and i. Hence, equation (6), which was derived using this simplification, will be used below to provide analytical expectations for the isolation by distance model.

Isolation by Distance Model.--As expected, immediately after the radiation event subpopulations are essentially panmictic; any differentiation among them is not correlated with distance (Fig. 4.5). As time passes, the equilibrium pattern described by equation (6) is approached, with the rate of approach dependent on Nm. As noted by Slatkin (1993), the isolation by distance pattern becomes apparent for subpopulations in close proximity first

(Fig. 4.5b), with more distant populations approaching expected levels of gene flow later in time (Fig 4.5d). If gene flow ceases after isolation by distance has been established, the pattern degrades over time (Fig. 4.6) with near populations deviating from expectations early in time (Fig. 4.6b) and distant populations deviating later in time (Fig. 4.6d).

The model was used to determine the minimum and maximum levels of gene flow under which an isolation by distance pattern might arise. From Figure 4.7 it is apparent that this pattern is quite robust. Not until *Nm* approaches the subpopulation size (i.e., $m = 1$; Fig. 4.7b) does the relationship begin to break down. Under these circumstances, migration is high enough that the entire array of subpopulations remains essentially panmictic and no geographic structuring is possible. For low migration rates, in Figure 4.7 it is shown that given a sufficiently long period of time, exchanging a single individual every 25 generations can produce the isolation by distance pattern (Fig. 4.7c). This corresponds to an exceptionally low migration rate ($m = 8 \times 10^{-5}$).

The disjunct scenario resulted in a high correlation between $\hat{M}$ and distance ($r^2 = 0.89$), but with a regression slope of -1.7 (Fig. 4.8). Thus, a large negative slope may be an indicator of barriers to gene flow, limiting the exchange of individuals to levels lower than expected based on distance alone.

The distributions of slope and intercept values generated from 100 Monte Carlo replicates (Fig. 4.9) indicate that *y*-intercept estimates follow a

well-behaved distribution, with a fairly narrow spread of estimates about the expected values. The distributions of slope estimates show larger variances about the expected values; however, there is little overlap in slope values between the isolation by distance and disjunct scenarios.

Dispersal Model.--Similarity patterns for the dispersal model also conform to expectations (equation [7]) over a wide range of $\Delta\tau$ values (Fig. 4.10b & d). At $\Delta\tau = 200$, which represents a total of 1600 generations between the founding of subpopulation 2 and the founding of subpopulation 10, the predicted pattern is still apparent. Under the population bottleneck scenario (Fig. 4.11), model results continue to match expectations fairly well over a range of founder sizes. Figure 4.10a and c show *Nm* plotted against distance, revealing the absence of an isolation by distance pattern under the dispersal scenario. Distributions of slope and intercept values for the dispersal model are shown in Figure 4.12. For low values of $\Delta\tau$, modal values of the slope and intercept distributions correspond fairly well to model expectations (Fig. 4.12a & b). In contrast, large $\Delta\tau$'s lead to systematic deviations of slope and intercept values from analytical predictions (Fig. 4.12c & d).

Range Contraction Model.--The pattern evident under the range contraction model (Fig. 4.13b & c) appears similar to the isolation by distance model (i.e., negative slope close to one). However, $\hat{M}$ is plotted against distance under the isolation by distance model and against *j* in the range contraction model (where *j* is the population founded earlier in time). Figure

4.13a and c demonstrate a lack of dependence of $\hat{M}$ on distance. The effect of a large central population can be seen from the distribution of slope and intercept values in Figure 4.14. As with the dispersal model, large $\Delta\tau$'s result in systematic deviations of slope and intercept values from predicted values. However, large central populations tend to bring these distributions more in line with model predictions.

Discussion

As noted by Slatkin (1987b), direct and indirect measures of gene flow in natural populations often give contradictory results. Such discrepancies are usually attributed to episodic gene flow or historical changes in species' ranges. Direct measures of gene flow, which monitor the movements of individuals, are necessarily based on short time scales and are thus not well suited for detecting episodic gene flow or rare dispersal events. In contrast, indirect estimates of gene flow are by definition long-term averages of genetic exchange among populations. As previously discussed, one of the difficulties with using allele frequency data for estimating gene flow has been the inability to distinguish similarity due to the actual exchange of individuals (or gametes) among populations and similarity due to ancestral relationships. The results presented here suggest a simple test to distinguish among these alternatives. By plotting pairwise estimates of gene flow against both distance and positional information, it may be possible to distinguish ongoing gene flow from historical

associations under a range of conditions. The computer model does not adhere to many of the assumptions used to derive the analytical models, such as the circular array condition, yet results derived from the computer simulations of population structure match theoretical expectations quite well. Numerical results from the random walk model demonstrate that the mathematical simplifications used to arrive at equation (6) actually describe a linear array of population structure better than they describe a circular array (upon which the theory is based) when the number of subpopulations is small.

Equation (9) and results from the isolation by distance model indicate that an estimate of *Nm* between subpopulations can be obtained from the *y*-intercept of the regression line when an isolation by distance pattern is evident (Slatkin 1993). There is a difficulty with this estimate when applying it to real data, however: The *y*-intercept is dependent on the scale used to measure distance. Ideally, a unit distance should be the scale at which one is attempting to measure gene flow. For example, if kilometers are used to measure physical distance, then the value of *Nm* derived from the *y*-intercept of a regression line will be an estimate of gene flow among groups separated by one kilometer. The population may be panmictic at this scale, resulting in an abnormally high value for *Nm*. Hence, the choice of an appropriate spatial scale is critical for obtaining a meaningful value of *Nm* from the *y*-intercept value.

The slope of $\log(\hat{M})$ versus $\log(distance)$ plots are not affected by scale

but should be noted nonetheless. Population structures based on a simple stepping-stone model of migration are expected to show slopes near negative one for *log-log* plots of $\hat{M}$ versus distance (equation [9]). A regression slope greater than -0.5 is an indication that the system under study is not at equilibrium with regard to gene flow and genetic drift. Slopes greater than -0.5 can be obtained under a dispersal scenario (e.g., Fig. 4.10) or, alternatively, may be due to a recent cessation of gene flow (Fig. 4.6). Regression slopes less than -1.5 are unlikely under a simple isolation by distance model and may be due to barriers to gene flow such as those modeled under the disjunct scenario (Fig. 4.8). Plots of $\hat{M}$ versus geographic distance in a recent study of disjunct populations of Canyon Treefrogs (*Hyla arenicolor*; Murray and Wolf; submitted) gave a highly significant regression ($r^2 = 0.84$) with a slope of -1.8. Gene flow estimates for Canyon Treefrogs showed a pattern of moderate to high gene flow within drainages and among adjacent drainages, but low gene flow among disjunct or widely separated drainages. Thus, the pattern of similarity values for Canyon Treefrog populations may be a result of a disjunct isolation by distance scenario, with few restrictions on gene flow within watersheds and disproportionately low gene flow among populations with no direct drainage connections.

Under the dispersal scenario, scatter plots of $\hat{M}$ versus position result in a distinctive pattern, as predicted by equation (7), which should be established before attempting a regression analysis using $1/\hat{M}$ values. Larger $\Delta\tau$'s

between founding events result in greater deviations about the expected slope and intercept values for empirical data sets, giving systematically high estimates of this parameter. Population bottlenecks during founder events will also tend to give overestimates of $\Delta\tau$. Similarly, for the range contraction model, large $\Delta\tau$'s between vicariant events will result in greater errors in estimating the $\Delta\tau$ parameter (data not shown).

As noted above, patterns observed under the dispersal and range contraction models are not equilibrium patterns and will tend to degrade over time. There are two possible outcomes as these patterns degrade: Under restricted gene flow, the dispersal and range contraction patterns will approach the equilibrium pattern expected under isolation by distance; if there is no gene flow among disjunct populations, measures of similarity among populations will become random with respect to geographic distance or position.

Given these results, it would appear that gene flow estimates should be plotted against positional parameters as well as distance whenever possible in studies of population structure. In Figure 4.10a, it is evident that plotting $\hat{M}$ against distance may produce a significant regression even when gene flow is not influencing gene frequencies. Plotting the same similarity values against a positional parameter may give a much stronger relationship. Patterns similar to Figure 4.5a are common in the literature (e.g., Slatkin 1993; Alvarez-Buylla and Garay 1994) but other possible relationships among the data are seldom considered. A step-by-step methodology for performing the analysis on

empirical data sets is given in the appendix.

Predictions from the dispersal and range contraction models need to be further validated against empirical data. This could be done in cases where the recent history of a particular species or population is known. Easteal's (1985) studies of genetic variation in *Bufo marinus* would provide an ideal test of the dispersal model because the introduction and dispersal history of these toads in Queensland, Australia is well documented.

