12-1-2009

The role of spatial and genetic modeling to biogeography.

Dolly Crawford

Dolly Lyne Crawford

Candidate

Biology

Department

This dissertation is approved, and it is acceptable in quality
and form for publication:

Approved by the Dissertation Committee:

_____ Felisa A Smith _____, Chairperson

_____

_____

_____

_____

_____

_____

_____

i

The role of spatial and genetic modeling to biogeography.

by

Dolly Lyne Crawford

B.A. Biology, Mansfield University of Pennsylvania, 1992
M.S. Museum & Field Studies, University of Colorado, 2003

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Ph.D. Biology

The University of New Mexico
Albuquerque, New Mexico

December, 2009

The role of spatial and genetic modeling to biogeography.

by

Dolly Lyne Crawford

ABSTRACT OF DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Ph.D. Biology

The University of New Mexico
Albuquerque, New Mexico

December, 2009

THE ROLE OF SPATIAL AND GENETIC MODELING TO BIOGEOGRAPHY.

by

DOLLY LYNE CRAWFORD

B.A., Biology, Mansfield University of Pennsylvania

M. S. Museum and Field Studies, University of Colorado at Boulder

Ph.D. Biology, University of New Mexico

ABSTRACT

One of the most persistent challenges in biology is explaining the distribution of animal taxa. Quantitative explanations for the distribution of organisms are challenged by the complexity of factors that potentially limit a species range, including topography, ecology, climate and biology. Because of limits on data and approaches, early biogeographers were hampered in their ability to explain patterns. Here, I benefit from two primary developments in biogeography. My research draws from the accumulation of data across several scientific disciplines and advances in spatial and statistical approaches to examine the distribution of animal taxa within an integrative biogeography framework. I combine ecological, climatic and genetic data and analyses to address these primary issues with biogeographical distributions: 1, reconstructing the colonization history of taxa,, 2, distinguishing between Pleistocene source and Pleistocene refugial populations, 3, understanding how distribution constraints influence population connectivity and the evolutionary potential of species populations and 4, modeling the ecological and genetic parameters that most influence the extirpation of local populations.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter I

INTRODUCTION

The dynamic influence of tectonics, climate, topography and other factors form the biotic realm of life on Earth. This complex physical template is matched by the diversity of ecosystems: from dry rocky deserts to tropical cloud forests. Explaining the distribution of organisms has a long and active history, from the early explorations of Carolus Linnaeus and Comte de Buffon to current work in some of the remote places on earth, such as Antarctica (Peat et al., 2007). Throughout this rich history, the common thread that ties most of these people and places together is also one of the most challenging and long standing questions in biology: why are organisms found where they are?

In the 17th and 18th century, diversity of life was thought to be the result of divine creation and as such, species could not fundamentally change. While the immutablity of species was the predominate explanation offered by scientists such as Georges Cuvier, many nineteenth century researchers sought a more mechanistic explanation for the distribution of species on earth. By that time, the observations and biological collections made by seafaring explorers were substantial. As such, it was clear that organisms were distributed throughout many habitats on earth and presented a challenge to explain. A more dynamic view of the earth and its species came into sharper focus after the publication of Charles Lyell's *Principles of Geology* in 1830. The dispersalist explanation was a consequence of this new view of a dynamic earth but seemed especially limiting to

1

Leon Croziat, who was dissatisfied with his dispersalist explanation for the distribution of a fossil seed fern, *Glossopteris*.

Croziat's difficulty with *Glossopteris* is an early illustration of the utility of additional, independent data to explain the distribution of organisms. Although he could not have realized it at the time, a more acceptable explanation for the distribution of *Glossopteris* promoted after his death was the result of integrating the principles of continental drift as advanced by Alfred Wegener (Wegener, 1966). This more synthetic explanation was possible because of an accumulation of biological and geological data. This is one of the earliest examples of integrative biogeography. Increasingly, biogeographic studies incorporate an integrative approach. Modern integrative biogeography is a consequence of many factors, including the accumulation of data on the past and present distribution of taxa, spatial approaches such as GPS, remote sensing and GIS, and the advent of molecular genetics (Riddle et al., 2008).

The focus of my dissertation research at the University of New Mexico is explaining the distribution of mammal lineages at high elevation. Like Croziat, I was faced with a challenging task of explaining the distribution of animal taxa that are embedded in a complex matrix of topography, climate and biotic interactions. However, unlike Croziat, I can draw from the accumulated data on the biophysical setting, and use the data to test hypotheses using multiple, independent lines of evidence. My solution to the problem of high elevation mammal distributions is the essence of an integrative biogeographical approach. I apply this approach in each of four chapters.

The distribution of the Mexican vole, *Microtus mexicanus* is the focus of chapter one. The species occupies a mountainous distribution throughout the southwestern United

States and Mexico. Despite the number of mammal taxa that appear to share this distribution, few studies have examined the distribution from a vicariance/dispersal perspective. I incorporate an ecological niche modeling and phylogeographical approach to reconstruct a plausible colonization and diversification scenario for the species. My analyses of the genetic variation in the Mexican vole include an examination of nuclear and mitochondrial DNA. My study highlights the influence of climate on the past and present distribution of organisms, and indicates the potential for future research in the biologically diverse region.

In his treatise on non-equilibrium dynamics, Brown (1971) suggested that the extant distribution of mammals in the Great Basin was the result of initial colonization from the Sierra Nevada or Rocky Mountains. The primary motivation for Chapter two is test Brown's hypothesis using an integrated molecular phylogenetic and simulation approach. To test the hypothesis, I examine the distribution of the montane vole, *Microtus montanus*. My study examined the genetic variation within and between populations of the montane vole using mitochondrial and nuclear DNA. The species is ideally suited to testing the hypothesis because its geographical distribution includes the Sierra Nevada and Rocky Mountains and a significant portion of the Great Basin. My analyses suggest support for the Sierra Nevada Mountains as the source region for the initial colonization of the Great Basin by the montane vole. This study represents one of only a few studies into the origins of Great Basin mammal taxa, and suggests the need for additional studies to examine concordance across mammal taxa in the Great Basin.

An understanding of the limits to a species distribution is fundamental to biogeography. Chapter three is an examination the distribution constraints in the montane

vole (*Microtus montanus*), and how the constraints influence the level of connectivity between populations in the Great Basin. I test the hypothesis that populations of the montane vole along the White River in the Great Basin are peripheral isolates using mitochondrial and nuclear DNA. I also conducted several simulations of the nuclear data to mitigate for small sample sizes. My analyses suggest a north to south cline in genetic diversity parameters: with northern populations exhibit greater genetic variation compared to southern populations. Overall, my results suggest that the connection between populations during the Pleistocene was severed in the late Pleistocene to middle Holocene, which precipitated the formation of peripheral populations along the southern end of the White River. My results also suggest the importance of simulations to test for adequate sample size, an issue that is not addressed in all studies of population connectivity.

One of the most challenging and timely questions in biogeography is understanding the biotic and biotic factors that promote the extinction of species. Over the course of my field surveys, I documented the extirpation of a population of the montane vole (*Microtus montanus*) at Ash Meadows in the Amarogosa Valley of southwestern Nevada. This survey result was later verified by additional independent surveys by the Nevada Division of Wildlife and private biological monitoring firms. The impetus for Chapter four is to understand the risk factors that influenced the extirpation of this population, in an effort to conserve remaining populations in the region. I conducted hypothesis testing within an integrative framework in which I examined abiotic, genetic and additional biotic and demographic data. My analyses suggest a

potential scenario that includes the action of post-Pleistocene habitat contraction and

recent anthropogenic habitat change as causal agents in the extirpation of the population.

Results of this study provide a potentially important baseline for understanding the

thresholds for population viability, and a timeframe for population response when the

thresholds are surpassed.

Chapter 1

Tracing the colonization and diversification dynamics of the Mexican vole, *Microtus*

mexicanus in the southwestern United States and Mexico.

Several mammal species have a discontinuous distribution across the mountains of the

southwestern United States and Mexico. Despite the exceptional level of mammal

diversity in this region, few studies have examined the diversification and colonization

dynamics of taxa found in this region. Such studies have the potential to elucidate the

processes that generated this diversity. Here I focus on one member of this community,

the Mexican vole (*Microtus mexicanus*) that inhabits high elevation grassland and pine

forests. I examined two different hypotheses proposed to explain the current distribution

of the taxon: colonization into the southwestern United States after the Last Glacial

Maximum, or late Pleistocene colonization into the southwestern United States with

subsequent diversification of lineages in Mexico. I assessed support for the hypotheses by

utilizing multiple lines of evidence including ecological niche modeling and

phylogeographic analyses. Ecological niche models representative of climate conditions

for the Last Interglacial (c. 0.12-0.14Ma) indicated that the species may have colonized

the Sierra Madre of Mexico, and the northwestern Mogollon Rim of the southwestern

United States by the Last Glacial Maximum (c. 0.015Ma). A final model of current

climate conditions indicated that more widespread colonization of the southwestern

United States was likely only after the Last Glacial Maximum. This recent colonization

of the southwestern United States is also supported by molecular analyses, which

demonstrated shallow branch lengths, late Pleistocene to early Holocene divergence

estimates and unimodal haplotype distributions. Deeper branch lengths, mid-Pleistocene divergence estimates and multimodal haplotype distributions also indicate the influence of vicariance on the Mexican vole distribution. My study highlighted the influence of both vicariance and colonization to the current distribution of the Mexican vole, which acted at different times in different regions.

INTRODUCTION

The mixed pine-oak woodlands of the mountainous southwestern United States and Mexico are biologically diverse communities (Eisenberg, 1981), and are threatened by human mediated habitat loss (Mittermeier et al., 2005) and global climate change (Peterson et al., 2002). Much of the region's vertebrate diversity is due to the number of rodent species that are distributed across the mountainous domain of the southwestern United States and Mexico (Ceballos and Oliva, 2006). Of the greater than 30 rodent species found in the region, at least half are restricted to pine-oak forest at high elevations. Despite the exceptional level of biodiversity, few studies have examined vertebrate diversification and colonization dynamics (Sullivan et al., 2000; Leon-Panigua et al., 2007; McCormack et al., 2008). Such data are needed to understand the processes that promote and maintain vertebrate diversity in the region.

Several models of biotic diversification are offered to explain the regions' high endemic biodiversity, including variation in latitude, habitat, climate or orogeny (Fa and Morales, 1993; Coates and Obando, 1996). The most widely accepted explanation invokes a Quaternary climate-induced desiccation of lowland habitats, which drove cool adapted species further upslope, promoting their isolation and eventual diversification (Vrba, 1993; Cracraft and Prum, 1988; Patterson, 1982). While such a vicariant model may explain diversity within some taxa (e.g., *Sciurus abert*i, Lamb et al., 1997), it does not offer a complete explanation for all vertebrates found in the region. Moreover, a purely vicariant hypothesis neglects the influence of colonization, which may be substantial. It is essential that approaches are developed that can simultaneously infer the

potential role of both vicariance and colonization in order to fully explore the factirs that led to the current levels of biodiversity in Mexico and the southwestern United States.

To date, most studies of biodiversity of the region rely on current or past species distribution records to frame biogeographic hypotheses (e.g., Lawlor, 1998; Davis et al., 1988). Two assumptions potentially limit the conclusions reached using such an approach; namely that the distribution of a species remains fixed over time, and that species distributions are evenly and adequately sampled. In reality, these assumptions are rarely met, which can confound interpretation. Clearly, there is a need for additional, independent measures of colonization and extinction parameters (Lamb et al., 1997; Fa and Morales, 1993; Lomolino et al., 1989).

The application of molecular genetics has revolutionized biogeography (Riddle et al., 2008) by providing an additional, independent source of information about species colonization and diversification dynamics. For example, genetically distinct clades identified using a phylogeographic approach can be constrained with divergence estimates and compared with paleoecological, geological or climatic data to produce a more synergistic scenario of species history. The cytochrome *b* gene of the mitochondrial DNA (cyt*b*) is commonly used to examine variation within and between species because of its relatively rapid coalescent time and ease of amplification. Of course, an inherent problem with the single-gene approach is distinguishing gene trees from species trees (Avise, 2000). While the low mutation rate, larger effective population size and longer coalescence time of nuclear genes currently used in higherlevel systematic studies (Moore, 1995; Vawter and Brown, 1986) may limit their use in resolving recent divergences, the higher mutation rate found in non-coding nuclear

sequences are potentially useful in assessing clade structure among closely related taxa or populations.

Although genetic information provides a way to correlate geography with the distribution of lineages, ecological data is also potentially vital to the reconstruction of a species history. More recent approaches in geographic information systems, and ecological niche modeling, in particular, contribute to methods that link patterns of genetic variation with abiotic factors, such as past or present climatic conditions (Swenson, 2008; Knowles et al., 2007; Hugall et al., 2002). In this capacity, ecological niche models represent *a priori* hypotheses about the distribution of lineages at different time periods that can be tested with genetic data (Hugall et al., 2002). This approach has utility in examining the relative role of vicariance and colonization in the montane mammal fauna distributed across Mexico and the southwestern United States.

The montane rodent species that are distributed across the southwestern mountains and Mexico include the Mexican vole, *Microtus mexicanus*. The disjunct distribution of the Mexican vole includes grasslands from the Colorado Plateau and the Sky Islands of the southwestern United States into southern Mexico (Frey et al., 2002; Hall, 1981; Armstrong, 1972; Findley and Jones, 1962). The distribution of the Mexican vole suggests three diversification hypotheses (Fig. 1). One hypothesis posits the role of post-Pleistocene climate change beginning about 0.015Ma as a vicariant agent that separated the populations in Mexico from the United States, followed by recent expansion in the United States (Davis and Callahan, 1992; Lomolino et al. 1989; Davis et al., 1988). Alternatively, the Mexican vole may have colonized the United States from the Sierra Madre beginning about 0.50 Ma and diversification of the southern populations

occurred during the Sangamon Interglacial, between 0.38-0.16Ma (Frey, 1989). Finally, populations of the Mexican vole may be the product of recent widespread range expansion. All of the hypotheses form testable predictions about the relationship between northern and southern populations of the Mexican vole (Figure 1). There remains, however, a surprising lack of consensus on the colonization and diversification dynamics of the Mexican vole.

Here, I analyze the ecological niche and genetic diversity of the Mexican vole to examine its potential colonization history. I assessed the likelihood of each colonization hypothesis by constructing ecological niche models (ENM) representative of climate conditions before and after the Last Glacial Maximum (LGM). I assessed whether the distribution of geographical lineages matched the expectations of the ENM through analyses of the cytochrome b gene of the mitochondrial DNA (cyt$b$) and the nuclear intron, the acid phosphatase V (AP5) gene. I further assessed the likelihood of each colonization hypothesis by constraining the tree topology to match the expectations of each hypothesis. The constrained topologies included: (a) expansion into the southwestern United States about 0.015 Ma, and (b) expansion into the southwestern United States from the Sierra Madre and the diversification through vicariance of southern populations between 0.38-0.16 Ma, and (c) a null hypothesis where rapid, relatively coincident expansion occurred throughout the region.