Results of the stochastic computer model of population structure closely match theoretical predictions based on coalescent and gene flow theory. The model was used to implement biogeographic scenarios of isolation by distance, dispersal, and range contraction (vicariance). Model results suggest that each of these scenarios leaves a distinct signature on levels of similarity (as measured by F_{ST}) among subpopulations that can be detected by plotting similarity estimates against distance or positional information. The method may be useful in the analysis of isozyme data from disjunct subpopulations when the goal is to distinguish between dispersal and vicariant hypotheses.

LITERATURE CITED

ALVAREZ-BUYLLA, E. R., AND A. A. GARAY. 1994. Population genetic structure of *Cecropia obtusifolia*, a tropical pioneer tree species. Evolution 48(2):437-453.

EASTEAL, S. 1985. The ecological genetics of introduced populations of the giant toad, *Bufo marinus*. III. Geographical patterns of variation. Evolution

39:1065-1075.

EXCOFFIER, L., AND P. E. SMOUSE. 1994. Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: Molecular variance parsimony. Genetics 136:343-359.

FELLER, W. 1957. An introduction to probability theory and its applications. 2d ed., vol. 1. John Wiley & Sons, Inc., New York.

HEY, J. 1991. A multi-dimensional coalescent process applied to multi-allelic selection models and migration models. Theor. Pop. Biol. 39:30-48.

HUDSON, R. R. 1990. Gene genealogies and the coalescent process. Pp. 1-44 *in* D. Futuyma and J. Antonovics, eds. Oxford surveys in evolutionary biology, vol. 7, Oxford University Press, Oxford, UK.

KIMURA, M., AND G. H. WEISS. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49:561-576.

MARUYAMA, T. 1970. Analysis of population structure. Ann. Hum. Genet. 34:201-219.

—. 1971. The rate of decrease of heterozygosity in a population occupying a circular or linear habitat. Genetics 67:437-454.

NEI, M. 1972. Genetic distance between populations. Am. Nat. 106:283-292.

—. 1973. Analysis of gene diversity in subdivided populations. Proc. Nat. Acad. Sci. 70(12):3321-3323.

NEI, M, T. MARUYAMA, AND R. CHAKRABORTY. 1975. The bottleneck effect and

genetic variability in populations. Evolution 29:1-10.

SLATKIN, M. 1987a. The average number of sites separating DNA sequences drawn from a subdivided population. Theor. Pop. Biol. 32:42-49.

—. 1987b. Gene flow and the geographic structure of natural populations. Science 236:787-792.

—. 1991. Inbreeding coefficients and coalescent times. Genet. Res. 58:167-175.

—. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47(1):264-279.

STROBECK, C. 1987. Average number of nucleotide differences in a sample from a single subpopulation: A test for population subdivision. Genetics 117:149-153.

TEMPLETON, A. R., E. ROUTMAN, AND C. A. PHILLIPS. 1995. Separating population structure from population history: A cladistic analysis of geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. Genetics 140:767-782.

WEIR, B. S. 1996. Intraspecific differentiation. Pp. 385-405 *in* D. M. Hillis, C. Moritz, and B. K. Mable eds. Molecular systematics, 2d ed., Sinauer Associates, Inc., Sunderland, MA.

WRIGHT, S. 1943. Isolation by distance. Genetics 28:114-138.

—. 1946. Isolation by distance under diverse systems of mating. Genetics 31:39-59.

—. 1951. The genetical structure of populations. Ann. Eugen. 15:323-354.

APPENDIX

A step-by-step approach to detecting various patterns of population similarity for empirical data sets should proceed as follows:

(1) Estimate pairwise values of $\hat{M}$ and geographic distance for all populations, keeping in mind the effect of scale when determining physical distance. Create a *log-log* scatterplot of these values and determine the least-squares regression equation for the scatterplot. A slope greater than -0.5 suggests that isolation by distance effects are not at work; gene flow levels may be too high, too low, or populations may not be in gene flow/genetic drift equilibrium. A slope less than -1.4 suggests a disjunct pattern of gene flow, possibly resulting from barriers to migration for widely separated groups.

(2) Number populations from one end of the range to the other. If some *a priori* knowledge regarding a center of origin is available, the population closest to the center of origin should be number one. Plot pairwise values of $\hat{M}$ against the lower population number in each pair. Scatterplots similar to those on the right in Figure 4.10 suggest a dispersal scenario. If this pattern is observed, $1/\hat{M}$ should be plotted against the population numbers to obtain an estimate of $\Delta\tau/N$ from the regression.

(3) To detect a range contraction pattern, use the same numbering system as for the dispersal scatterplot, but in this case plot $\hat{M}$ against the larger of the

two population numbers for each pair. Plots similar to those on the right in Figure 4.13 suggest that a range contraction process has been influencing patterns of genetic similarity. Again, $1/\hat{M}$ can then be plotted against population position to obtain estimates of $\Delta\tau/N$.

(4) If none of the expected patterns are obvious from the scatterplots, then populations are either in a transition period from one pattern to another (e.g. from the dispersal pattern to the isolation by distance pattern if gene flow has resumed), or populations have been isolated from one another for a sufficient period of time that patterns of similarity are random with respect to distance or position.

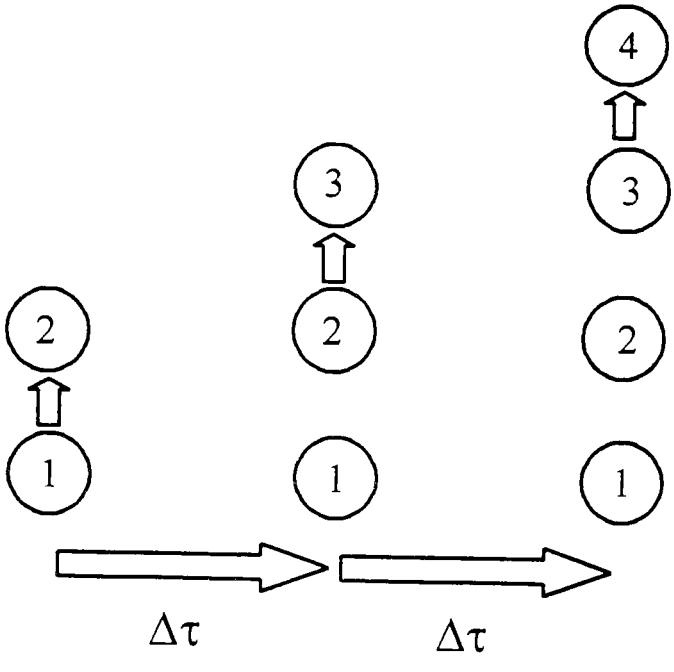

FIG. 4.1. Depiction of the dispersal model. Every $\Delta\tau$ generations, a new subpopulation is founded from the most recently established subpopulation.

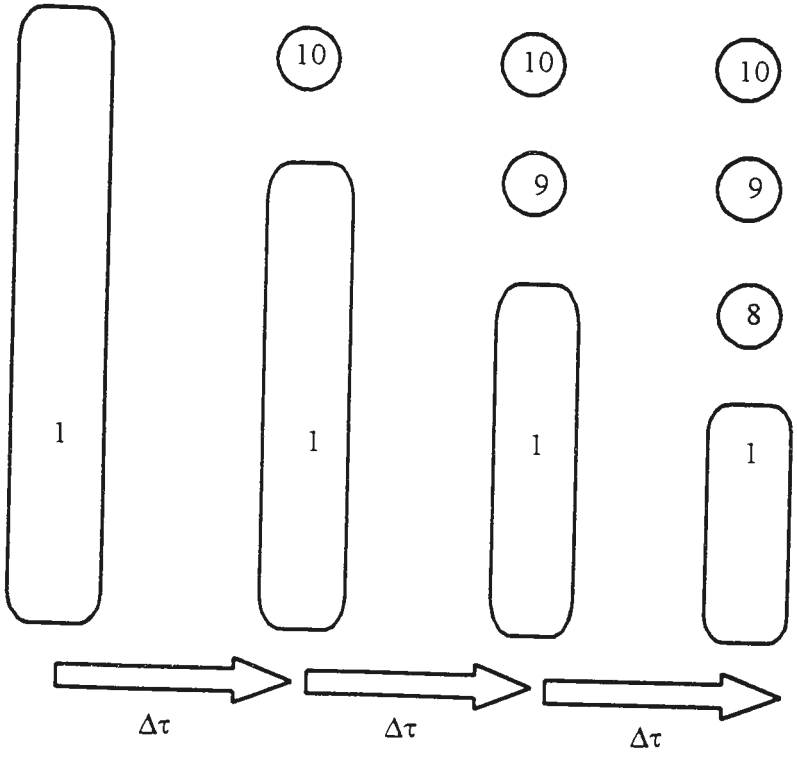

FIG. 4.2. Depiction of the range contraction model. Every $\Delta\tau$ generations, a new subpopulation is founded via fragmentation. In the computer model, the size of subpopulation 1 remains constant despite its smaller range.

FIG. 4.3. Depiction of the disjunct scenario of the isolation by distance model. In this figure, there are two groups: subpopulations 1, 2, and 3 form a group and subpopulations 4, 5, and 6 form a group. Gene flow within groups is high (bold arrows) and gene flow among groups is low (faint arrows).

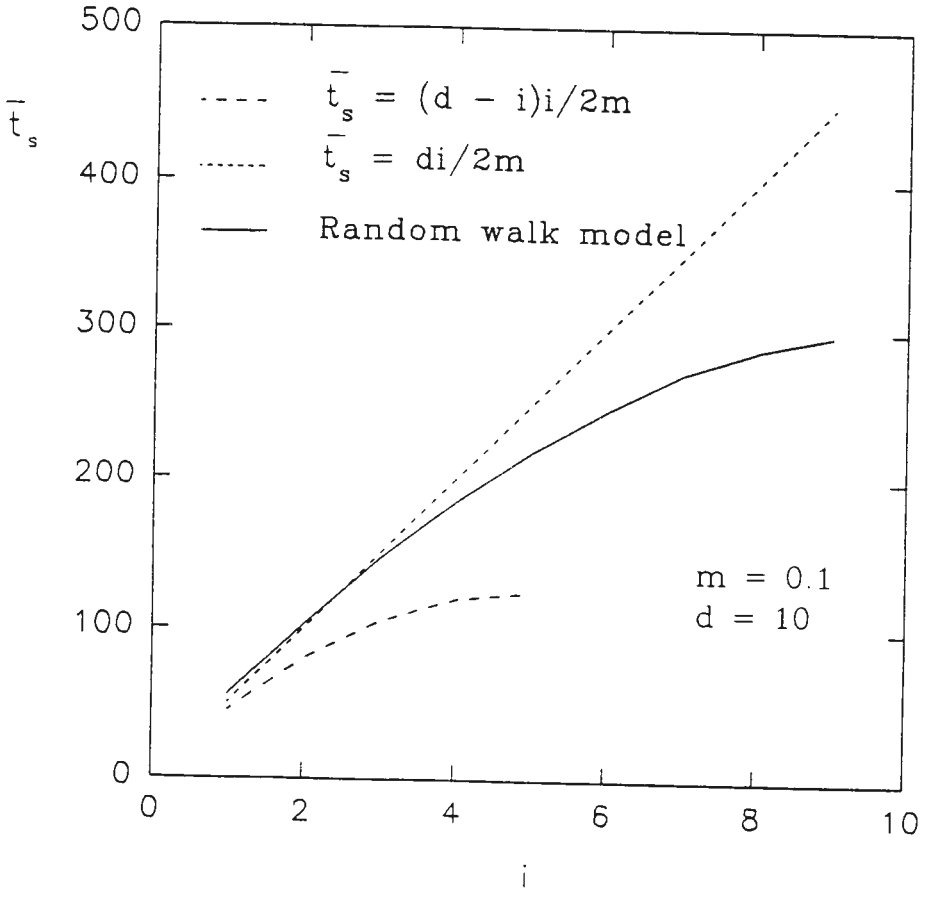

FIG. 4.4. Results from numerical random walk model (solid line) compared with exact and approximate analytical models. $\bar{t}_s$ is the mean time for genes initially separated by *i* steps to arrive in the same subpopulation.

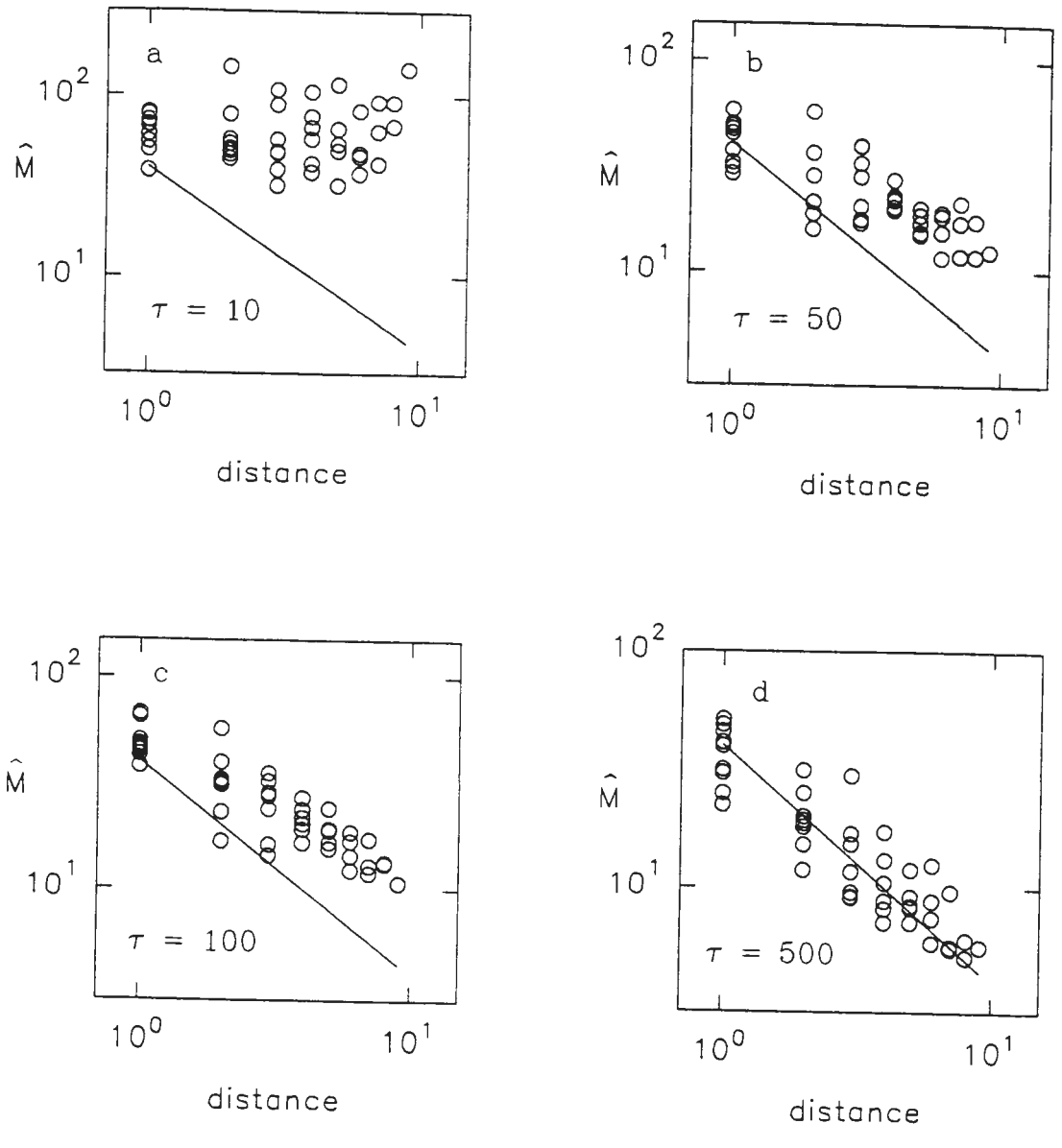

FIG. 4.5. Scatterplots show results of computer implementation of isolation by distance model with an increasing number of generations for each graph (Nm = 20). Solid lines are results predicted by analytical model of isolation by distance equation (6).