MATERIALS AND METHODS

Ecological modeling

The primary difference among the colonization hypotheses is the timing of range expansion into the United States. I first examined the support for the hypotheses (Fig. 1)

11

within an ecological niche modeling framework. The first hypothesis (Fig. 1a) suggested that northward expansion of the Mexican vole into the United States occurred after the LGM, circa 0.015Ma (Davis and Callahan, 1992). An underlying assumption is that suitable habitat for dispersal between Mexico and the United States was present over that time interval. The second hypothesis suggests that the Mexican vole entered the United States from the Sierra Madre Oriental (SMOr) beginning about 0.5Ma and diversified in Mexico during the Sangamon Interval, between 0.38-0.16Ma (Fig. 1b, Frey, 1989). The second hypothesis is supported if the niche models depict suitable habitat for dispersal into the United States from the SMOr during the Pleistocene. If the hypothesis of rapid, fairly coincident expansion is to be supported (Fig. 1c), ecological niche models should indicate the presence of suitable habitat over the entire distribution over one time interval. I also examined a hypothesis where colonization among populations occurred coincidently, resulting in paraphyletic relationships among Mexican samples (Fig. 1). I constructed models of suitable habitat over each time period using the maximum entropy algorithm in MAXENT v.3.3.1 (Philips et al., 2006)

The advantages of MAXENT include a robust model output without the need for absence data and flexibility for the coincident analysis of topographical variables, climate and/or categorical data (Phillips and Dudik, 2008; Dudik et al., 2004). I parameterized the model using 4 steps: data proofing, regularization, threshold selection and an assessment of model performance. I collected and vetted a dataset of 1486 locality records from the Global Biodiversity Information Facility (http://data.gbif.org/datasets/resource), MANIS (http://www.manis.org) and the museum server ARCTOS

Figure 1.1. Colonization scenarios proposed for *M. mexicanus.* Abbreviations are United States (USA), Mexico (MX) and the Sierra Madre Oriental (SMOr). A null hypothesis of rapid expansion to all regions is also indicated

Hypothesis	Expected topological relationships	Predictions
Expansion into United States from Mexico after 0.015Ma, and fragmentation of populations in Mexico before 0.015Ma	USA MX	Paraphyletic northern clade Diversification between northern and southern clades date to about 15,000 ybp
Dispersal into United States along Sierra Madre Oriental beginning about 0.50Ma, diversification of populations in Mexico ~0.38-0.16Ma	USA SMOr MX	Divergence of populations in Mexico between 0.38-0.16 Ma Populations sample from Sierra Madre Oriental most closely related to northern populations in the United States
Expansion after Last Glacial Maximum to all areas	USA USA MX MX	Paraphyly throughout phylogenetic topology

(http://arctos.database.museum/). The locality records were proofed for errors or missing latitude and/or longitude data. I eliminated duplicate records and records without detailed locality information. I also excluded data with a coordinate precision below 15km as calculated in Biogeomancer, as the coarse resolution limits their accuracy. After proofing, my final input dataset was comprised of 300 records distributed throughout the extant range (Fig. 2). Finally, I divided the data into a randomly selected training dataset, comprised of 75% of the input data and a testing dataset, comprised of the remaining 25% of the input data.

Because the number of external factors that potentially influence a species distribution can increase rapidly, the algorithms used in most distribution modeling approaches are prone to overfitting. I mitigated for overfitting by setting a regularization parameter. The regularization parameter acts to increase predictive power by weighting the influence of each feature in the model: a more complex model is penalized more compared to a less complex model. In this way, the MAXENT distribution maximizes the likelihood of presence over the study area while minimizing the error associated with more complex models (Phillips and Dudik, 2008).

I evaluated model performance using a threshold-dependent and a threshold-independent approach (Phillips et al., 2006). A threshold is needed to convert a species' probability distribution into presence/absence data in MAXENT (Phillips and Dudik, 2008; Phillips et al., 2006). I applied two thresholds suited for presence-only data (Phillips et al., 2006). I set a minimum training presence threshold, which defines the smallest range of suitable habitat while also including all the training data used in model construction (Phillips et al., 2006). I also selected an equal training threshold that

minimizes the number of presence falsely labeled as absent (false negatives) and the

number of absences falsely predicted as present (false positives) in order to maximize

model sensitivity and specificity, which is recommended for model performance (Liu et

al., 2005). I evaluated the performance of each model by comparing the fraction of the

predicted area with the omission rate for each model.

I also assessed model performance using a threshold-independent approach. To

assess the predictive power of the ENM, I used the area under the ROC curve (AUC)

statistic. The AUC is a robust measure of model accuracy and has been applied

successfully to evaluate several modeling approaches (Elith, 2002). Under the ROC, the

possible outcomes in a presence/absence analysis are: 1, true positives (assigning

presence when actually present), 2, false positives (assigning present when actually

absent), 3, true negatives (assigning absent when actually absent) and 4, false negatives

(assigning absent when actually present). The AUC ranges from 0 to 1 and is a measure

of the probability that a given occurrence is a true positive. An AUC equal to 1 represents

a model with 100% model sensitivity and specificity. With presence-only data, the

maximum AUC is less than 1 (Phillips et al., 2006) and a random prediction has an AUC

= 0.5.

If the Mexican vole colonized the southwestern United States after the Last

Glacial Maximum (circa 0.15Ma) (Davis and Callhan, 1992), I would expect that habitat

suitable for dispersal of the species would be present between the southwest and Mexico

during this time interval. To assess this possibility, I constructed the MODERN ENM by

projecting 300 Mexican vole locality points onto 19 environmental grids in 1km spatial

resolution (Hijimans et al., 2005). I conducted tests of omission by selecting 25% of the

15

occurrence localities as test data, selecting 10,000 randomly chosen points from the

background as random occurrences, and performing 1,000 iterations of the model. I used

the remaining 75% of the locality records as a training data to determine omission errors.

If the Mexican vole colonized the southwestern United States during the

Sangamon Interglacial (c. 0.28-0.16Ma, Frey, 1989), I expect habitat suitable for

dispersal existed between the southwest and Mexico during the time interval. I examined

the scenario by constructing an ENM representative of LGM climatic conditions. I

projected the current niche model onto environmental data calculated for the LGM as

depicted in the Community Climate System Model (CCSM) from the PMIP2 website

(http://www.pmip2.cnrs-gif.fr). An ENM that predicts a past or future climate scenario is

unlikely to have analogs in the current climate (MODERN), and the non-analog

conditions are represented in the model as clamped areas. I used a fade-by-clamping

approach on the LGM model to subtract from the model output any area that includes

clamping. To maximize model performance, I again selected 25% of the data as test data,

and the remaining 75% of the data to train the model. I performed 1,000 iterations of the

model and used the AUC to measure model performance.

Sampling

The Mexican vole is an ideal proxy for an examination of vertebrate colonization

and diversification dynamics in Mexico and the southwestern United States because its

distribution includes pine forest grasslands throughout the region. To conduct

phylogeographic analyses, I sequenced the cytochrome *b* gene of the mitochondrial DNA

(cyt*b*) and the nuclear intron, the acid phosphatase V (AP5) gene. I examined

mitochondrial and nuclear genes because of the difference in the putative mutation rate

(μ) and differences in effective population size (N_e) for each gene (Avise, 2000). The cyt*b* gene is recognized as a data rich site, and mutates at a rate suitable to resolve nodes younger than 2 Ma (Galewski et al., 2006). Nuclear exons and introns are thought to mutate at a slower rate (Friesen, 2000). Nuclear introns are nonfunctional and as such each base position can potentially provide phylogenetic signal (compared to the 1_{st} and 2_{nd} positions in mtDNA) and accumulate mutations at a faster rate compared to coding regions (Friesen, 2000). Most phylogenetic studies of the variation within nuclear introns have primarily been at the generic level (e.g., Matthee et al., 2006; Creer et al., 2003; Pacheco et al., 2002; Debry and Sheshadri, 2001), although variation within the nuclear intron AP5 was recently examined in a species–level study of *Microtus californicus* (Conroy and Neuwald, 2008).

Prior to genetic analyses, I obtained *M. mexicanus* tissue samples from the University of New Mexico's Museum of Southwestern Biology (MSB) and skin samples from Michigan State University Museum (MSUM). I selected samples to provide a representation of the geographic extent of the species (Fig. 1.2). I included tissue from 8 of the 12 subspecies of *Microtus mexicanus* recognized by Hall (1981), including *M. m. navaho* (n = 1), *M. m. hualpaiensis* (n = 6), *M. m. mogollonensis* (n = 13), *M. m. guadalupensis* (n = 8), *M. m. subsimus* (n = 4), *M. m. mexicanus* (n = 3) *M. m. fulviventer* (n = 3). Museum skin samples were obtained from an additional population of *M. m. madrensis* at MSUM (n = 7) (Fig. 1.2, Appendix A). Material for the remaining 4 subspecies was not available. For convenience in subsequent analyses, I grouped the samples according to the respective mountain regions from which they were collected. The resulting populations and their respective abbreviations are the mountains of the

Figure 1.2 The geographical distribution of *M*. *mexicanus* samples analyzed in this study. The distribution of the 8 subspecies examined in this study, its location on the map and its abbreviation include 1) *M. m. Navaho* (Utah, (UT)), 2) *M. m. hualpaiensis* (northwestern Arizona, COP (AZ)), 3) *M. m. mogollonensis* (eastern New Mexico, SKY (NM)), 4) *M. m. guadalupensis* (western New Mexico, SKY (NM)), 5) *M. m. subsimus*, 6) *M. m. fulviventer* (Oaxaca, Mexico, OAX)*, 7) M. m. mexicanus* (Veracruz, Mexico, VZ and MX), and 8) *M. m. madrensis* (Durango, Mexico, DUR)). Mexican vole localities used in the ecological niche modeling are shown in blue; molecular samples are shown in red. Map projection is geographic World Geodetic System 1984
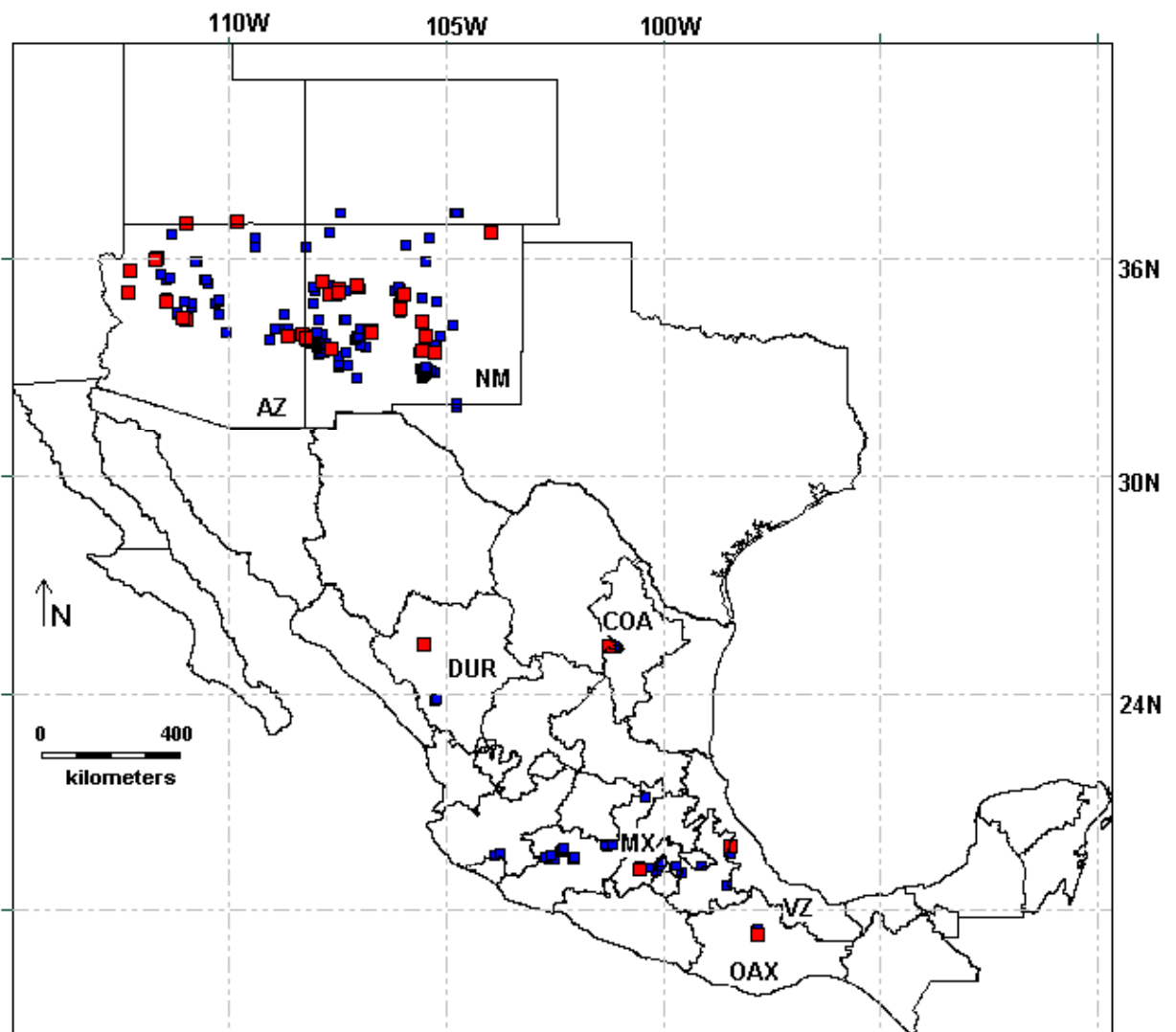
Colorado Plateau (COP (AZ)), the mountains of New Mexico (SKY (NM)), the Sierra

Madre Occidental (SMOcc (DUR)), and the Sierra Madre Oriental, which included

samples from the states of Coahuila, Veracruz, Mexico and Oaxaca in Mexico (Fig. 1.2).

I placed the observed divergence within *M. mexicanus* in a broader context by including

2 additional sequences from GENBANK as outgroup taxa (Appendix A). These included

sequence data for *Myodes gapperi*, a recognized outgroup for *Microtus* (Jaarola et al.,

2004; Conroy and Cook, 1999), and *Microtus californicus*, the hypothesized sister taxon

to *M. mexicanus* (Conroy and Cook, 1999).

Sample preparation and sequencing

   I extracted nucleic acids from 46 *M. mexicanus* samples using a modified salt

extraction procedure (Medrano et al., 1990). I amplified the cyt*b* gene of the

mitochondrial DNA (cyt*b*) via the polymerase chain reaction (PCR) and primer pairs

L14724 and H15915 (Irwin et al. 1991), and internal primer pairs MVZ23-MVZ14,

MVZ05-MVZ04, and MVZ11-MVZ26 (Smith and Patton, 1993). Museum skin samples

were soaked twice in 10% SDS solution at 42° C for 1 hour to remove potential

contaminants prior to extraction. Reaction conditions for PCR were 94°C (30 s),

denaturing, 50°C (25 s) annealing, and 72°C (1 min) extension. PCR products were

purified using the QIaquick PCR Purification Kit (Qiagen, Valencia, California). I

prepared PCR templates for automated sequencing using dye-labeled terminators and

cycle sequencing conditions of 96°C (10 s) denaturing, 50°C (5 s) annealing, and 60°C (4

min) extension. PCR products were sequenced using the same primer sets, and purified

using a Qiagen DyeEx Spin kit. Sequence samples were read using an ABI 3100

Automated Sequence Analyzer and manually aligned using SEQUENCER v.4.7

(Bromberg et al., 1995).

I examined the temporal nature of variation within Mexican vole populations by

also sequencing the AP5 gene from 22 Mexican vole samples. I included tissue from 12

vole samples from within the United States and 10 samples from Mexico. These included

M. m. hualpaiensis (n=5), *M. m. mogollonensis* (n=5), *M. m. guadalupensis* (n=2), *M. m.*

subsimus (n=6), *M. m. mexicanus* (n=2), and *M. m. fulviventer* (n=2) (Appendix B). The

AP5 gene was amplified and sequenced using gene specific primers (Debry and Seshadri,

2001) and reaction conditions similar to that for the cyt*b* gene.

Analysis of molecular variation

If the extant distribution of the Mexican vole is the product of recent range

expansion into the southwestern United States from Mexico, I would expect that the

genetic signature of the southwestern population to include: a) reduced genetic diversity,

b) an unresolved tree topology, c) divergence estimates following the LGM, d) a

unimodal pairwise mismatch distribution, and e) reject neutrality. I analyzed the

molecular variation from the cyt*b* and AP5 to assess the degree of congruence between

the observed and expected levels of molecular diversity.