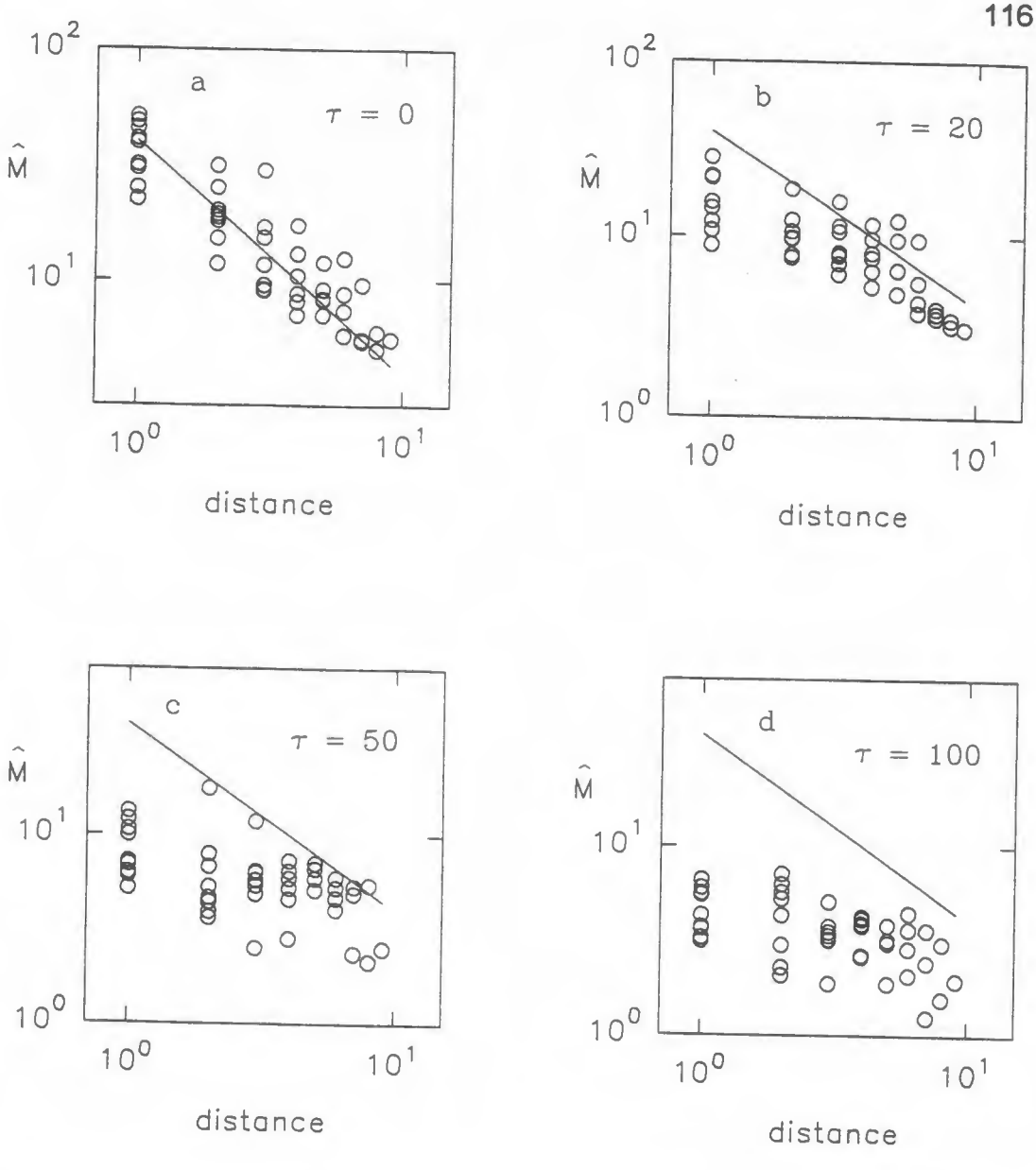

FIG. 4.6. Scatterplots show consequences of interrupted gene flow after an isolation by distance pattern has been established. τ is the number of generations since gene flow was interrupted. Solid lines represent analytical prediction (equation [6]) before cessation of gene flow ($Nm = 20$).

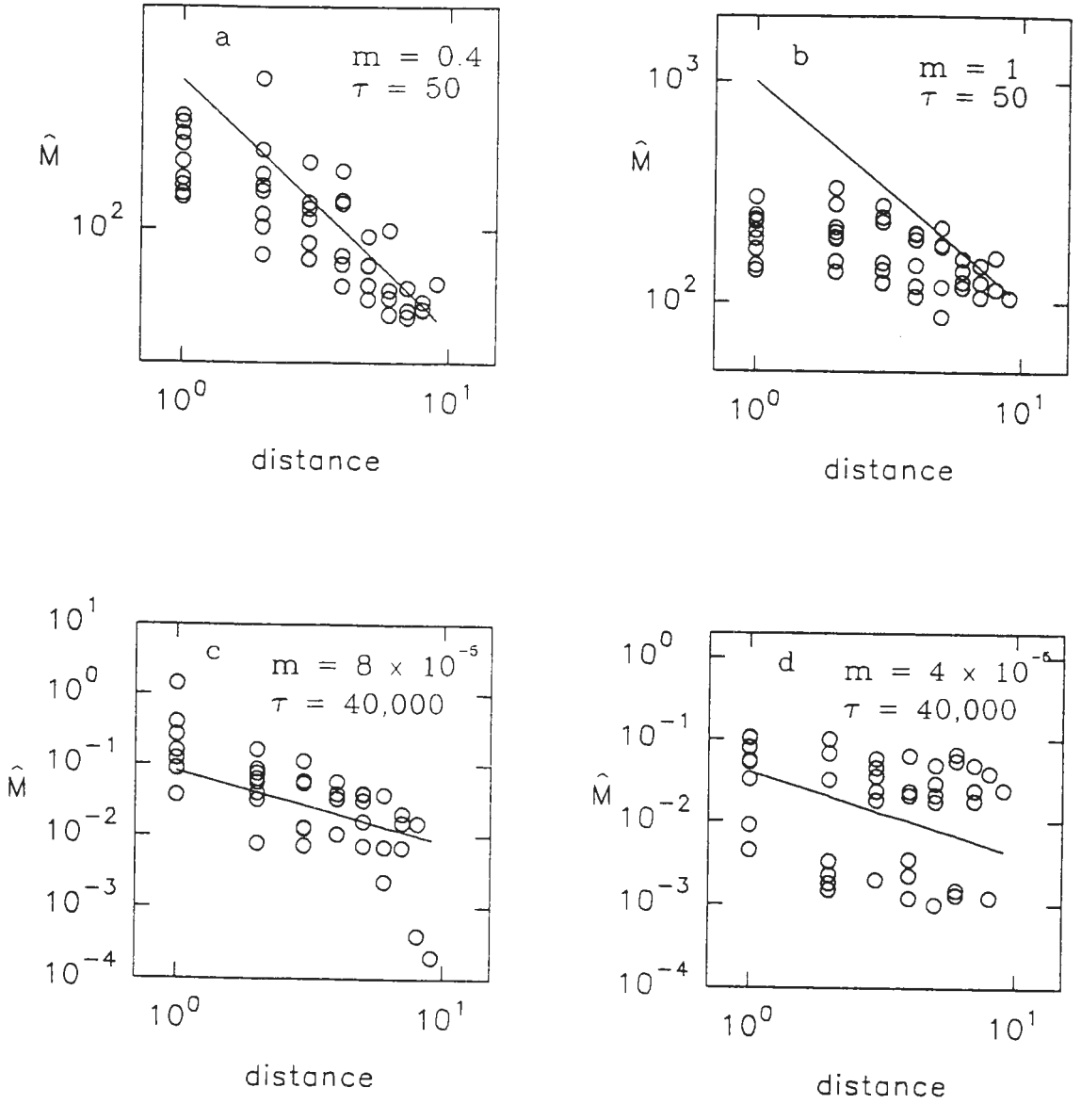

FIG. 4.7. Limits on migration rates for which the isolation by distance pattern is evident for $N = 500$. Isolation by distance pattern breaking down due to high levels of gene flow (a & b) and low levels of gene flow (c & d).

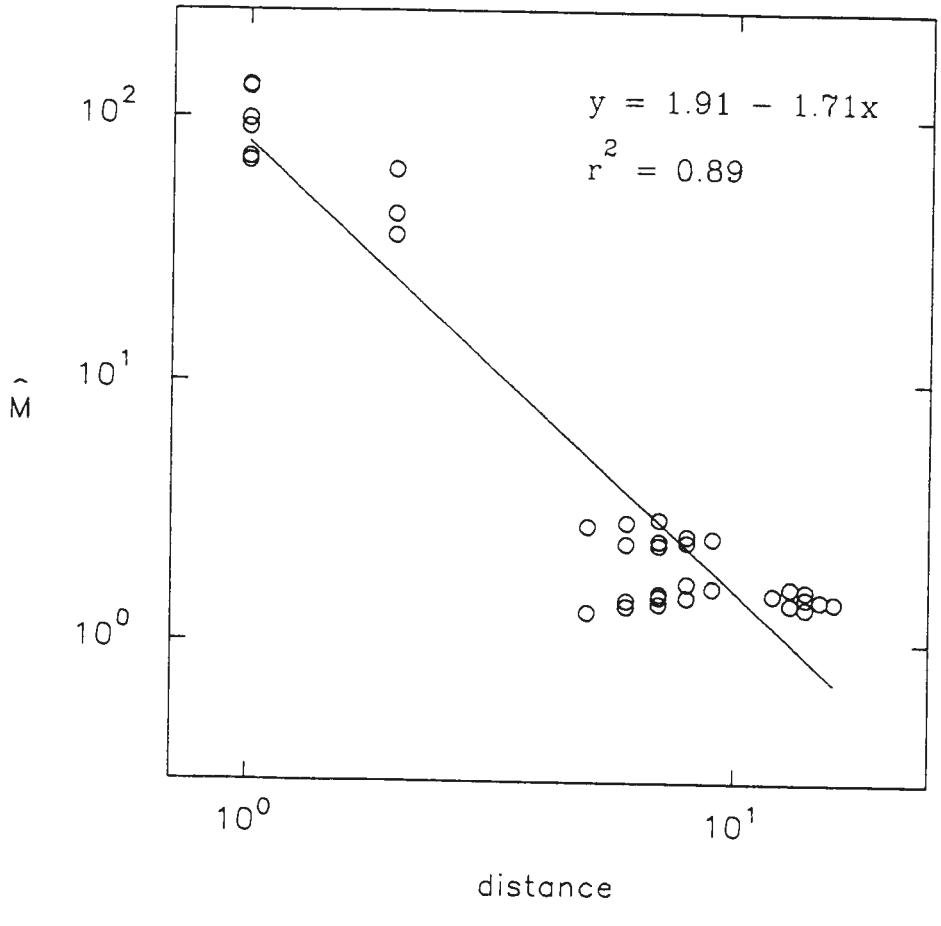

FIG. 4.8. Results of disjunct scenario of isolation by distance model. Nm = 80 within groups and 0.2 between groups, τ = 1000. Groups are separated by five unit steps. Note large negative slope of regression equation (solid line).

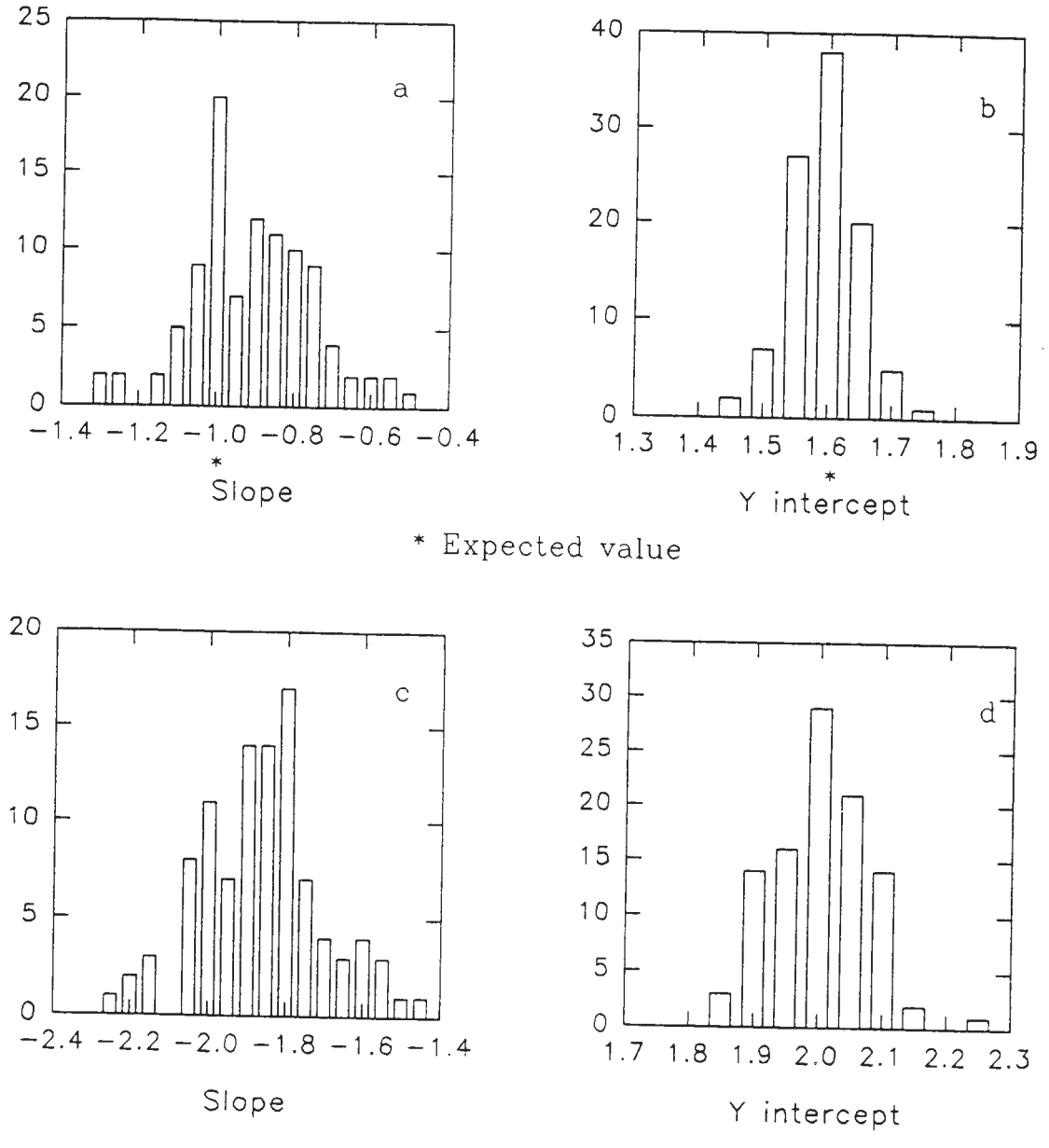

FIG. 4.9. Distributions of slope and intercept values generated from 100 Monte Carlo replicates of isolation by distance model (a & b). * are the predicted values from the analytical model. (c & d) Distributions generated from the disjunct implementation.

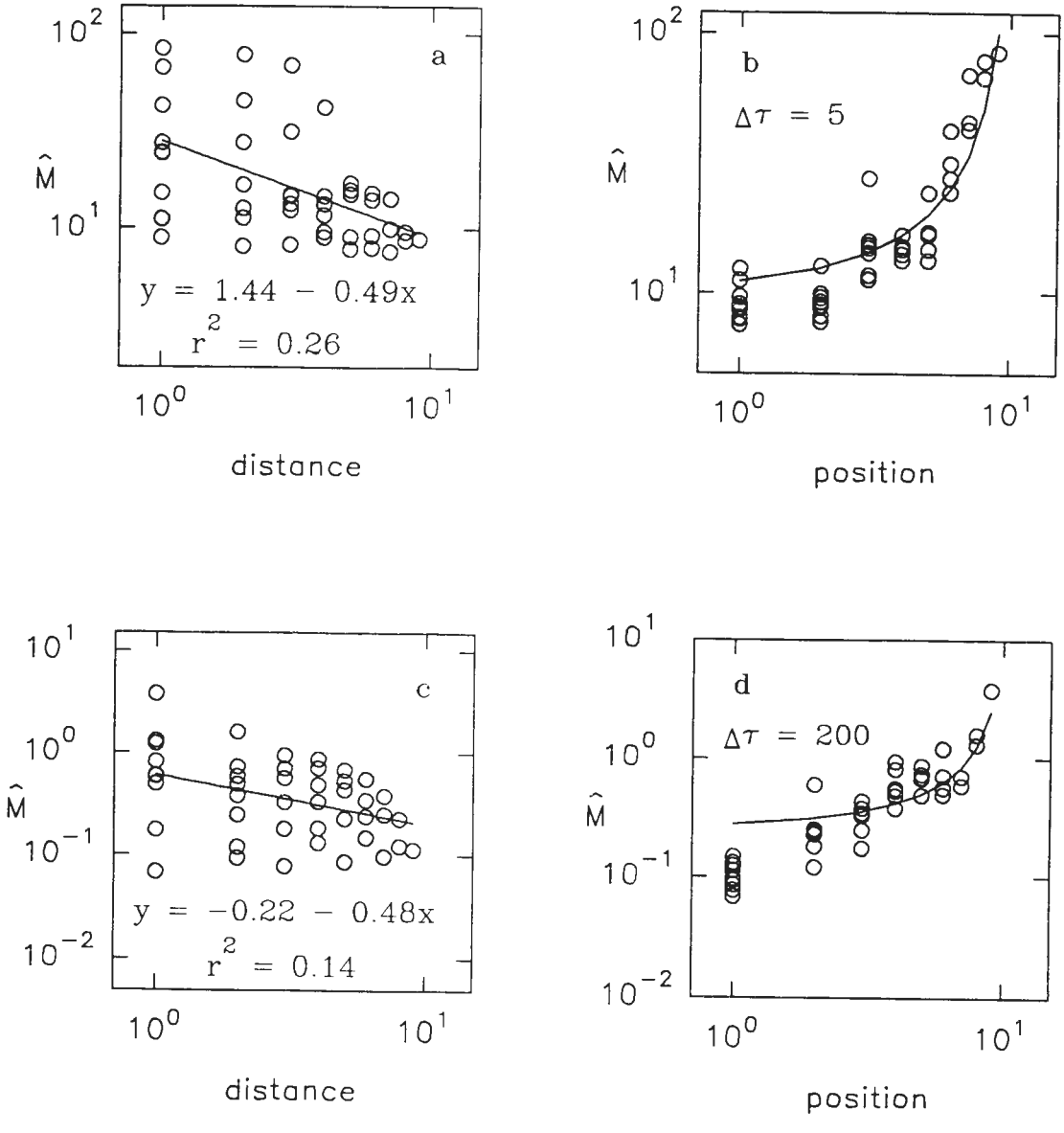

FIG. 4.10. Scatterplots showing results of computer implementation of dispersal model. (a & c) $\hat{M}$ values plotted against distance between subpopulations with corresponding regressions. (b & d) $\hat{M}$ values plotted against the position of the subpopulation; solid lines are predictions from analytical model (equation [7]).