I examined molecular variation within and between populations of the Mexican

vole using both mitochondrial (cyt*b*) and nuclear (nDNA) DNA. The differences in the

effective population size and coalescent time expected between cyt*b* and nDNA can

potentially provide independent assessments of population history (Sunnucks, 2000). I

examined the level of diversity at the cyt*b* through genetic parameters theta pi (π), theta

(Θ), effective population size (Ne) and the number of segregating sites (*S*). I calculated π,

Θ and *S* in DnaSP v.4.90 (Rozas et al., 2003). Sites that are missing characters or gaps are ignored from phylogenetic analyses in DnaSP, ARLEQUIN, MIGRATE and MRBAYES. However, I excluded from all analyses 5 samples that were missing large portions of data. I calculated effective population size (N_e) using MIGRATE v.3.0 (Beerli and Felsenstein, 2001), where I set θ for haploid data equal to $\theta = 2 N_e \mu$, and N_e is the effective population size and a mutation rate (μ) of 2.0%/lineage/million years (Brown et al., 1979).

An analysis of molecular variance (AMOVA) was conducted on the cyt*b* data in ARLEQUIN v.3.1 (Harpending, 1994; Excoffier et al., 2005) to assess partitioning of molecular variation. To mitigate for samples with missing data, I set the search criteria to include only those sites that were sequenced in all individuals. Molecular variation within and between populations was estimated with the Kimura 2-parameter (D) distance, which estimates the pairwise distances while correcting for the different rate between transitional and transversional substitutions. I divided the data into eastern and western clades, as identified from phylogeographic analyses (below). I estimated the hierarchical F-statistics in ARLEQUIN, which estimate the degree of subdivision while correcting for the influence of unequal sampling (Weir and Cockerham, 1984)

I also summarized molecular diversity within and between vole populations from the nuclear intron AP5. I calculated the summary statistics of the number of segregating sites (S), nucleotide diversity (π), and theta (Θ) in DnaSP v.4.90 (Rozas et al., 2003). Nuclear genes are expected to have a higher N_e compared to mitochondrial genes due to longer coalescent times (Avise, 2000). I calculated N_e for the nuclear intron, AP5 using the relationship, $\theta = 4 N_e \mu$ (i.e. $Ne = \theta/4\mu$), where the expected μ for the AP5 gene is 0.072

(Weinreich, 2001; Yang and Nielsen, 1998) and the rate for nuclear introns $\mu = 0.79$ x 10^{-9} (Yu et al., 2001).

Phylogenetic analyses

I conducted several phylogenetic analyses to examine the partitioning of genetic variation within and between clades in *M*. *mexicanus*. I examined the likelihood that *M*. *mexicanus* is monophyletic using a maximum likelihood approach in MRBAYES v.3.1.2 (Huelsenbeck and Ronquist, 2001). Missing characters and gaps are treated as missing data and do not contribute to phylogenetic analyses in MRBAYES v.3.1.2. Therefore, I included the sequences from all individuals sampled. Because of unequal sampling, phylogenetic analyses were conducted on the cyt*b* and AP5 gene separately.

For the cyt*b* data, I selected the model of evolution that best fit the data using MODELTEST (Posada and Crandall, 1998). The model of evolution selected by the Akaike Information Criterion (AIC) was HKY +I +G with 6 free parameters. The Bayesian analysis was conducted using the parameters of two chains run simultaneously; each chain ran for 2,000,000 generations and every 1000th tree was sampled. The runs converged onto the stationary distribution of likelihood scores at 21,600 generations. The first 540 trees were discarded so that only parameters estimated after stationarity were used. Because of differences in sample size between the cytb and AP5 data, I conducted a second Bayesian analysis for a reduced cyt*b* dataset that I trimmed to match the samples represented in the AP5 dataset.

I conducted analyses of the AP5 gene to estimate the temporal nature of variation within the Mexican vole and to examine congruence between the cyt*b* and nuclear gene topologies. The model of evolution that best fit theAP5 data selected by the Akaike

Information Criterion (AIC) in MODELTEST (Posada and Crandall, 1998) was HKY + I + G with 6 free parameters. I analyzed the distribution of variation within the AP5 gene in MRBAYES v.3.1.2 (Huelsenbeck and Ronquist, 2001).

Coalescent simulations

I conducted coalescent simulations of constrained topologies on the cyt*b* data to examine the timing of diversification and expansion events. I performed coalescent simulations to estimate the likelihood of each topology suggested from a colonization scenario using MESQUITE v.2.6 (Maddison and Maddison, 2009). The input parameters were set to the estimated N_e for mtDNA and the number of iterations to 1000 for each simulated genealogy. I manipulated the tree topology to match the expectations of each colonization scenario (Fig. 1). My simulations included manipulating the location of clades to reflect a paraphyletic topology, as well as a topology with clades in the United States and Mexico. I also modified the length of tree branches to reflect pre-Pleistocene (branches ≥ 2 Ma) and post-Pleistocene (branches ≤ 0.01 Ma) divergence. I estimated the likelihood of each topology using the distribution of the *S*-statistic (Slatkin and Maddison, 1989). The smaller the value of *S*, the better the fit of the model given the data.

Divergence estimation

I calculated divergence estimates for each clade to evaluate scenarios regarding the timing of diversification in the Mexican vole. A test for rate homogeneity was conducted using the Likelihood ratio test with *s-2* degrees of freedom (Felsenstein, 1981). I estimated the divergence times for *M. mexicanus* clades using BEAST v.1.4.6

(Drummond et al., 2006). The program uses a Bayesian Markov chain Monte Carlo algorithm to estimate divergence times between clades under a relaxed molecular clock assumption.

The purpose of relaxed molecular clocks is to obtain divergence estimates while allowing differential evolutionary rates within lineages. When fossils are available to serve as external calibration points, model priors are constrained to be lognormal. In this way, rates of evolution are allowed to vary along the entire branch, rather than concentrating the rate on the node, which yields more robust estimates (Ho, 2007). Divergence times between Mexican vole clades were estimated in two ways and the results compared. First, I set the divergence between *Microtus* and *Myodes* using a lognormally distributed prior of 1.3 ± 0.2 Ma based on fossil evidence (Chaline and Graf, 1988; Catzeflis et al., 1987), and an internal lognormal prior of 0.3 ± 0.1 Ma for the divergence of *Microtus* in Mexico (Conroy et al., 2001). In an effort to maximize the effective sample size, and minimize the error of a given divergence estimate, I set the Bayesian search to 10,000,000 generations (Drummond et al., 2006). I performed a second set of divergence estimates by setting the divergence between *Microtus* and *Myodes* using the same lognormally distributed prior of 1.3 ± 0.2 Ma, with an internal lognormal prior representative of divergence within *Microtus* at 0.5 ± 0.2 Ma (Steppan et al., 2004; Conroy and Cook, 1999; Repenning, 1990), and a lognormal prior of 0.3 ± 0.1 Ma for the divergence of *Microtus* in Mexico (Conroy et al., 2001). The search was again conducted using 10,000,000 generations.

Tests of neutrality

If the extant distribution of the Mexican vole is the product of recent range expansion into the southwestern United States from Mexico, I would expect that the southwestern population to violate neutrality. I assessed population expansion through tests of neutrality on the mtDNA. A Fu's *Fs* (Fu, 1997) test of the mismatch distribution is expected to yield significantly negative values in expanding populations because the number of segregating sites (S) is expected to change more rapidly given changes in population size compared to the pairwise % nucleotide diversity (π). I assessed the indices of molecular diversity such as theta pi (π) and theta S (*S*) in ARLEQUIN v.3.1 (Harpending, 1994; Excoffier et al., 2005) to test for expansion in *M. mexicanus*. I conducted a mismatch analysis of northern and southern populations using 10,000 permutations in ARLEQUIN v.3.1 (Harpending, 1994; Excoffier et al., 2005). I tested neutrality by conducting a Fu's *Fs* test (Fu, 1997) in ARLEQUIN v.3.1 using the number of rare alleles (Harpending, 1994; Excoffier et al., 2005).

Although nuclear introns are thought to be non-coding, recent studies suggest the action of selection on some introns (Matthee et al., 2006). I tested the neutrality of the AP5 gene using coalescent simulations in DnaSP (Rozas et al., 2003). I evaluated the influence of selection on the AP5 genes in populations of the Mexican vole through tests of neutrality Fu's *Fs*, Tajima's *D* (Tajima, 1989) and Fu and Li's *D** (Fu and Li, 1993) test statistics. The action of selection is suggested when *Fs* and *D* are non-significant but *D** is significant. Neutrality test settings included the default of free recombination, and 5000 replications.

RESULTS

Ecological modeling

Ecological niche models (ENM) depict the probability distribution that represents the likelihood that a species will be found at a particular location, given its climatic preferences and geographic distribution. My examination of the LGM and MODERN models suggested two main patterns in the distribution of the Mexican vole over time. First, the MODERN model predicted more area suitable for range expansion throughout the southwestern United States by the Mexican vole compared to the LGM model (Fig. 3). While the MODERN model depicted suitable habitat in central/eastern Arizona and New Mexico, the LGM model did not. The LGM model predicted relatively greater distribution of suitable habitat compared to the MODERN. The relatively widespread distribution of suitable habitat predicted in the LGM model (Fig. 3a) appeared fragmented and discontinuous in the MODERN model (Fig. 3b). The MODERN model predicts two disjunct regions of suitable habitat for the Mexican vole found north and south of approximately 32° latitude.

I report greater support for the minimum training threshold because it predicted a greater fractional area (fa) and lower omission rate (or) (Appendix E, fa = 0.488 and or = 0.000) compared to the predicted fractional area and omission rate for the equal training threshold (Appendix E, fa = 0.057, or = 0.059). The lower omission rate for the minimum training presence threshold suggests that no test locations fell into pixels predicted as unsuitable for the species. Both threshold approaches preserved a region of suitable habitat along the north-south axes of the Sierra Madre, which may be a historical connection between populations in the United States and Mexico (subsequent paragraph).

Both the LGM model (AUC_{train}= 0.986, AUC_{test}= 0.969) (Fig. 1.3a) and the modern

niche model (AUC_{train} =0.980, AUC_{test} = 0.976, Fig. 1.3b) were well supported by the

AUC criterion. For both models, the small difference between AUC_{test} and AUC_{train}

suggests the MAXENT algorithm is able to detect small differences in the environmental

conditions for test and training data. The overall model performance from both threshold-

dependent and threshold-independent tests suggest the MAXENT distribution is a close

approximation of the probability distribution of the species actual distribution.

Analysis of molecular variation

Results from the AMOVA indicated relatively high molecular diversity (π) for

cyt*b* data in Veracruz (0.034 ± 0.001), Oaxaca (0.031± 0.023) and Durango (0.049 ±

0.029), Mexico (Table 1.1, Appendix C). I partitioned vole populations into the western

and eastern clades identified from phylogeographic analyses (below) and estimated the

parameters of nucleotide diversity (π) and theta (Θ). The eastern clade demonstrated

greater genetic diversity across most parameters (π = 0.191 ± 0.044) compared to the

western clade (Table 1.1, π = 0.032 ± 0.001). I estimated Ne for the cyt*b* data using the

relationship $N_e = \Theta/\mu$, where μ = 0.02 substitutions per site per million years and a range

of 0.021-0.029 (Brown et al., 1979; Smith and Patton, 1993) corrected by the generation

time of 7 per year (Keller, 1985). I evaluated estimates of Θ from MIGRATE v.3.0

(Beerli and Felsenstein, 2001). My estimates of N_e resulted in smaller N_e for population

Figure1.3a. Ecological niche models for the Mexican vole. Ecological niche models representative of the Last Glacial Maximum during the late Pleistocene (a) and modern climate conditions (b) to examine the potential timeframe of colonization proposed by each hypothesis. The LGM model predicted a large region of contiguous habitat throughout the Sierra Madre into western Arizona and unsuitable habitat in central/eastern Arizona and Mexico. The MODERN model depicted significant fragmentation of LGM habitat, with the exception of northwestern Arizona and predicted newly available habitat in central/eastern Arizona and New Mexico. Map projection is geographic World Geodetic System 1984
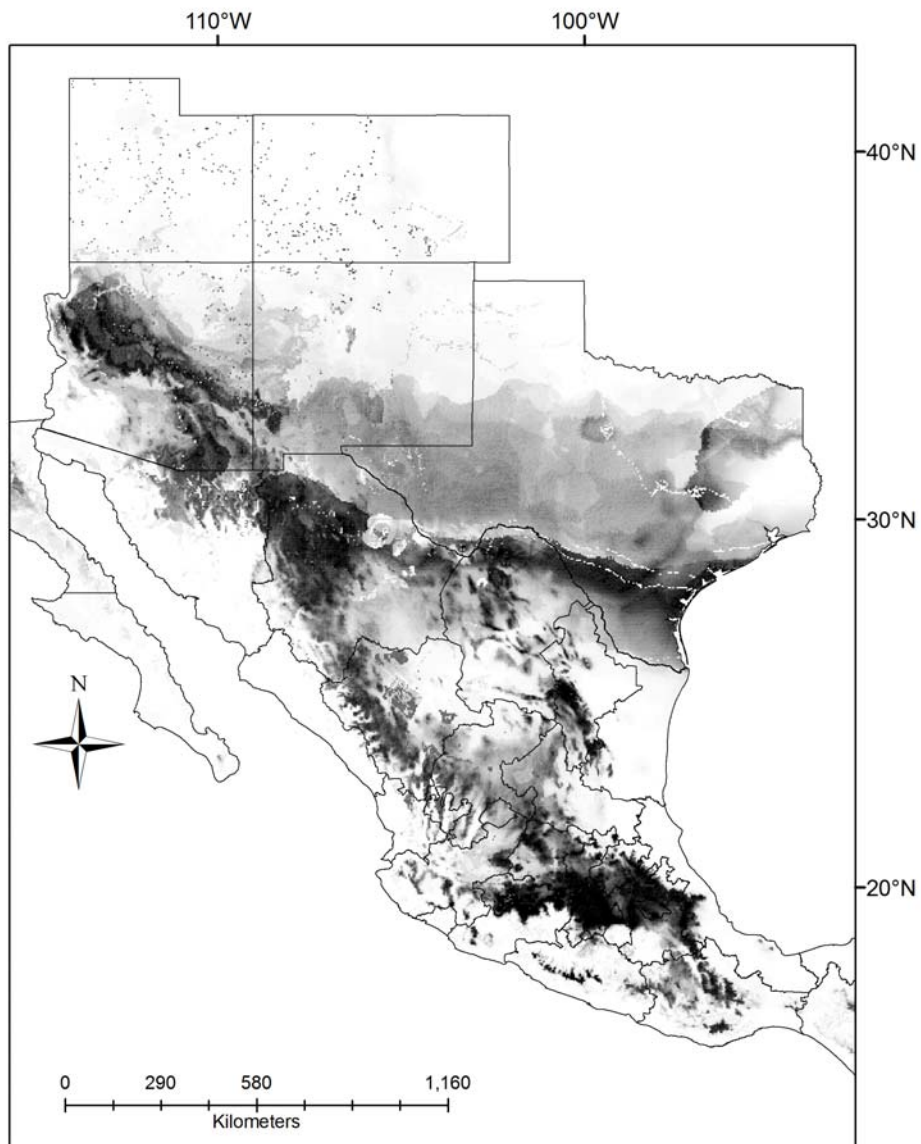
Figure1.3b
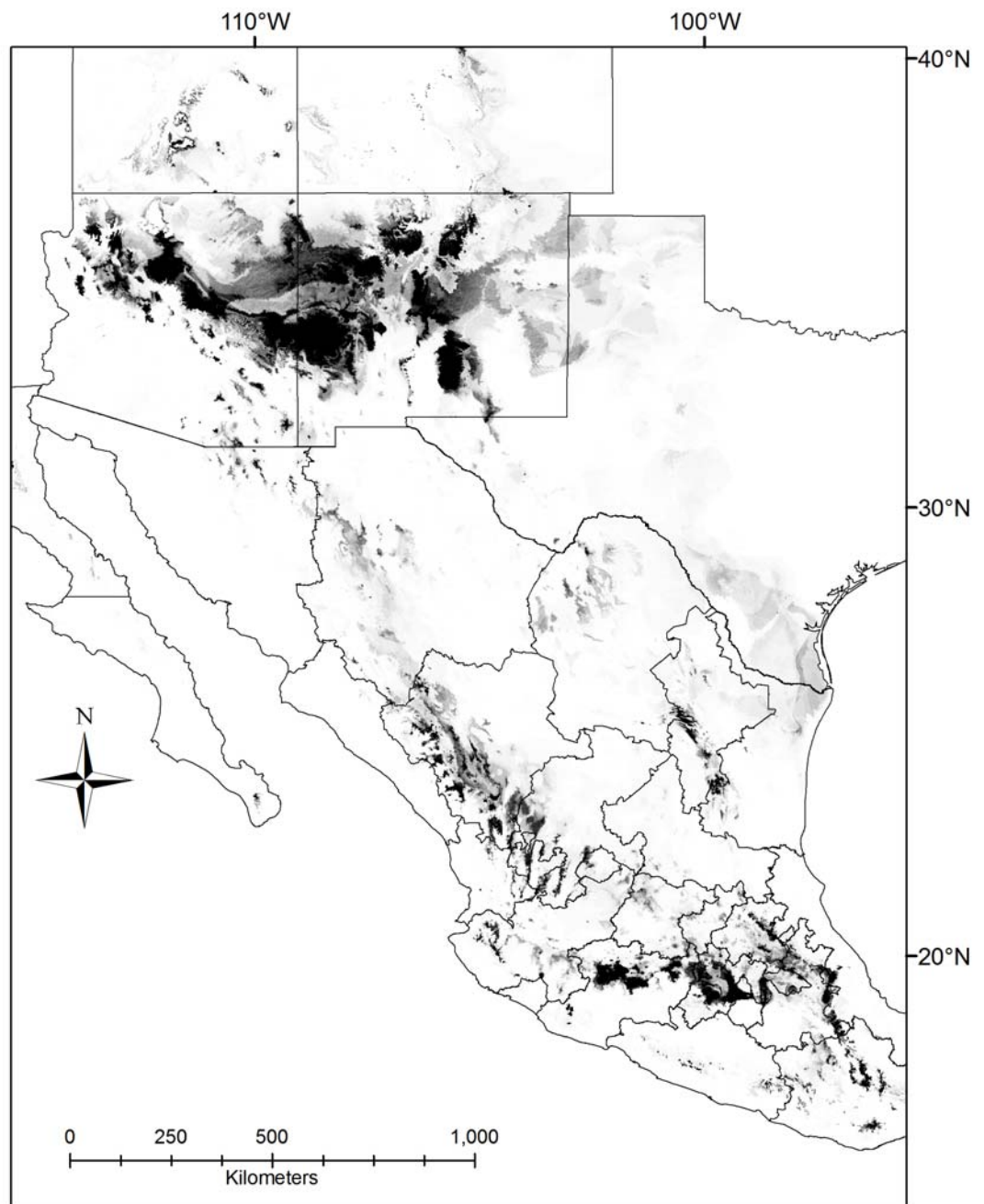
Table 1.1 Summary statistics for the Mexican vole. Parameters include the number of polymorphic sites (*S*), nucleotide diversity (π), and the number of unique sequences sampled per population (US). Data are summarized from the cyt*b* gene of the mitochondrial DNA from western (COP + SKY + SMOcc) and eastern (SMOr + SmdelSur) clades. Population abbreviations as in Figure 2. Sample sizes (N) are also included.