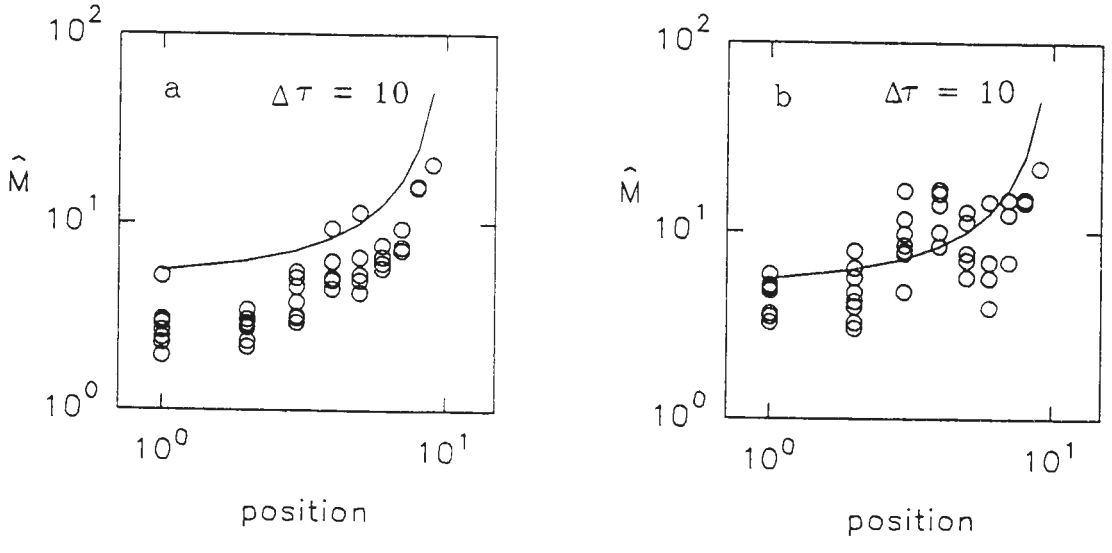

FIG. 4.11. Effect of population bottlenecks during founding events on dispersal model. The predicted dispersal pattern is still evident in the computer results but with $\hat{M}$ values systematically lower than predicted in the absence of bottlenecks. (a) τ = 10; subpopulations founded by sampling 50 individuals from parent population. (b) τ = 10; subpopulations founded by sampling 25 individuals from parent population.

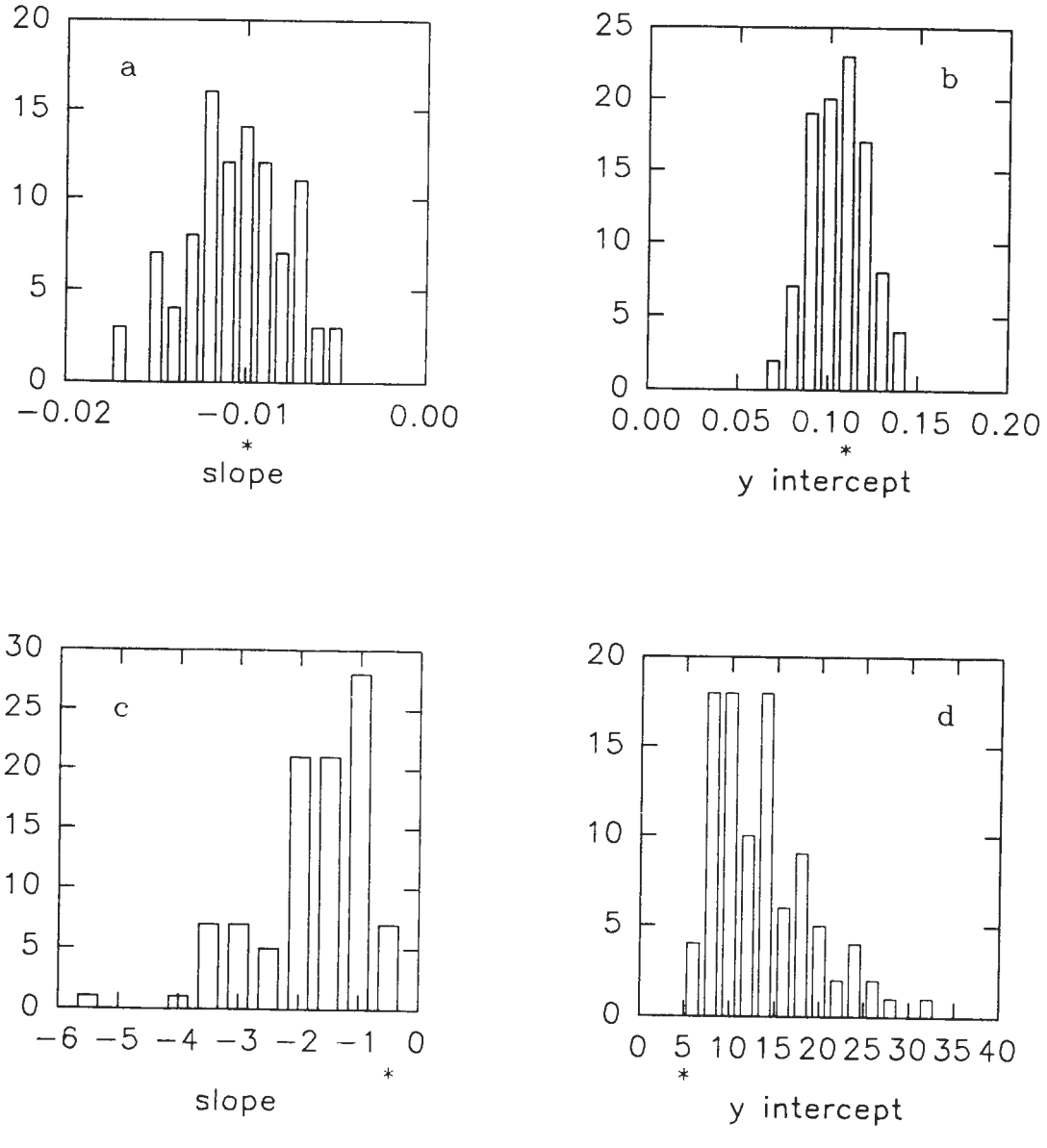

FIG. 4.12. Distribution of slope and intercept values generated from Monte Carlo replicates of the dispersal model. * are predicted values from the analytical model (equation 10). Distributions become biased from predicted results as $\Delta\tau$'s increase. (a & b) $\Delta\tau$ = 5. (c & d) $\Delta\tau$ = 200.

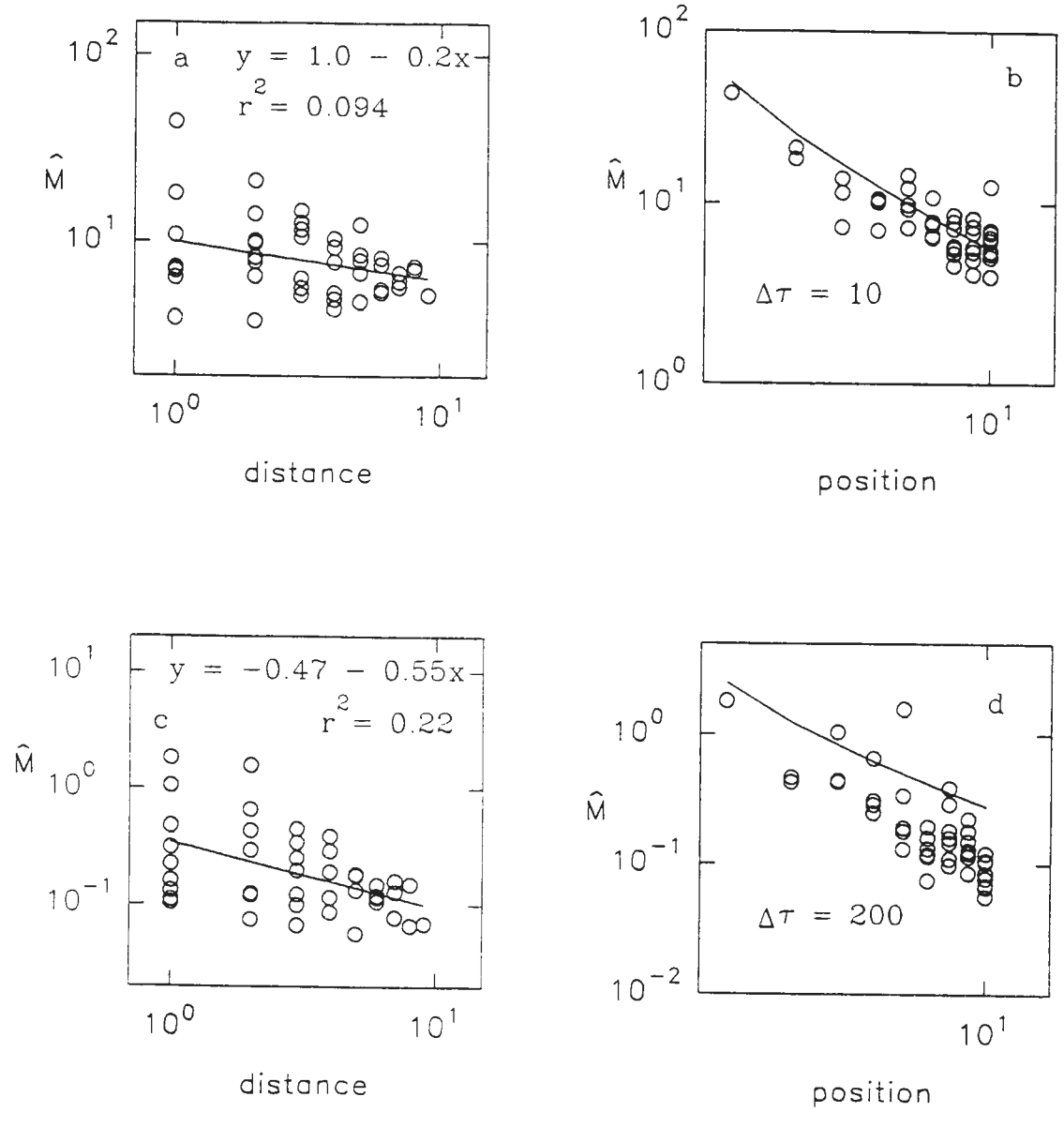

FIG. 4.13. Scatterplots of $\hat{M}$ values vs. distance (a & c) and position (b & d) generated from computer implementation of the range contraction model. Solid lines on left are regressions; solid lines on right are analytical predictions (equation [8]).

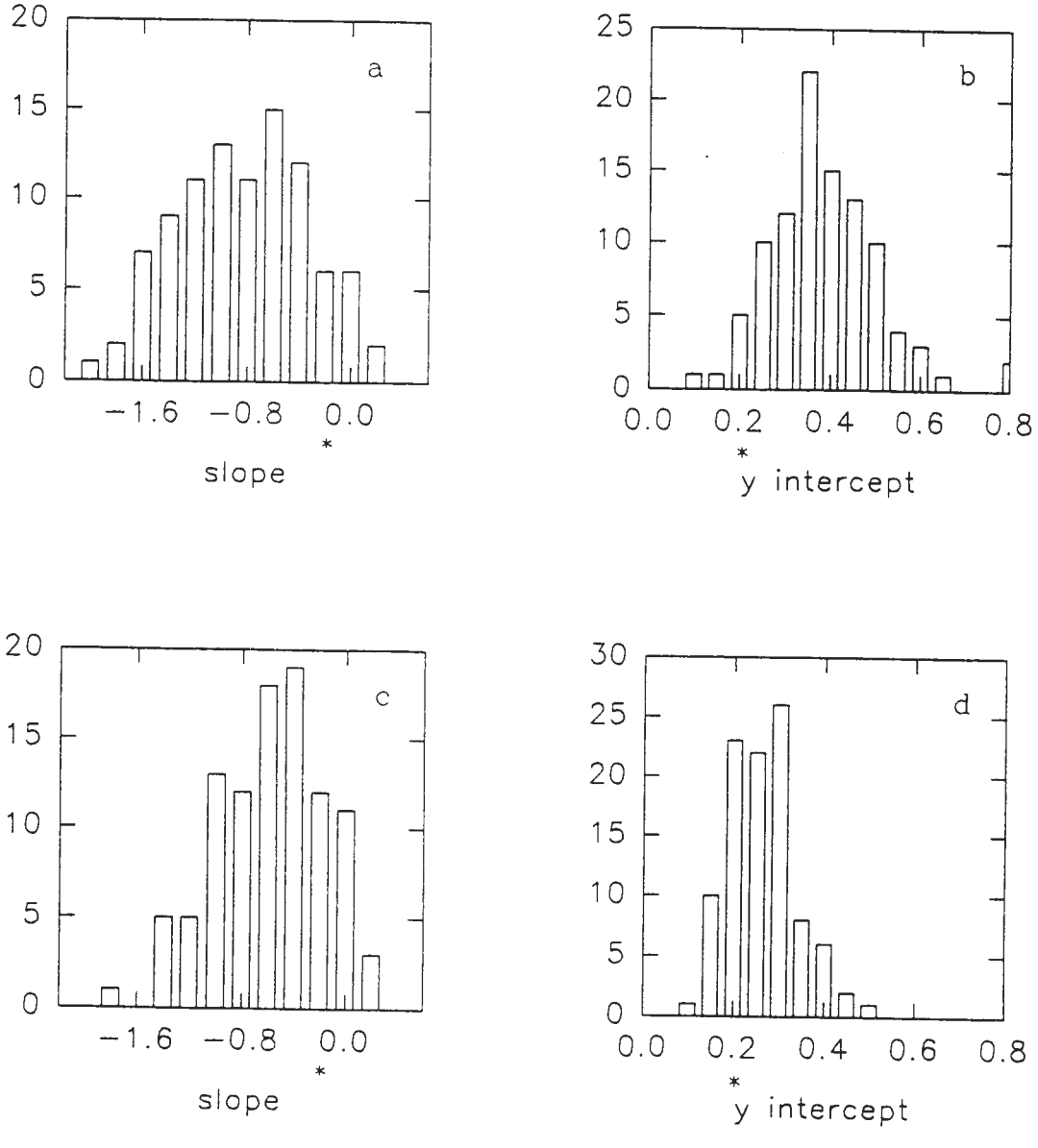

FIG. 4.14. Distributions of slope and intercept values generated from Monte Carlo replicates of the range contraction model, showing effect of large central populations. * are predicted values from the analytical model (equation 11). A large central population brings distributions closer to predicted values. (a & b) τ = 100; central population size = 500. (c & d) τ = 100; central population size = 5000.

CHAPTER 5

SUMMARY

Results of a study of genetic variation at protein coding loci in isolated populations of Canyon Treefrogs (*Hyla arenicolor*) indicate that effective population sizes in this species are influenced to a greater extent by historical factors than by current population numbers. Genetic variation within populations was low in peripheral populations and sharply higher in more centrally located populations. This is to be expected if peripheral populations were recently founded or if peripheral populations are more isolated from potential migrants than central populations. Genetic differentiation among widely separated populations was high, indicating a period of low or non-existent gene flow at the regional scale. In particular, populations in the Canadian River drainage system show divergence levels that are consistent with interspecific levels in other amphibians. Overall, these results are consistent with other studies that have examined the partitioning of genetic variation in natural populations in that the importance of preserving independently evolving demes for the maintenance of genetic diversity is demonstrated.

A computer model of gene flow and geographic structure was used to show that observed patterns of genetic similarity among populations of *H. arenicolor* could be a result of a disjunct scenario of isolation by distance.