Clade	Population	N	*S*	π	Θ	US
West	COP (AZ)	6	15	0.005 ± 0.003	0.002	3
	SKY (NM)	22	11	0.036 ± 0019	0.026	5
	SMOcc (DUR)	7	30	0.049 ± 0.029	0.051	5
Overall		35	35	0.032 ± 0.001	0.033	3
East	SMOcc (COA)	4	3	0.002 ± 0.004	0.003	1
	SMOcc (VZ)	3	2	0.034 ± 0.001	0.034	3
	SmdelSur (OAX)	3	33	0.031 ± 0.023	0.025	3
Overall		10	3	0.191 ± 0.044	0.175	5

in the United States ($\Theta = 0.032$, $N_e = 202,041$) compared to those in Mexico ($\Theta = 0.191$, $N_e = 1,356,463$). From these data, I estimated the average theta (Θ) for the total population ($\Theta = 0.112$, $N_e = 779,252$).

Despite the relatively paraphyletic relationships observed in some interspecific comparisons using nuclear genes, sequence data obtained from nuclear genes can provide potentially important insights into species dynamics. I assessed inter-clade diversity through the statistics π, Θ, S, N_e and tests of neutrality. My analyses of the AP5 data demonstrated significantly greater nucleotide diversity in the western clade ($\pi = 0.034$) compared to the eastern clade ($\pi = 0.004$) despite similar sample sizes (Table 1.2, Appendix D). AP5 sequence was unavailable for the population at Durango, Mexico.

I estimated N_e for the eastern clade, western clade and the overall N_e for the populations using the expression, $\Theta = 4N_e\mu$ as the nuclear gene was expected to be 4 times that for the cyt*b* gene. The expected mutation rate for the non-synonymous to synonymous substitutions (d_N/d_S) for the AP5 genes in rodents is estimated $d_N = 0.049$ and $d_S = 0.680$ (Weinrich, 2001; Yang and Nielsen, 1998). Assuming neutrality, the substitution rate is equal to the mutation rate (Yang and Nielsen, 1998; Kimura, 1983). Although the assumptions of neutrality may not be valid across Mexican vole lineages, the value provides a rough estimate of the expected mutation rate for the entire AP5 gene. The d_N/d_S ratio was estimated for rodents $\omega = 0.072$ (Weinrich, 2001), or estimated as a mutation rate of $0.74 \times 10\text{-}9$ substitutions per site per million years (Yu et al., 2001). Using an average generation time of 7 per year, the values of Θ I calculated from

Table 1.2 Summary statistics for clades of the Mexican vole. Parameters include the number of polymorphic sites (S), nucleotide diversity (π), and the number of unique sequences sampled per population (US). Data are summarized from theAP5 nuclear intron from western (COP + SKY + SMOcc) and eastern (SMOr + SmdelSur) clades. Population abbreviations as in Figure 2. Sample sizes (N) are also included.

Clade	Population	N	S	π	Θ	US
West	COP (AZ)	5	15	0.029 ± 0.008	0.025	3
	SKY (NM)	7	26	0.035 ± 0.017	0.044	2
	SMOcc (DUR)	na	--	--	--	--
Overall		12	22	0.032 ± 0.010	0.034	5
East	SMOcc (COA)	6	1	0.002 ± 0.000	0.002	2
	SMOcc (VZ)	2	1	0.001 ± 0.001	0.001	1
	SmdelSur (OAX)	2	3	0.008 ± 0.003	0.008	2
Overall		10	3	0.005 ± 0.002	0.004	4

MIGRATE v.3.0 (Beerli and Felsenstein, 2001) produced a larger N_e for the United States population ($\Theta = 0.032$, $N_e = 6,274,864$) compared to the populations in Mexico ($\Theta = 0.005$, $N_e = 705,244$) and a total $N_e = 3,490,054$.

Phylogenetic analyses

Phylogeographic analyses of the cyt*b* gene strongly support the monophyly of the Mexican vole. Base frequencies observed in this study for cyt*b* are similar to those reported in other studies of *Microtus* (Conroy and Cook, 1999), and no significant deviations were found ($\chi^2 = 68.820$, df =180, P =1.00). There is also additional topological support for *Microtus californicus* as sister to *M. mexicanus* (Conroy and Cook, 2000); thus, *M. californicus* was used as an outgroup for *M. mexicanus* in subsequent analyses.

My analysis of the cyt*b* gene data in MODELTEST v.3.7 (Posada and Crandall, 1998) demonstrated a level of complexity best modeled by the HKY+I+G model of evolution with 6 free parameters; base frequencies were estimated as: A, 35.31%, C, 33.36%, G, 8.06%, and T, 23.27%. The transition/transversion ratio was estimated at 5.54, and values for the proportion of invariable sites (I), and the gamma shape parameter (G) were estimated as 0.5207, and 0.9725, respectively. I conducted a parsimony analysis on 210 unordered, equally weighted, parsimony informative characters. The parsimony analysis generated 660 trees of 638 length, with a consistency index of 0.677, and a retention index of 0.780. A majority consensus tree depicted two clades within *M. mexicanus*; a western clade composed of samples from the southwestern United States and Durango, Mexico and an eastern clade composed of samples from northeastern and southeastern Mexico.

I examined additional support for the eastern and western clades depicted in the parsimony analysis through a Bayesian analysis of the cyt*b* data in MRBAYES v.3.1.2 (Huelsenbeck and Ronquist, 2001). The Bayesian analysis demonstrated a 90% posterior probability for an east/west clade division. There was moderate support for a subclade comprised of samples from northwestern Arizona. In all topologies, the eastern clade was basal to the western clade. A phylogram constructed using a Bayesian algorithm in MRBAYES v. 3.12 (Huelsenbeck and Ronquist, 2001) depicted a similar topology, with longer branch lengths indicative of greater genetic divergence for samples from Mexico, including samples from within the eastern and western clades (Fig. 1.4). A similar topology was recovered from a Bayesian analysis of a reduced cyt*b* dataset, which I trimmed to match the samples from the AP5 dataset.

The reduced mutation rate and larger effective population size expected for nuclear genes suggest that a phylogenetic topology will reflect the sorting of ancestral polymorphisms from an earlier coalescent time compared to cyt*b* gene. I examined this expectation from 453bp of the AP5 nuclear intron for 26 Mexican vole samples. The model selected by the AUC criterion in MODELTEST v.3.7 (Posada and Crandall, 1998) was the HKY+I+G model, with 6 free parameters. The parameters estimated for the data included a transition/transversion ratio of 5.53, the proportion of invariable sites, $I = 0.5207$ and the gamma shape parameter of 0.9725.

Figure 1.4. A Bayesian phylogram derived from 43 cytb sa,mples. The western clade, comprised of samples from the Colorado Plateau (COP (AZ)), the mountains of New Mexico (SKY (NM)) and samples from the Sierra Madre Occidental in Durango, Mexico (SMOcc (DUR)) is depicted as diagonal (COP AZ)) and diagonal striped bars. The eastern clade, made up of populations in the Mexico states of Coahuila and Veracruz along the Sierra Madre Oriental (SMOr (COA) and SMOr (VZ), respectively) and in the Sierra Madre del Sur in Oaxaca (SMdelSur (OAX)), is depicted as horizontal bar. Support for major nodes is indicated as a Bayesian posterior probability to the left of each bracket, and bootstrap values to the right

COP (nw AZ)
COP (nw AZ)
0.92/100 COP (nw AZ)
COP (nw AZ)
COP (nw AZ)
COP (AZ)
SKY (NM)
SKY (NM)
SKY (NM)
SKY (NM)
SKY (NM)
SKY (NM)
SKY (NM)
SKY (NM)
0.96/90 SKY (NM)
SKY (NM)
SKY (NM)
SKY (NM)
SKY (NM)
SKY (NM)
0.90/80 SKY (NM)
SKY (NM)
COP (UT)
SKY (NM)
0.59/60 SKY (NM)
SKY (NM)
SKY (NM)
COP (AZ)
1.00/70 SMOcc (Durango)
SMOcc (Durango)
0.77/70 SMOcc (Durango)
0.61/70 SMOcc (Durango)
SMOcc (Durango)
SMOcc (Durango)
0.90/70 SMOcc (Durango)
SMOr (Coahuila)
0.99/100 SMOr (Coahuila)
SMOr (Coahuila)
SMOr (Coahuila)
0.95/90 SMOr (Veracruz)
1.00/100 SMOr (Veracruz)
SMOr (Veracruz)
1.00/100 SMdelSur (Oaxaca)
SMdelSur (Oaxaca)
SMdelSur (Oaxaca)
0.92/100 *M californicus*

0.1 Sequence Difference

The likelihood topology demonstrated 341 informative characters and a tree length = 822, CI=0.9112, RI=0.9280. My analysis of the AP5 intron within a Bayesian framework in MRBAYES v. 3.12 (Huelsenbeck and Ronquist, 2001) yielded a similar topology: a polytomy except for genetic structure for samples from northwestern Arizona (Fig. 1.5).

Coalescent simulations

Results of simulations suggested a mid-Pleistocene divergence and expansion from a single refugium was not supported (P>0.99), or the hypothesis of recent expansion from a single refugium (P>0.99). Instead, results suggested mid-Pleistocene divergence of eastern populations and expansion from 2 refugium in the late Pleistocene (P<0.0001). The S-statistic was calculated as S=8 for the empirical data.

Clade divergence estimation

A test for rate homogeneity was conducted using the Likelihood ratio test with *s-2* degrees of freedom (Felsenstein, 1981); the assumption of rate homogeneity among all branches was rejected. Results of both analyses were similar, with an estimated mid-Pleistocene divergence for the eastern clade and a late-Pleistocene divergence for the western clade (Table 1.3).

Tests of neutrality

I also examined the expectations for neutrality for the parameters of mutation rate (μ) and the effective population size (N_e) in the cyt*b* and AP5 genes. My analyses of data from mitochondrial and nuclear data demonstrate two results. Demographic expansion

Figure 1.5 A Bayesian phylogram derived from the nuclear intron, acid phosphatase V (AP5). The 24 samples analyzed are abbreviated as in Figure 1 and a sample from the southern Rocky Mountains (S Rcky Mts). A genetically distinct clade from the Hualpai Mountains of northwestern Arizona is shown (COP, nw AZ).
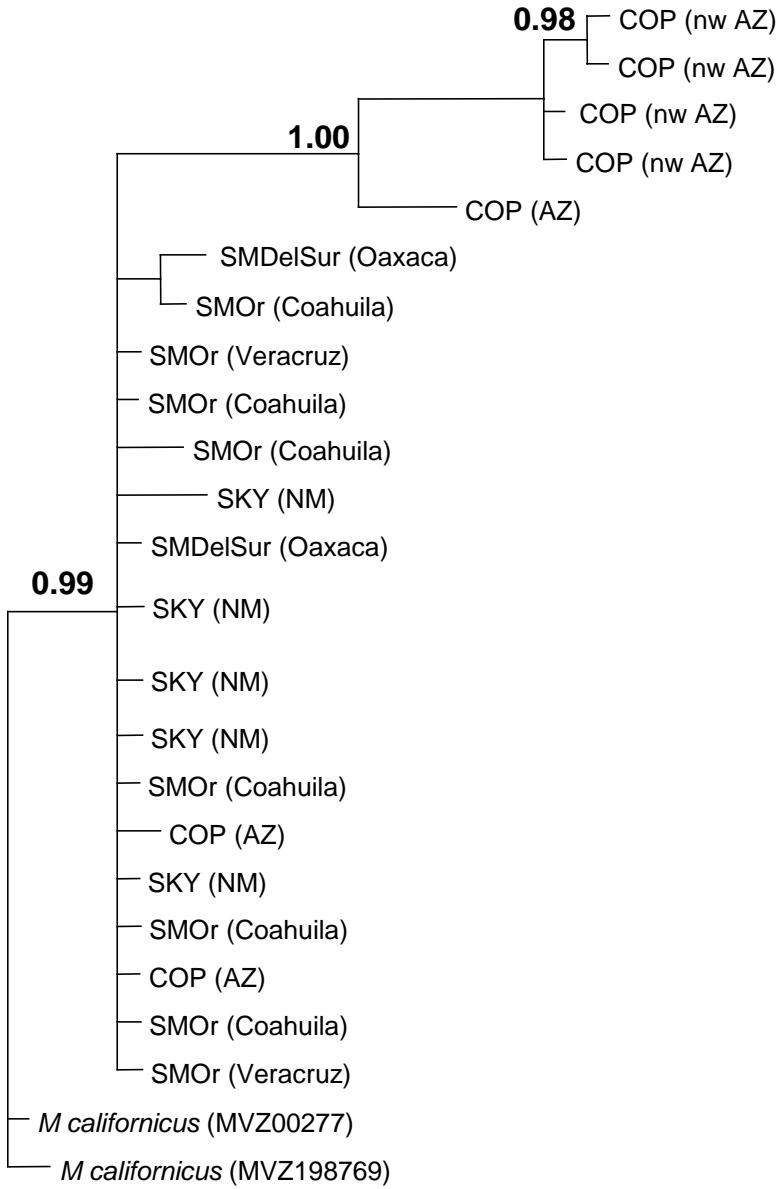
Table 1.3. Estimates of divergence for 6 clades of the Mexican vole. Divergence was estimated using a lognormal uncorrelated molecular clock and 10,000,000 generations in BEAST v.1.4.6 (Drummond et al., 2006). The first test (A) included the lognormal prior of 1.3 ± 0.2 divergence as an external calibration point for the split between *Myodes* and *Microtus* (Conroy and Cook, 2000), and a lognormal prior of 0.75 ± 0.1 Ma as an internal calibration point for divergence of *Microtus* in Mexico (Conroy et al., 2001). The second test (B) incorporated a prior of 2.4 ± 0.5 Ma for diversification within *Microtus* (Conroy and Cook, 1999). The mean divergence (T_{MRCA}), effective sample size (ESS) and range of divergence estimate (highest posterior density) are reported. An ESS value of ≥ 100 is considered appropriate for robust parameter estimation (Drummond et al., 2006).

Clade	Method	Mean Divergence (Ma)	ESS	Range (Ma)
Northwestern Arizona	A	0.08	298.58	0.02-0.18
	B	0.07	678.41	0.14-0.15
New Mexico	A	0.30	61.14	0.08-0.55
	B	0.31	415.65	0.12-0.28
Durango	A	0.42	651.67	0.04-0.91
	B	0.48	544.03	0.11-0.90
Coahuila	A	0.20	1180.21	0.03-0.44
	B	0.16	1093.87	0.03-0.35
Veracruz	A	0.36	384.41	0.09-0.71
	B	0.28	673.69	0.07-0.55
Oaxaca	A	0.61	298.13	0.13-1.2
	B	0.48	195.99	0.12-0.92

Table 1.4. Mismatch analyses and tests of neutrality for cyt*b* data by clade. Parameters calculated include the *Fu's Fs* test of neutrality (Fu, 1997) and associated *P*-value, a measure of the fit of the data to a model of population expansion (raggedness index, *r*), and the associated *P*-value. Under a model of population expansion, *P*-values for the raggedness index, *r* are expected to be non-significant (*P*> 0.05), and *Fu's Fs* test are expected to be significantly negative. Clade and population abbreviations are the same as those used in Table 1. Sample sizes (N) are also provided.

Gene	Clade	Population	N	*Fs* (p-value)	*r* (p-value)
cyt*b*	West	COP (AZ)	6	-1.83 (0.07)	0.16 (0.43)
		SKY (NM)	17	**-16.93 (0.00)**	**0.1 (0.26)**
		SMOcc (DUR)	7	-1.07 (0.15)	0.13 (0.27)
	Overall		30	0.44 (0.56)	0.74 (0.43)
	East	SMOr (COA)	4	**-2.18 (0.02)**	**0.22 (0.98)**
		SMOr (VZ)	3	-1.22 (0.07)	1.00 (0.49)
		SMdelSur (OAX)	3	2.44 (0.53)	0.66 (0.77)
	Overall		10	0.26 (.046)	0.11 (0.00)

was supported for the cyt*b* for the samples from New Mexico (Table 1.4) by a

nonsignificant raggedness index (*r*), and a significant negative value from a Fu's *Fs* test

in ARLEQUIN (*r* = 0.04, *Fs* = -2.54). Tests of neutrality on the AP5 data indicated a

significant *D** test, but a non-significant *Fs* and *D* test, suggestive of positive selection on

samples from the western clade (Table 1.5). The D* test was not significant when I

included the samples from northwestern Arizona (*D** = -0.003, P = 0.87).

Table 1.5. Results from neutrality tests and summary statistics for clades of the Mexican vole. Data is summarized from the AP5 gene from eastern (SMOr + SmdelSur) and western (SMOcc + SWMtns) clades of the Mexican vole. Parameters calculated include the Tajima's D test (*D*), Fu's *Fs* test of neutrality (*Fs*) and the Fu and Li's *D* test for selection (*D**). Significant tests are indicated (*). Under a model of population expansion, *Fu's Fs* test are expected to be significantly negative. Positive selection is indicated as a significant *D** test. Clade and population abbreviations are the same as those used in Table 1. Sample sizes (N) are also provided.

Gene	Clade	N	*D*	*Fs* (p-value)	*D** (p-value)
AP5	East	13	-0.01	0.19 (0.98)	-0.04 (0.65)
	West	11	-0.04	0.26 (0.98)	-0.09 (0.02)*

DISCUSSION

My analyses of ecological and genetic parameters suggested the action of both vicariance and colonization in the extant distribution of the Mexican vole. Additionally, results indicated that the influence of the two processes varied spatially and temporally. The original colonization of Mexico notwithstanding, the distributions of genetic lineages in the Sierra Madre Mountains of Mexico appeared to be the result of late Pleistocene vicariance. Conversely, the colonization of the southwestern United States occurred relatively recently, following the Last Glacial Maximum.

The greater relative amount of suitable habitat depicted in the Last Glacial Maximum model suggested that conditions were conducive to widespread movement by the Mexican vole throughout the Sierra Madre and the northwestern Mogollon Rim of the Colorado Plateau. The climate conditions maintained Sierra Madre pine-oak forests (Toledo, 1982; van Devender, 1987), which supported the grassland habitat required by the species. Although modern climate conditions in the Mexican lowlands are largely unsuitable for dispersal of the Mexican vole, midden records indicate that pinyon-juniper woodlands occupied the present day Chihuahuan desert during the late Pleistocene (Betancourt et al., 2001; Holmgren et al., 2003). The Mexican vole is thought to readily disperse through wooded habitat (Harris, 1985; Davis and Callahan, 1987; Lomolino et al., 1989). In addition, a Pleistocene woodland corridor connected the southwestern United States and northwestern Mexico (Morafka, 1977; van Devender, 1987), and is a suspected dispersal corridor for several taxa (Castoe et al., 2007). The corridor may have facilitated gene flow between *M. mexicanus* populations between the northwestern

Mogollon Rim in Arizona and the Sierra Madre Occidental during environmentally conducive periods of the Pleistocene (Fig. 1.6).

The fragmentation of suitable habitat depicted in the model of current climate conditions may have changed the level of connectivity between vole populations, promoting diversification through vicariance. My analyses of genetic data, in which southern populations were characterized by longer branch lengths and earlier divergence estimates, also implicated the influence of vicariance. Assuming large size and isolation, the number of mutations is a function of the time elapsed since lineages diverged; as a consequence older populations were thought to harbor greater molecular diversity (Hoffman and Blouin, 2004). Vole populations in Mexico harbored greater molecular diversity for the cyt$b$ gene compared to populations in the United States (Table 1.1). Conversely, the populations sampled from Mexico demonstrated reduced molecular diversity for the AP5 gene compared to those in the United States. Under neutrality, the effective population size for nuclear introns and exons is expected to be approximately 4 times larger than that for cyt$b$. The larger than expected N_e reported from the AP5 gene for the United States samples is due to the larger than expected value for theta (Θ). The same population yielded a significant Fu and Li's D^* test (Table 5). As a result, the larger Θ noted for the population in the United States may be due to the effects of selection, and not a measure of diversification in *situ*. The older age of the southern populations was confirmed from mismatch analyses and estimates of clade divergence. A multitimodal haplotype distribution, indicative of long-term diversification *in situ*, characterized the subclades of Oaxaca and Durango. My simulations support a mid-Pleistocene divergence for populations in Mexico, and a Late Pleistocene divergence for the recently expanded

populations. These corroborate the 0.38-0.16 Ma divergence suggested from a morphological work (Frey, 1989) and fit within the time frame suggested for closely related *Microtus* (Conroy et al., 2001; Conroy and Cook, 2000).

Ecological niche models parameterized with modern climate conditions supported range expansion in the Mexican vole. The modern niche model predicted regions of suitable habitat in northwestern Arizona, central Arizona and New Mexico; areas not predicted as suitable in the Last Glacial Maximum model. This result supported that expansion from Arizona into New Mexico was possible and even likely after the Last Glacial Maximum (Fig. 1.6). My molecular analyses confirmed the recent expansion as suggested by the model. Relatively shallow branch lengths, later divergence, and lower nucleotide diversity characterize samples that represent more recent colonization events. I observed reduced molecular diversity for the clades in New Mexico, Arizona and Coahuila. I also recovered unimodal haplotype distributions for the Coahuila, New Mexico and northwestern Arizona subclades, which were also indicative of recent expansion.

Fossil evidence from plants and animals demonstrates the suitability of the Colorado Plateau as a source population for the Mexican vole. The relict pine forests of

Figure 1.6. Colonization and diversification scenario for *M. mexicanus* suggested from ecological and genetic data. Although post-LGM vicariance is suggested for the populations in Mexico, the Sierra Madre may have acted as a source region for expansion into northern Mexico (indicated as dashed arrow). Suggested dispersal between the Hualpai Mountains of northwestern Arizona and the SMOcc is indicated (solid bi-directional arrow). The post-LGM colonization of central/eastern Arizona and New Mexico is also indicated (one way solid arrow). The relative age of each clade is indicated by: mid-Pleistocene divergence (oval), pre-LGM divergence (diamond) and post-LGM divergence (square).   Map projection is geographic World Geodetic System 1984
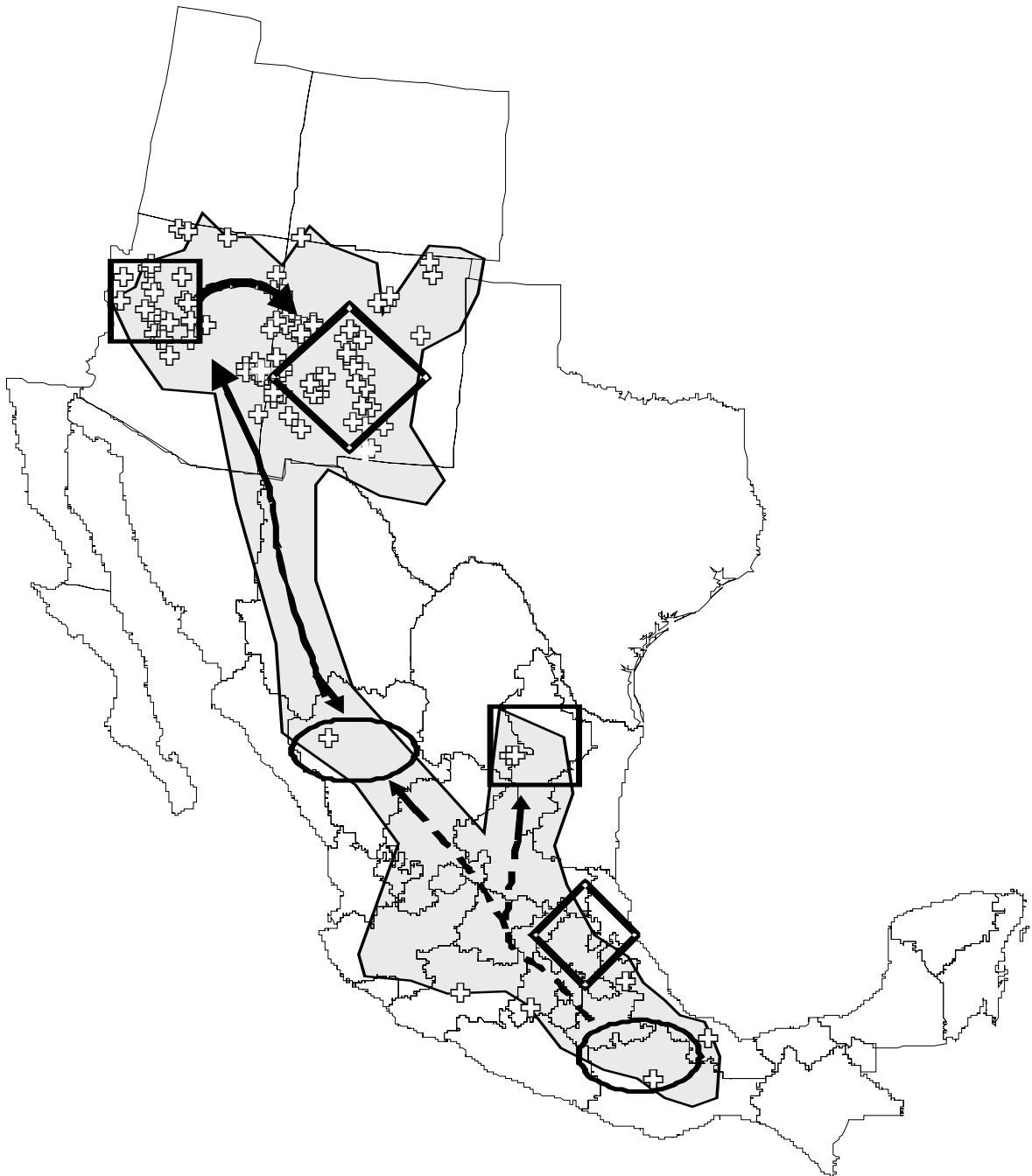
the Colorado Plateau (the Hualpai Mountains and Prospect Valley) would have been suitable late Pleistocene refugia for the Mexican vole (Davis and Callahan, 1992; Conroy et al., 2001), and fossil *M. mexicanus* were documented from the area (Jass, 2004). The deeper branch lengths and higher molecular diversity within the Colorado Plateau population compared to neighboring populations suggested that northwestern Arizona was a source population. The deeper branch lengths and earlier divergence estimates for vole samples from northwestern Arizona appeared to confirm the refugial hypothesis: the Mexican vole arrived in Arizona earlier to New Mexico, and expanded into New Mexico from Pleistocene refugia in Arizona in the late Pleistocene. This would also explain the suggestion that the Mexican vole was a recent arrival in New Mexico (Davis and Callahan, 1992; Findley et al., 1975). If woodland corridors encouraged continuity, northern and southeastern populations may have subsequently been bisected by the expansion of desert communities beginning in the late Pleistocene (Holmgren et al., 2003; Allen and Anderson, 2000).

I note that these analyses also indicated several taxonomic implications. The taxonomic status of the Mexican vole has been the subject of debate. Research into the karyotypes (Modi, 1987; Wilhelm, 1982; Judd, 1980; Lee and Endler, 1977) and morphology of *M. mexicanus* suggested the Mexican vole is a species complex. For example, Mexican voles north of 32° latitude in the United States possess a diploid chromosome number of 44, while those south of 32° latitude in Mexico have a diploid chromosome number of 48 (Modi, 1987). Based on morphology, Frey (1989) suggested the elevation of the distinct form found in the southwestern United States to *Microtus mogollonensis*. My results supported this proposal. Note that the two highly supported

46

clades recovered in this study are characterized by Kimura 2-parameter pairwise differences of between 7.13-10.82%; levels that define other recognized species of *Microtus* (Conroy et al., 2000). Moreover, I uncovered deep divergence between the populations in southern Mexico, which may also meet the criteria for species assignment (Fig. 6). My taxonomic conclusions are tentative and beyond the scope of this study, but it is interesting that other studies of co-distributed rodent taxa demonstrate divergence values similar to those for *M. mexicanus.* For example, the level of genetic differentiation observed between recognized species of *Neotoma mexicana* and *Sigmodon hispidus* approach that found between currently recognized species, including distinct clades in Mexico and the United States (Edwards and Bradley, 2002; Peppers and Bradley, 2000).