Under this scenario, adjacent populations within local drainage systems are subject to moderate levels of gene flow, whereas populations in widely separated or disjunct drainages exchange genes at a disproportionately low rate. Similarity patterns among populations of *H. arenicolor* do not conform to patterns predicted from the computer model for either a dispersal or a vicariant model of population isolation. However, the expected patterns of genetic divergence for dispersal and vicariant scenarios are transient and will degrade over time. Thus, the absence of these patterns suggests only that these events have not occurred recently. Given the limited dispersal ability of Canyon Treefrogs based on their behavior and physiology, and the geologic evidence for shifting drainage patterns in the southwest region, it is likely that vicariant events have played a substantial role in the establishment of disjunct populations of *H. arenicolor*. This suggests that populations of *H. arenicolor* have been in place for a period of time sufficient to alter the genetic signal of population establishment. Gene flow among populations appears to be the main factor influencing patterns of similarity at the local scale, while the time since populations shared a common ancestor is likely to be the critical factor establishing patterns of similarity at the regional scale.

Mitochondrial DNA sequence variation among populations of *H. arenicolor* is somewhat contradictory to isozyme results. MtDNA results support the major division between northwestern and eastern clades indicated by isozyme analysis; however, the sequence data do not support high levels of

differentiation among populations within the eastern clade. Among these populations, overall sequence divergence is low and the number of shared character states is insufficient to further resolve the clade. This outcome is expected if the eastern clade was founded relatively recently during a period of rapid range expansion or if a moderate level of gene flow has been maintained among populations. The latter explanation is possible for populations in southern Arizona and New Mexico given the moderate rates of gene flow estimated from isozyme data for these populations. The Canadian River population is an anomaly in this sense because isozyme results suggest that it exists in extreme isolation.

The most important result of the mtDNA analysis is the possibility of mtDNA of hybrid origin in two populations of *H. arenicolor* in southern Arizona. Sequences from the Madera Canyon and Marijilda Creek populations show a level of divergence that places them midway between other populations of *H. arenicolor* and the outgroup *Hyla versicolor.* This level of divergence is difficult to explain without invoking a hybrid origin for the divergent haplotypes. The most likely source for these haplotypes is the species *Hyla eximia*, which has a similar distribution to *H. arenicolor* in Mexico. Confirmation of the hybridization hypothesis will require an analysis of mtDNA sequence data from *H. eximia*.

According to our current understanding, *H. arenicolor* and *H. eximia* have distinctly different breeding strategies. *H. arenicolor* normally breeds according to time of year, with breeding in most populations running from late

April to mid-June. In contrast, *H. eximia* is a synchronous breeder, cued in to mid to late summer rainfall events. Thus, natural hybridization between these two species would require polymorphic breeding behavior in at least one of the species. This would represent an interesting system for future study in that intraspecific differences in breeding behavior can represent a prelude to speciation. Given regional variation in weather patterns and moisture availability, these frogs are likely to experience selection pressures for optimal breeding strategies.

CURRICULUM VITAE

ROY A. MURRAY
April, 1997

Address:
64 North 200 West
Logan, Utah 84322-5305
Telephone:
(801) 753-1714 (office)
(801) 753-0832 (home)
Internet:
roy@harvestmaster.com

DOB: April 11, 1959
Ithaca, New York

EDUCATION

1977-1982: Georgia Institute of Technology, Atlanta, GA. Bachelor of Electrical Engineering, 1982.
1987-1988: University of Alabama in Huntsville, Huntsville, AL. Post-graduate Studies in Biology.
1989-present: Utah State University, Logan, UT.
Ph.D. Candidate: Population Genetics
Major Professor: Paul G. Wolf
Dissertation Title: *Evolutionary Genetics of Canyon Treefrogs* (Hyla arenicolor).
Interdisciplinary Program in Ecology

EMPLOYMENT

1996-present: Research Scientist, Juniper Systems, Logan, UT.
1991-1996: Teaching Assistant, Department of Biology, Utah State University (USU), Logan, UT.
1994-present: Consultant, Prince William Sound Science Center (PWSSC), Cordova, AK.
1990-1993: Consultant, Omnidata International, Logan, UT.
1989: Systems Engineer, System Dynamics Incorporated (SDI), Huntsville, AL .
1983-1988: Research Engineer, Georgia Tech Research Institute (GTRI), Huntsville, AL.
1979-1981: Cooperative Student, Georgia Institute of Technology, Atlanta, GA.

AWARDS AND HONORS

1997: Phi Kappa Phi Honorary Society, Utah State University
1992: President's Fellowship, Utah State University ($8,500).
1992: Distinguished Service Award, Associated Students of Utah State University.
1989: Graduate Fellowship, Utah State University ($7,000).
1989: Theodore Roosevelt Memorial Fund, American Museum of Natural History.
1978: Phi Eta Sigma Honorary Society, Georgia Institute of Technology.

SOCIETY MEMBERSHIPS

Society for the Study of Evolution
The Herpetologists' League

PROFESSIONAL EXPERIENCE

Global Positioning System (GPS) Interface Design (Juniper Systems): Currently designing and coding a user interface for a GPS handheld computer to be called FieldScout™. FieldScout™ will be used in both research and business applications for field data collection, including the capture of accurate position location information from the Navstar Global Positioning System.

Wireless Networking (PWSSC): Tasked to design and implement a wireless data communications network for the Prince William Sound Science Center. The system is being used for transmittal of information to and from remote data collection platforms as part of the Sound Ecosystem Assessment (SEA) project ongoing at PWSSC.

Hardware/Software Design and Testing (Omnidata International): Wrote custom data collection software for hand-held computers and dataloggers. Designed and manufactured a buffered data collection interface and driver software for Calsonic environmental control unit installed in Nissan automobiles.

Digital Simulation and Modeling (GTRI): Involved in modeling of weapon system capabilities and performance. Developed environmental models as they apply to missile and radar system performance, including effects related to monostatic and bistatic clutter, multipath, and terrain reflectivity.

Data Acquisition (GTRI & SDI): Project Director for program to record bistatic terrain reflectivity data using specially instrumented missile seeker. The project was conducted in cooperation with the 3246th Test Wing at Eglin Air Force Base. Developed software for extraction and analysis of targeting/engagement data from Pedestal Mounted Stinger tests conducted

at Ft. Hunter-Liggett Test Range.

Test Evaluation (GTRI): Project Director for program to evaluate test procedures as proposed by MIT Lincoln Labs for collection of clutter data. Test involved firing of specially instrumented HAWK missiles.

Radar Modeling (GTRI): Project Director for continued development of an existing HAWK HPI radar model.

Computers: Extensive experience on mainframe, mini-, and microcomputers. Familiar with VMS, UNIX, Windows and DOS operating systems. Fluent in Pascal, C, FORTRAN, BASIC, and assembly computer languages.

REPORTS AND PUBLICATIONS

Murray, R.A. and P.G. Wolf. Submitted to Herpetologica. An Electrophoretic Analysis of Genetic Variation and Population Structure in Canyon Treefrogs (*Hyla arenicolor*).

Wolf, P.G., R.A. Murray, and S.D. Sipes. Submitted to American Journal of Botany. A molecular test of secondary intergradation in hybrid zones of *Ipomopsis*.

Murray, R.A. and P.G. Wolf. Submitted to Molecular Phylogeny and Evolution. Phylogeny and Biogeography of Canyon Treefrogs (*Hyla arenicolor*) based on mitochondrial DNA sequences.

Murray, R.A. In preparation. Genes and Biogeography: Patterns of Population Divergence under Isolation by Distance, Vicariance, and Dispersal.

Murray, R.A. 1990. Ecology of Baird's Tapir in the Monteverde Cloud Forest Reserve. Final report submitted to the American Museum of Natural History for grant in aide of research.

HAWK Radar Systems Engineering Support, Final Report, Contract DAAH01-87-D-0082, Delivery Order 0018, GTRI Project A-4810, January 1988.

Clutter Analysis and Modeling, Final Report, Contract DAAH01-83-D-A013, Delivery Order 0117, GTRI Project A-4591, April 1987.

HAWK Clutter Measurements/Effects Program Support,Final Report, Contract DAAH01-83-D-A029, Delivery Order 0049, GTRI Project A-4257, March 1986, coauthor.

Phase III PIPs, Final Report, Contract DAAH01-83-D-A013, Delivery Order 0095, GTRI Project A-4075, May 1985, coauthor.

Clutter Measurements/Effects Program Support, Final Report, Contract DAAH01-83-D-A013, Delivery Order 0033, GTRI Project A-3628, October 1984, coauthor.

HAWK Continuous Wave Acquisition Radar (CWAR) System Simulation

Support, Final Report, Contract DAAH01-83-D-013, Delivery Order 0063, GTRI Project A-3800, November 1984, coauthor.

Effects of Bistatic Clutter on Low Altitude Air Defense Missile Systems, Milestone Report, Contract DAAH01-81-D-A003, Delivery Order 007, GTRI Project A-2847, January 1981, contributor.