Results from this study supported the role of Pleistocene climate variability to the distribution of the Mexican vole in the mountains of the southwestern United States and Mexico. The bioregion is noted for its high levels of biodiversity. In the light of potential threats from human-mediated habitat and climate change, research that explores a mechanistic explanation for the origin and maintenance of the regions' biodiversity are of fundamental and timely importance. A cadre of abiotic influences is proposed to explain the regions' biodiversity, such as habitat heterogeneity, elevation and Pleistocene refugia (Fa and Morales, 1993). While a few intriguing studies of rodent taxa suggest the action of vicariant processes to diversity in the region (Lamb et al., 1987; Sullivan et al., 2000), a surprising paucity of studies have examined taxa distributed across the spatial extent of the domain. My analyses that link climate variability over the last 30,000 years to the phylogeographic structure in *M. mexicanus* suggested a plausible scenario for the disjunct distribution of the species. This approach could easily be applied in other studies of co-

distributed taxa such as *N. mexicana*, *Peromyscus melanotis*, *P. truei* and *Limoys irroratus* (Alvarez-Castenada, 2005; Durish et al., 2004) to add additional support for the influence of vicariance in this exceptional region.

ACKNOWLEDGEMENTS

Chapter II

Refugial dynamics of the montane vole, *Microtus montanus* in the Great Basin

and the western United States

The diverse mammalian fauna of the Great Basin have focused studies on extinction and colonization dynamics. Despite the suggestion that post-Pleistocene colonization occurred from nearby mountain ranges, the origin of mammalian taxa in the Great Basin is largely unknown. Source populations are expected to demonstrate higher levels of molecular diversity and demographic stasis compared to non-source populations. Here, I examine the Sierra Nevada and/or Rocky Mountains as source populations for colonization of the Great Basin through an analysis the cytochrome *b* gene (cyt*b*) of the mitochondrial DNA and the nuclear intron, acid phosphatase V (AP5) gene from the montane vole, *Microtus montanus*. The southern Sierra Nevada Mountain population demonstrated relatively high levels of cyt*b* diversity and demographic stasis. Coalescent simulations indicated that the southern Sierra Nevada might have acted as a source population based on the geographical distribution of haplotypes. The southern Sierra Nevada demonstrated only moderate diversity for the AP5. However, a haplotype network indicated the population was a central haplotype, related to adjacent populations through a hierarchy of connections. Results indicated expansion from the southern Sierra Nevada Mountains during the middle to late Pleistocene, and possible colonization of the Rocky Mountains and the eastern Great Basin during the late Pleistocene. This study highlights the utility of a molecular approach to elucidating the role of the Great Basin as colonization and diversification nexus in mammalian taxa.

INTRODUCTION

Since the Great Basin was first named in 1845, much has been learned about its climatic, hydrological and volcanic history. The volcanic legacy of the region is reflected in its topography: the Great Basin is bordered on the west by the Sierra, White and Inyo Mountains, to the east by the Wasatch Range and to the north by the Rocky Mountains (Maxson and Tikoff, 1996). A diverse vertebrate fauna is known from the area during the Pleistocene (Grayson, 1993). By contrast, there is a surprising lack of detailed examinations of extant Great Basin mammal fauna. In his treatise on non-equilibrium biogeography, Brown (1971) suggested that montane mammals first colonized the Great Basin Mountains from the Sierra Nevada and/or the Rocky Mountains. The initial conclusions outlined by Brown (1971) were modified with the addition of new mammal distribution data (Lawlor, 1998; Grayson and Livingston, 1993; Brown, 1971), but a question fundamental to the predictions of non-equilibrium biogeography remains largely untested: what is the source of Great Basin mammal fauna?

The ideal taxon to examine the source of Great Basin mammalian fauna is one that has a distribution that includes the Sierra Nevada and Rocky Mountains and also the Great Basin. The montane vole, *Microtus montanus* is a specialist in montane or intermontane meadow or grassland habitat at elevations that range from 1000-4500m (Findley and Jones, 1962; Anderson, 1959) from the Cascades and Sierra Nevada the Southern Rockies, the Colorado Plateau and the Great Basin (Fig. 2.1). The distribution is characterized by the presence of numerous peripheral populations along the southern margin (Hall, 1981). Most notable among these are populations found in the White

51

Figure 2.1 Locations of *Microtus montanus* in the western and southwestern
United States used in the phylogeographic study. Hypothesized colonization routed for
microtine rodents are: (A) a coastal route between 1.5-1.8Ma, and (B) an interior route
beginning about 0.085Ma (Repenning, 1990). The approximate distribution is represented
by known localities (blue squares) and includes the 7 clades examined: Sierra Nevada
(SNV), White/Inyo Mountains (WH/INYO), Amargosa Valley (ASHMDWS), eastern
Great Basin (EGRTBSN), northern Rocky Mountains (NROCKY), and southern Rocky
Mountains (SRCKY) and the mountains of the Colorado Plateau (COP). Samples
sequenced in molecular analyses are also shown (red squares). Map projection
is World Geodetic System 1984 Geographic.
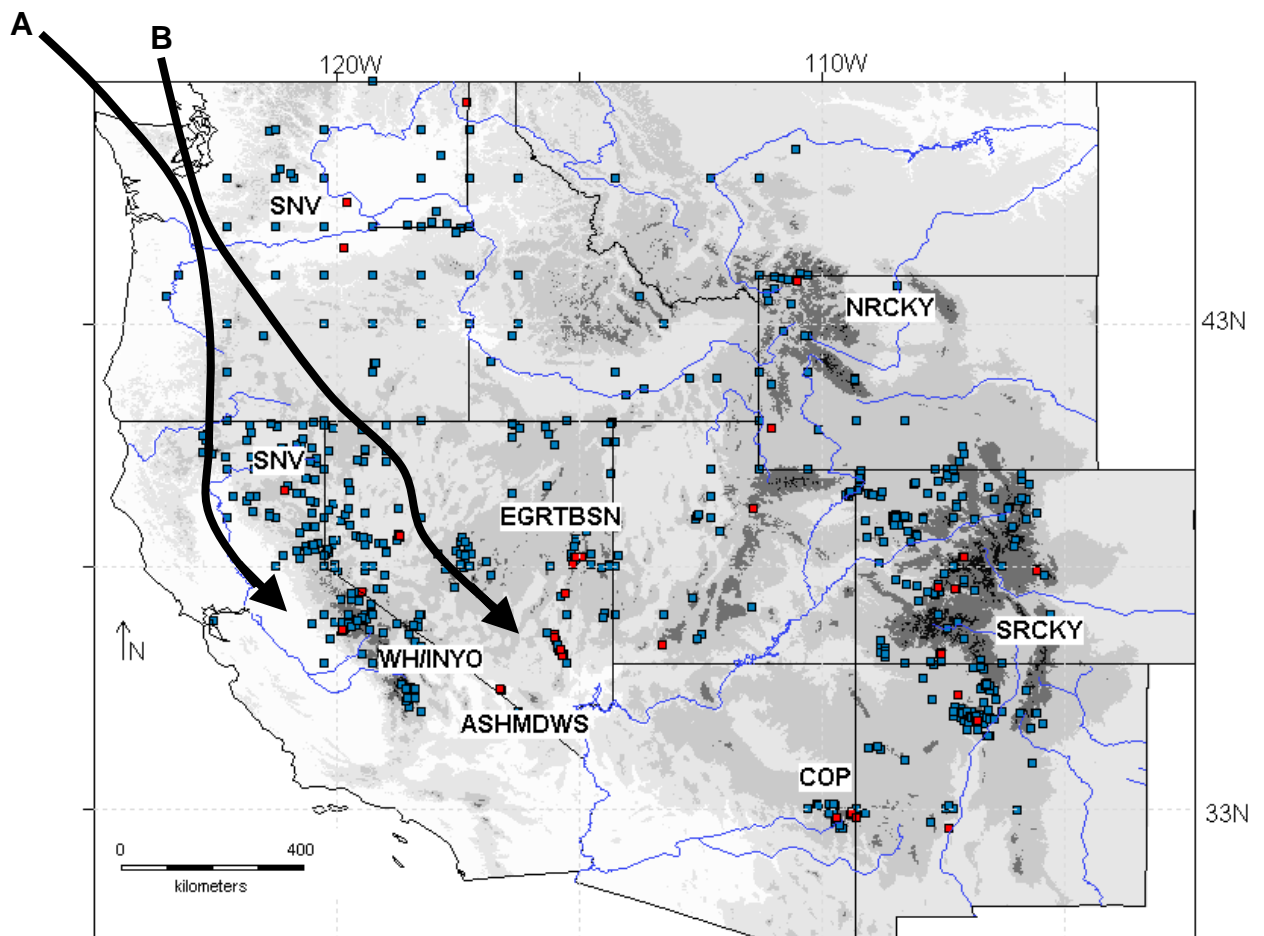
Mountains of southeastern Arizona and southwestern New Mexico, and those found along the southern edge of the Great Basin (Hall, 1981).

The curiously disjunct distribution of the montane vole is largely unexplained both in terms of its timing and its mechanism. Some previous workers believed that the montane vole colonized much of western North America during the early Pleistocene approximately 2.0Ma (Findley and Jones, 1962; Anderson, 1959) with inter-population divergence evident by the late Wisconsin, about 0.15Ma (Patterson, 1982). Alternatively, it has been proposed that ancestral *Microtus* colonized the United States from Asia along two temporally disjunct pathways: about 1.8Ma, animals moved along a western coastal route, and approximately 0.8Ma, animals colonized using a route east of the Sierra Nevada (Fig.2.1, Repenning, 1990). The general lack of knowledge concerning the species colonization history and dynamics of the western United States includes the possible source/s of extant montane vole lineages in the Great Basin.

The southern Rocky Mountains and Colorado Plateau (i.e. southern mountains) have been considered as the source of mammals in the Great Basin (Davis and Callahan, 1992; Brown, 1971). Paleoecological evidence suggests large areas of the Colorado Plateau were ice-free during the Pleistocene and the region probably functioned as a refugium for a diversity of animals (Baars, 2000; Davis and Callahan, 1992). A species response to ecological change will vary with parameters such as body size and habitat affinity (Patterson, 1982). As a result, species are likely to exhibit independent responses to dispersal opportunities (Carstens et al., 2005; Isaac et al., 2005; Faunmap Working Group et al., 1996), and it is unlikely that all extant taxa found in the region originated from glacial refugia in the southern mountains.

Mountain top habitat may have functioned as vital late Pleistocene refugia as animals moved up along an elevation gradient to escape the warming climate (Arbogast et al., 2001; Conroy and Cook, 2000; Hafner and Sullivan, 1995; Patterson, 1982; Findley, 1969). However, it is equally possible that species may have persisted in protected mountain valleys. Although glaciers covered the highest peaks in the Great Basin, many glaciers drained into the interior basin, which potentially sustained low elevation habitats (Osborn and Bevis, 2001). The impact of glacial expansion and contraction dynamics may have had a lesser impact on these protected areas, as found in other studies (DeChaine and Martin, 2004; Dyurgerov and Meier, 2000; Hewitt, 2000). Clearly, additional work is needed to examine the Great Basin as a post-Pleistocene source population.

Pleistocene refugia may leave one of two distinctive genetic signatures on species populations. A refugial population that was permanently isolated after the end of the Pleistocene will demonstrate regional structuring, higher levels of genetic diversity, and eventual monophyly. On the other hand, a refugium that acted as a source for post-Pleistocene colonization will be characterized by the absence of significant geographical structuring and decreased genetic diversity (Crespi et al., 2003). In the first case, genetic analyses documented the location of putative refugia (Arbogast et al., 2001; Conroy and Cook, 2000; Hafner and Sullivan, 1995; Patterson, 1982; Finley, 1969). In the latter case, genetic studies documented the presence of expanding lineages (Lessa et al., 2003; Hayes and Harrison, 1992). However, few studies have documented the location of the source population and its expanding lineages (see Steele and Storfer, 2006 for one exception). Establishing the linkage between refugia and expanding populations is essential to

reconstructing a species colonization history and dynamics (Wagner et al., 2005; Hewitt, 1996). Source populations tend to demonstrate high levels of allelic and haplotype diversity and demographic stasis (Rowe et al., 2004; Wright, 1943). Lineages that expanded from a source population are characterized by reduced intraclade differentiation (i.e. shorter branch lengths) compared to other populations (Rowe et al., 2004) and demographic expansion (Steele and Storfer, 2006; Haffer, 1969).

The cytochrome *b* gene of the mitochondrial DNA (cyt*b*) is commonly used to examine variation within and between species because of its relatively rapid coalescent time and ease of amplification, and has been used to identify putative Pleistocene refugia (e.g., Lessa et al., 2003). The relatively rapid mutation rate and smaller effective population size expected for the cyt*b* gene may limit its utility in resolving lineage relationships at deeper coalescent times. In addition, an inherent problem with the single gene approach is distinguishing gene trees from species trees (Avise, 2000). The lower mutation rate, larger effective population size and longer coalescence time of some nuclear genes (Galewski et al., 2006; Moore, 1995; Vawter and Brown, 1986) provide an additional way to potentially link expanding populations to their source.

Here, I examined the molecular diversity within *Microtus montanus* to reconstruct the species colonization history and to identify potential source areas for the Great Basin. To meet these objectives, I evaluated the genetic variation from the cyt*b* gene and the nuclear intron, the acid phosphatase V (AP5) gene. I first tested the hypothesis that the montane vole is composed of geographically unique and reciprocally monophyletic genetic lineages. I assessed the timing of colonization and diversification within the

species by estimating divergence for montane vole clades identified from phylogeographic analyses. Finally, I examined whether the distribution of geographically distinct lineages were consistent with range expansion from a source population in the Sierra Nevada Mountains using coalescent simulations.

MATERIALS AND METHODS

I examined the timing and mechanism of colonization in the montane vole using a molecular approach. I obtained samples representing 15 of the recognized 16 subspecies (Hall, 1981) for molecular analyses (Appendix F), which included 34 fresh *M. montanus* tissue samples collected from field surveys (D. L. Crawford and Nevada Division of Wildlife). An additional 25 frozen tissue samples were obtained from the University of New Mexico's Museum of Southwestern Biology Division of Genomic Resources, the Burke Museum at the University of Washington (Burke), and the Vertebrate Zoology at the University of California at Berkeley (MVZ). The subspecies *M. montanus nevadensis* was sampled from 5 museum skin samples from MVZ. The distribution of samples covers a significant portion of the species known geographic range, including populations east and west of the Sierra Nevada Mountains, the Rocky Mountains, the Colorado Plateau and the eastern Great Basin (Fig. 2.2, Appendix F).

Total nucleic acids were extracted using a modified salt extraction procedure (Medrano et al., 1993). To quantify molecular variation in *M. montanus*, I amplified and sequenced the cyt*b* gene and the nuclear intron acid phosphatase V (AP5). The cyt*b* gene is recognized as a data rich site, particularly for nodes younger than 2 Ma (Galewski et al., 2006). I amplified the cytochrome *b* gene (cyt*b*) using the polymerase chain reaction (PCR), and primer pairs MVZ05-MICRO-06 and ARVIC07-VOLE14 (Hadly et al.,

Figure 2.2.The distribution of the montane vole (redrawn from Hall, 1981). The approximate distribution is represented by known localities (blue squares) and includes the 7 clades examined: Sierra Nevada (SNV), White/Inyo Mountains (WH/INYO), Amargosa Valley (ASHMDWS), eastern Great Basin (EGRTBSN), northern Rocky Mountains (NROCKY), and southern mountains of the southern Rocky Mountains (SRCKY) and the mountains of the Colorado Plateau (COP). Map projection is World Geodetic System 1984 Geographic.
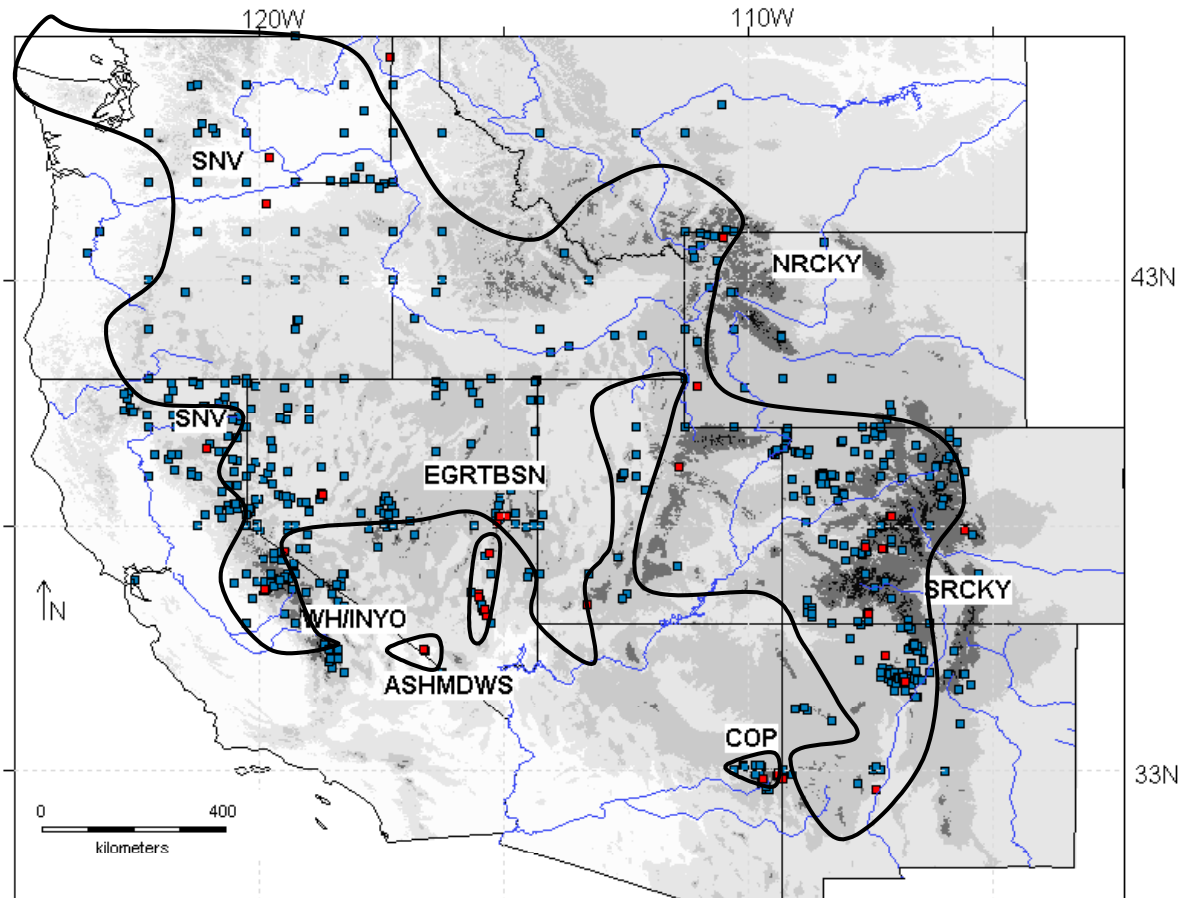
2002). Reaction conditions for PCR were 95oC (30 s), denaturing, 50oC (25 s) annealing,

I characterized the genetic diversity within and between populations of the montane vole

from the cyt*b* gene of the mitochondrial DNA (cyt*b*) and the nuclear intron, acid

phosphatase V (AP5). I sequenced 724bp of the cytochrome *b* gene from 64 samples of

M. montanus (Appendix F) using dye-labeled terminators and cycle sequencing

conditions of 96oC (10 s) denaturing, 50oC (5s) annealing, and 60oC (4 min) extension.

PCR products were sequenced using the same primer sets as above, and purified using a

sodium acetate and ethanol precipitation procedure. Sequence samples were read using an

ABI 3100 Automated Sequence Analyzer. Sequences were manually aligned using

Sequencher v.4.7 (Bromberg, 1995).

I also amplified the AP5 nuclear intron. I sequenced 433bp of AP5 from 29

samples to test for phylogeographic congruence across multiple molecular markers

(Appendix F). Variation in the AP5 gene was recently examined in a phylogeographic

study of *Microtus californicus* (Conroy and Neuwald, 2008). The samples represented the

Sierra Nevada Mountains, the eastern Great Basin, the Rocky Mountains, the Colorado

Plateau, and the White/Inyo region. An AP5 sequence of *Microtus californicus* (Conroy

and Neuwald, 2008) was used as an outgroup. I amplified and sequenced the AP5 gene

using gene specific primers (Debry and Sheshadri, 2001) and reaction conditions similar

to that for the cyt*b* data.

Phylogeographic analyses

The level of geographic distinctiveness and reciprocal monophyly was assessed

through separate phylogenetic and phylogeographic analyses of the cytochrome *b*

and AP5 genes. Because of unequal sample sizes, I conducted separate analyses of the cytb and AP5 genes.

I rooted the genealogical structure in the cyt*b* gene of *M*. *montanus* with *M*. *pennsylvanicus*, which is the suggested sister taxon to *M*. *montanus* (Conroy and Cook, 2000). The model of evolution selected using MODELTEST (Posada and Crandall, 1998) was the HKY +I+ G model (-ln*L*=2512.2903) with 6 rate parameters. Parameters estimated for this model included empirical base frequencies (A=0.3028, C= 0.2874, G= 0.1200, and T= 0.2897) and a transition/transversion ratio of 4.416. The values for the proportion of invariable sites and the gamma shape parameter were 0.5210 and 0.6631, respectively.

I examined reciprocal monophyly of montane vole populations via tree topology, which was estimated in MRBAYES v.3.1.2 (Huelsenbeck and Ronquist, 2001) using the selected model parameters. Bootstrap support for the likelihood topology was estimated using the Genetic Algorithm for Rapid Likelihood Inference in GARLI v.0.96b8 (Zwickl, 2006; Felsenstein, 1985) using 100 replicates, and the selected model parameters.

I expected that source populations would exhibit higher levels of genetic differentiation compared to other populations. To examine this expectation, I first estimated the amount of molecular variation within and between populations using the Kimura 2-parameter (Kimura 2P) distance, which estimates the pairwise distances between sequences while correcting for the different rate between transitional and transversional substitutions. I quantified the genetic variation in the software program DnaSP (Rozas et al., 2003). I also expected that source populations would demonstrate multiodal haplotype distributions and demographic stasis. The haplotype distribution of

non-expanding populations is expected to be multimodal because many haplotypes are maintained over time (Harpending, 1994). I assessed the haplotype distribution in ARLEQUIN v. 3.11 (Excoffier et al., 2006) to distinguish refugial populations from expanding lineages.

The reduced mutation rate, larger effective size and longer coalescent times expected for nuclear genes may provide an independent source of information on species the temporal dynamics of colonization, isolation and diversification. The model of evolution that best fit the AP5 data was selected in MODELTEST (Posada and Crandall, 1998) as the HKY +I+G model. The HKY model distinguishes between transition substitution (α) and transversion substitutions (β) and allows unequal base frequencies ($\Pi_A \neq \Pi_T \neq \Pi_C \neq \Pi_G$), which yields a model with 6 free parameters. Parameters estimated for this model included empirical base frequencies (A=0.3531, C= 0.3336, G= 0.0806, and T= 0.2327) and a transition/transversion ratio of 5.53. The values for the proportion of invariable sites and the gamma shape parameter were 0.5207 and 0.9725, respectively.

I assessed reciprocal monophyly of montane vole clades for the AP5 gene by constructing a haplotype network in TCS v.1.2.1 (Clement et al., 2000). I further examined the congruence between mitochondrial and nuclear gene data by assessing the phylogenetic structure of the montane vole for the AP5 nuclear intron in MEGA v.4 (Tamura et al., 2007). I also quantified molecular diversity parameters such as nucleotide diversity (π), theta (Θ) DnaSP (Rozas et al., 2001). Similar to the cyt*b* analysis, I examined whether the populations violated neutrality at the AP5 gene through a Fu's *Fs* test (Fu, 1997), a Fu and Li's *D** test (Fu and Li, 1993) and a Tajima's *D* test (Tajima, 1989). Each test is sensitive to different violations of neutrality, and as such can help

determine the potential processes influencing genetic structure for species populations.

Divergence estimates

If southwestern montane vole populations represent Pleistocene relicts, then divergence should precede the end of the Pleistocene. I tested the assumption of rate homogeneity among sequences by comparing the likelihood values for cyt*b* topologies with and without a molecular clock enforcing *s-2* degrees of freedom (Felsenstein, 1981).

I calculated the divergence estimates for each clade using cyt*b* data. The algorithm in BEAST V.1.4.6 searches the tree landscape using a Bayesian Markov chain Monte Carlo algorithm to estimate divergence times between clades under a relaxed molecular clock assumption. The parameters estimated in BEAST method include the time of the most recent common ancestor (T_{MRCA}), the highest posterior density for the T_{MRCA} (a measure of the variance around the T_{MRCA} estimate), and the effective sample size (ESS), which reflects the approximate number of independent samples drawn from the posterior distribution in the Markov Chain. An ESS value of ≥ 100 is considered necessary for the estimation of parameters, and the higher the ESS value, the less likely that autocorrelation is influencing the parameters.

The purpose of relaxed molecular clocks is to obtain divergence estimates while allowing differential evolutionary rates within lineages. When fossils are available to serve as external calibration points, model priors are constrained to be lognormal. In this way, heterogenous rates of evolution are allowed to vary along the entire branch, rather than concentrating the rate on the node, which is suggested to yield more robust estimates (Ho, 2007). I estimated the divergence between clades by setting the divergence between *Microtus* and *Myodes* using a lognormally distributed prior of 1.3 ± 0.2 Ma based on

61

fossil evidence (Chaline and Graf, 1988; Catzeflis et al., 1987; Zakrzewski, 1985), and an internal lognormal prior representative of divergence within *Microtus* at 0.5 ± 0.2 Ma (Steppan et al., 2004; Conroy and Cook, 1999; Repenning, 1990). I set the search to 10 million generations.

Coalescent simulations

The degree of genetic variation present in a refugial population is probabilistically associated with its function as a source population. I reconstructed the most likely ancestral node for the sampled populations using coalescent simulations on the cyt*b* data in MESQUITE v.2.5 (Maddison and Maddison, 2009) to examine the likelihood that the southern mountains served as a source population. Each individual montane vole in the dataset was coded to each of 5 geographical divisions identified from the phylogenetic analyses. The ancestral regions were elucidated using a Markov model (Mk1) on the tree topology. Because the rate of dispersal for the montane vole is not known, I used an equal likelihood for the rate of change on the tree. The likelihood that a region served as a source population is represented as a proportional likelihood (PL).

RESULTS

Sampling summary

I examined the genetic variation among 63 cyt*b* samples. I found 29 distinct haplotypes; 14 haplotypes were represented from populations in the western Great Basin. The most frequently sampled haplotype, designated as 'Step15', was reported from 9 individuals within the eastern Great Basin clade (Fig. 2.3, Appendix G). Sites that are missing characters or gaps are ignored in phylogenetic analyses in ARLEQUIN and MRBAYES. However, I excluded from all analyses 2 samples that were missing large

portions of data.

Phylogeographic analyses

My phylogenetic analyses of the cyt*b* gene yielded 159 parsimony informative characters. The best fit topology strongly supported three principal clades within *M. montanus*: an eastern, southern and western division (Fig. 2.3). The eastern clade was composed of the eastern Great Basin and the northern Sierra Nevada Mountains. Within the division, samples from Steptoe Valley and Kirch Wildlife Management Area comprised one group, and samples from Pahranagat Valley and Key Pittman Wildlife Management Area comprised a second group. The southern clade was composed of samples from the southern Rocky Mountains, Arizona and New Mexico. The western clade was composed of animals from the White Mountains and the southern Sierra Nevada Mountains. The western clade appeared to meet the expectations for a Pleistocene source population: branch lengths were longer, indicative of greater genetic differentiation relative to other clades (Fig. 2.3). A fourth clade includes samples collected in 1933 from a recently extirpated population of *M. montanus* from Amargosa Valley of southwestern Nevada.

My analyses demonstrated genetic divergence between vole populations for the cyt*b* gene. Genetic differentiation, measured as Kimura 2P pairwise distances, suggested minimal differentiation within clades (Table 2.1a) but values of approximately 2-5% characterized most inter-clade distances (Table 2.1b). The average haplotype diversity for the data is 0.928 ± 0.038 (range 0-1), and haplotype diversity values approaching or equal to 1 are reported for the Sierra Nevada, Colorado Plateau and the southern Rocky Mountains (Table 2.2). Populations sampled from the Sierra Nevada demonstrated the

63

Figure 2.3 The Bayesian phylogram of 65 *M. montanus* samples. The analysis identifies (3) primary clades within *M. montanus*: an eastern clade comprised of samples from the eastern Great Basin and the northern Sierra Nevada Mountains (horizontal striped bar), a southern clade made up of samples from the Colorado Plateau, the mountains of New Mexico and the southern Rocky Mountains (diamond bar), a clade composed of Ash Meadows, the northern Rocky Mountains and the northern Sierra Nevada (bricked bar) and a western clade composed of samples from the southern Sierra Nevada, White and Inyo Mountains (diagonal bar). Each clade is represented by Bayesian posterior probability (above bar) and bootstrap values (below bar).
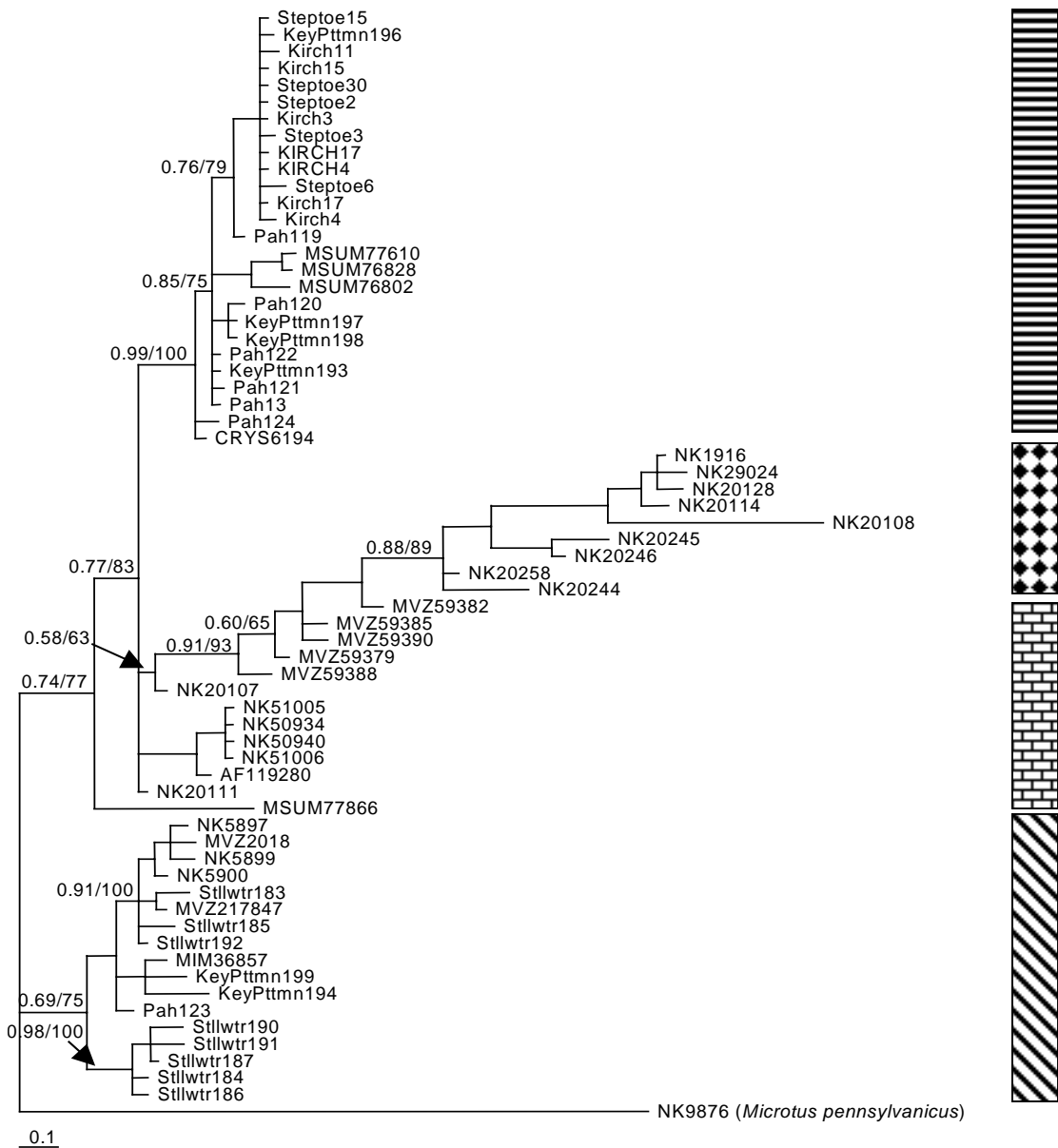
highest nucleotide diversity (π) and the highest number of haplotypes compared to other populations (Table 2.2).

If populations in the southern Rocky Mountains and/or Colorado Plateau were post-Pleistocene source population, the respective pairwise haplotype distributions should be multimodal, indicative of a diversification *in situ*. My analyses of the haplotype distribution of each population indicated that expansion was not supported for the Colorado Plateau or populations in the southern Sierra Nevada Mountains. Demographic expansion was indicated for the eastern Great Basin and the northern Rocky Mountains (Table 2.3).

Molecular diversity parameters estimated from the AP5 gene also indicated division among clades of the montane vole. Samples from the Sierra Nevada demonstrated relatively higher nucleotide diversity compared to other clades (Table 2.4, Appendix H), potentially indicative of long-term stasis and/or a larger refugial population size (Table 2.4). A minimum spanning network of AP5 haplotypes depicted a network in which samples from the Sierra Nevada and White/Inyo Mountains were centrally located, and connected to samples from other regions (Fig. 2.4). My analyses of population haplotype distributions for the AP5 gene demonstrated that neutrality was supported, and the null hypothesis of demographic expansion was rejected for all populations (Table 2.5). Differences in the value for *r* between mitochondrial and nuclear data may be due to differences in sample size (Tables 2.3, 2.5).

Divergence estimates

I estimated the divergence of each montane vole population to place molecular

65

Table 2.1a. Kimura 2P pairwise differences (D) within clades of the montane vole. The clades examined included the White/Inyo Mountains (Wht/Inyo), the Colorado Plateau (COP), the mountains of western Nevada and California (WGrtBsn), Pahranagat Valley in southern Nevada (Pah), the southern Rocky Mountains (SRcky), Steptoe Valley (Stptoe), Kirch Wildlife Management Area (Kirch), the western Sierra Nevada Mountains (SNev), the northern Rocky Mountains (NRcky), Key Pittman Wildlife Management Area (KP) and Amargosa Valley in southwestern Nevada (AshMdws).

Clade	D
Wht/Inyo	0.009
COP	0.036
Pah	0.011
SRcky	0.017
Stptoe	0.002
Kirch	0.001
SNev	0.019
NRcky	0.001
KP	0.019
AshMdws	0.013

Table 2.1b. Kimura 2P pairwise differences between clades of the montane vole. Population abbreviations are the same as those in Table 1a.

1	2	3	4	5	6	7	8	9	10
1 Wht/Inyo									
2 COP	0.043								
3 WGrtBsn	0.010	0.040							
4 Pah	0.034	0.039	0.027						
5 SRcky	0.031	0.034	0.028	0.030					
6 Stptoe	0.042	0.045	0.038	0.013	0.047				
7 Kirch	0.042	0.045	0.033	0.015	0.038	0.001			
8 SNev	0.036	0.042	0.028	0.019	0.038	0.013	0.013		
9 NRcky	0.032	0.034	0.026	0.015	0.037	0.025	0.025	0.020	
10 KP	0.026	0.042	0.032	0.017	0.040	0.019	0.020	0.024	0.022
11 AshMdw	0.033	0.037	0.029	0.032	0.026	0.039	0.040	0.037	0.029	0.029

Table 2.2. Genetic diversity parameters calculated for the cyt*b* gene from 7 clades of the montane vole. Clades are characterized by differences in the number of polymorphic loci (*S*), nucleotide diversity (π), theta (Θ), the average number of pairwise differences *(k)* and the number of haplotypes (h). Abbreviations for populations are as follows: the Colorado Plateau and the mountains of New Mexico (COP), the White and Inyo Mountains of southern California and Nevada (Wht/Inyo), the Sierra Nevada (SNev), the eastern Great Basin (EGrtBsn), the southern Rocky mountains in Colorado (SRcky), the northern Rocky Mountains in northern Wyoming (NRcky), and an Extirpated population from Ash Meadows in the Amargosa Valley of southwestern Nevada (AshMdws). Genetic parameters included the number of polymorphic sites(*S*), nucleotide diversity based on the number of pairwise differences (π), nucleotide diversity based on the expected level of homozygosity (Θ), the number of pairwise differences (*k*), and the number of haplotypes (h) for each population. Sample size (N) also provided.

Clade	N	*S*	π	Θ	*k*	h
COP	7	18	0.013 ± 0.006	0.015	7.6	5
Wht/Inyo	5	16	0.005 ± 0.003	0.004	6.57	8
SNev	13	15	0.011 ± 0.003	0.012	5.3	15
EGrtBsn	13	22	0.011 ± 0.009	0.011	6.26	8
SRcky	4	18	0.017 ± 0.009	0.016	10.5	4
NRcky	4	0	0.007 ± 0.005	0.007	2.5	1
AshMdws	5	9	0.014 ± 0.008	0.015	4.8	5

Table 2.3. Tests of neutrality for 7 clades of the montane vole while controlling for sample size for the cyt*b* gene. Parameters calculated include measure of the fit of the data to a model of population expansion (r), and the associated *P*-value, *Fu's Fs* test of neutrality (*Fs*, Fu, 1997) and the Tajima's *D* test (*D,* Tajima, 1989). Under a model of population expansion, *P*-values for the raggedness index are expected to be nonsignificant (*P*> 0.05), and *Fu's F* and Tajima's D tests are expected to be significantly negative. Significant results (*) indicate the populations for which a neutrality test was violated. Clade abbreviations are the same as in Table 2.

Clade	*r* (p-value)	*Fs*	*D*
COP	0.144 (0.000)	-0.74	-0.40
Wht/Inyo	0.345 (0.000)	0.42	-0.04
SNevada	0.176 (0.011)	0.43	-0.03
EstGrtBsn	0.249 (0.620)	**2.35***	-0.33
SRocky	0.277 (0.670)	0.59	0.43
NRcky	0.680 (0.460)	**-4.29***	0.43
AshMdws	0.467 (0.000)	0.48	1.12

* significant at P<0.05

Table 2.4. Molecular diversity parameters estimated from 433bp of the nuclear intron AP5. Clades are characterized by differences in the number of polymorphic sites (S), nucleotide diversity (π), theta (Θ) and the average number of pairwise differences (k). The sample size for each clade is also shown (N). Clade abbreviations are as follows: populations sampled from the southwest and the southern Rocky Mountains (RckyMtns), the mountains of California and western Nevada (Sierra Nevada), populations from the middle eastern Great Basin (Key Pittman), and populations in the eastern Great Basin (EstGrtBsn).

Clade	N	S	π	Θ	k
RckyMtns	4	21	0.031 ± 0.017	0.031	11.5 ± 2.56
Sierra Nevada	6	13	0.013 ± 0.005	0.014	5.6 ± 3.92
Key Pittman	5	25	0.035 ± 0.011	0.035	12.5 ± 2.08
EstGrtBsn	14	37	0.034 ± 0.005	0.040	10.58 ± 3.74

diversity within a temporal context. Because the assumption of rate homogeneity among sequences was violated, I calculated divergence estimates with a relaxed phylogenetics approach that co-estimates phylogeny and divergence as implemented in BEAST v.1.4.6 (Drummond et al., 2006). Clade divergence estimated using a relaxed molecular clock approach on the cyt*b* gene indicated the earliest divergence for the western clade, which occurred between 0.80-1.5Ma (Table 2.6). The eastern Great Basin and the southern Rocky Mountain clades demonstrated a late Pleistocene divergence, which averaged 0.433Ma and 0.273Ma, respectively (Table 2.6)

Coalescent simulations

My simulations of ancestral nodes indicated that the amount of molecular variation that was traced back to the southern Sierra Nevada was the result of its role as a source population (PL = 0.74, P = 0.02, node 9 in Fig. 2.5, Appendix I). The eastern Great Basin and the southern Sierra Nevada were indicated as significant sources for vole colonization of the northern Sierra Nevada Mountains, the eastern Great Basin, the Rocky Mountains and Colorado Plateau (node 10, Fig. 2.5, Appendix I), although the eastern Great Basin was less likely (PL = 0.22, P=0.04). My simulations also indicated that the refugial population in the southern Sierra Nevada Mountains was a likely source population for post-Pleistocene colonization of the northern Sierra Nevada Mountains, the eastern Great Basin and the northern Rocky Mountains (P = 0.02, Appendix I, Fig. 2.5).

Figure 2.4 The haplotype network for the nuclear intron AP5. The number of substitutions that separate each haplotype cluster are indicated as numbers adjacent to lines perpendicular to the lines connecting the network. The populations sampled included: the southern Sierra Nevada (diagonal slashing), the eastern Great Basin (horizontal lines), the southern Rocky Mountains (diamonds) and samples from White/Inyo Mountains (irregular polygons).
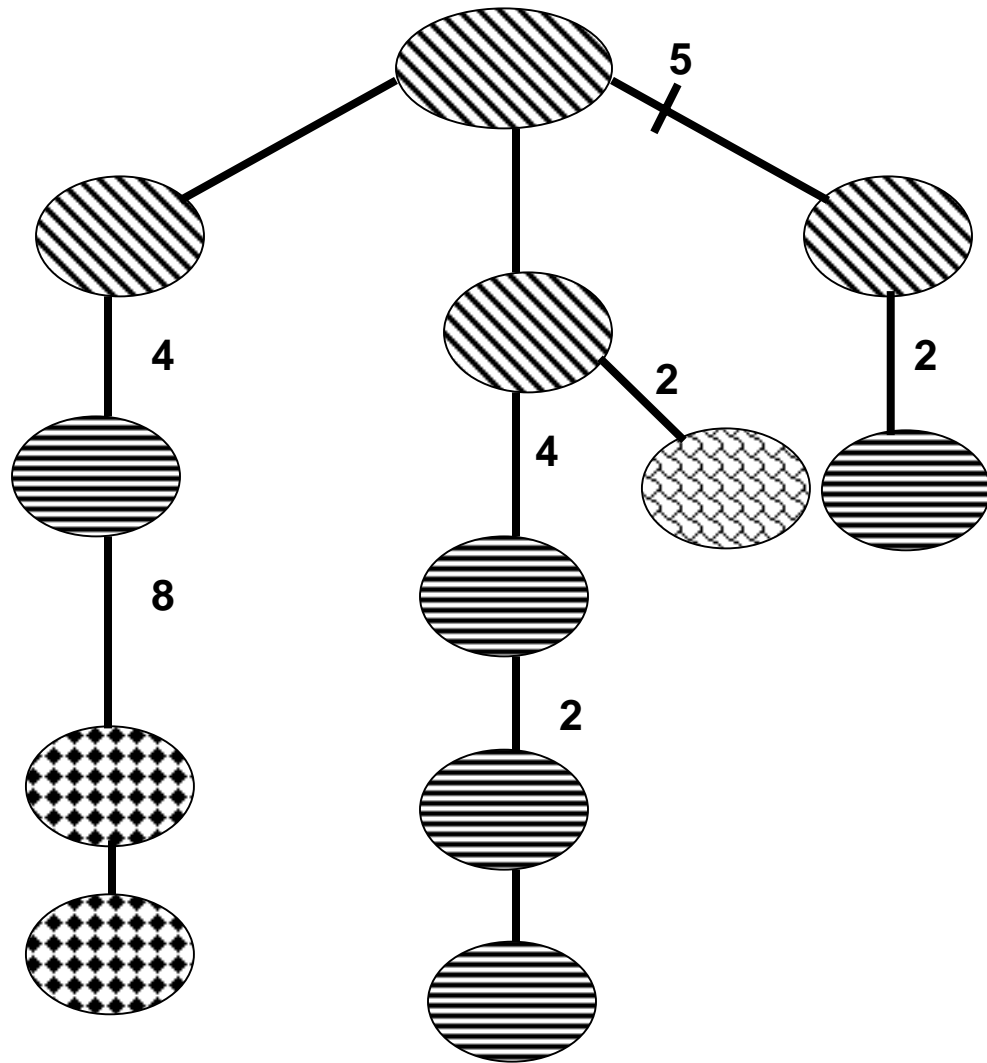
Table 2.5. Tests of neutrality for the nuclear intron AP5 from 4 clades of the montane vole. Parameters calculated include measure of the fit of the data to a model of population expansion (*r*) and the associated *P*-value, *Fu's F* test of neutrality (*Fs,* Fu, 1997), Tajima's *D* test (*D,* Tajima, 1989), and the number of unique sites (sites). Under a model of population expansion, *P*-values for the raggedness index are expected to be non-significant (*P*> 0.05), and *Fu's Fs* and Tajima's *D* tests are expected to be significantly negative. Significant results (*) indicate the populations for which a neutrality test was violated. Clade abbreviations are as in Table 4. Sample size (N) is also indicated.

Clade	N	*r* (P-value)	*Fs*	*D*	sites
RckyMtns	4	0.39 (0.00)	1.24	-0.05	21
Sierra Nevada	6	0.13 (0.03)	0.79	-0.07	6
Key Pittman	5	0.34 (0.00)	1.37	-0.11	37
EstGrtBsn	14	0.05 (0.03)	0.09	0.07	40

Table 2.6. Divergence estimates for primary clades within the montane vole. I estimated divergence using a lognormally distributed prior of 1.3 ± 0.2 Ma based on fossil evidence (Chaline and Graf, 1988; Catzeflis et al., 1989), and an internal lognormal prior representative of divergence within *Microtus* at 0.5 ± 0.2 Ma The mean divergence (T_{MRCA}), effective sample size (ESS) and range of divergence estimate (highest posterior density) are reported. An ESS value of ≥100 is considered appropriate for robust parameter estimation (Drummond et al., 2006). Clade abbreviations are as in Table 2.4.

Clade	Mean	ESS	+/-
COP	0.19	243.32	0.09-0.31
SRcky	0.27	114.74	0.10-0.47
EGrtBsn	0.43	125.21	0.15-0.81
NRcky	0.02	473.53	0.001-0.03
WGrtBsn	1.16	460.32	0.80-1.5
AshMdws	0.15	240.32	0.07-0.22

Figure 2.5. Results from coalescent simulations of the cyt*b* gene. The tree depicts the summary of divergence times and the most likely source area for populations of the montane vole. Mean divergence dates for primary clades are shown (above bar) with the 95% confidence intervals given (below bars) for the major clades of the montane vole (Table 6). The capital letters at each node represent the most likely geographical location for the source of each clade, reconstructed using proportional likelihoods (Appendix I).
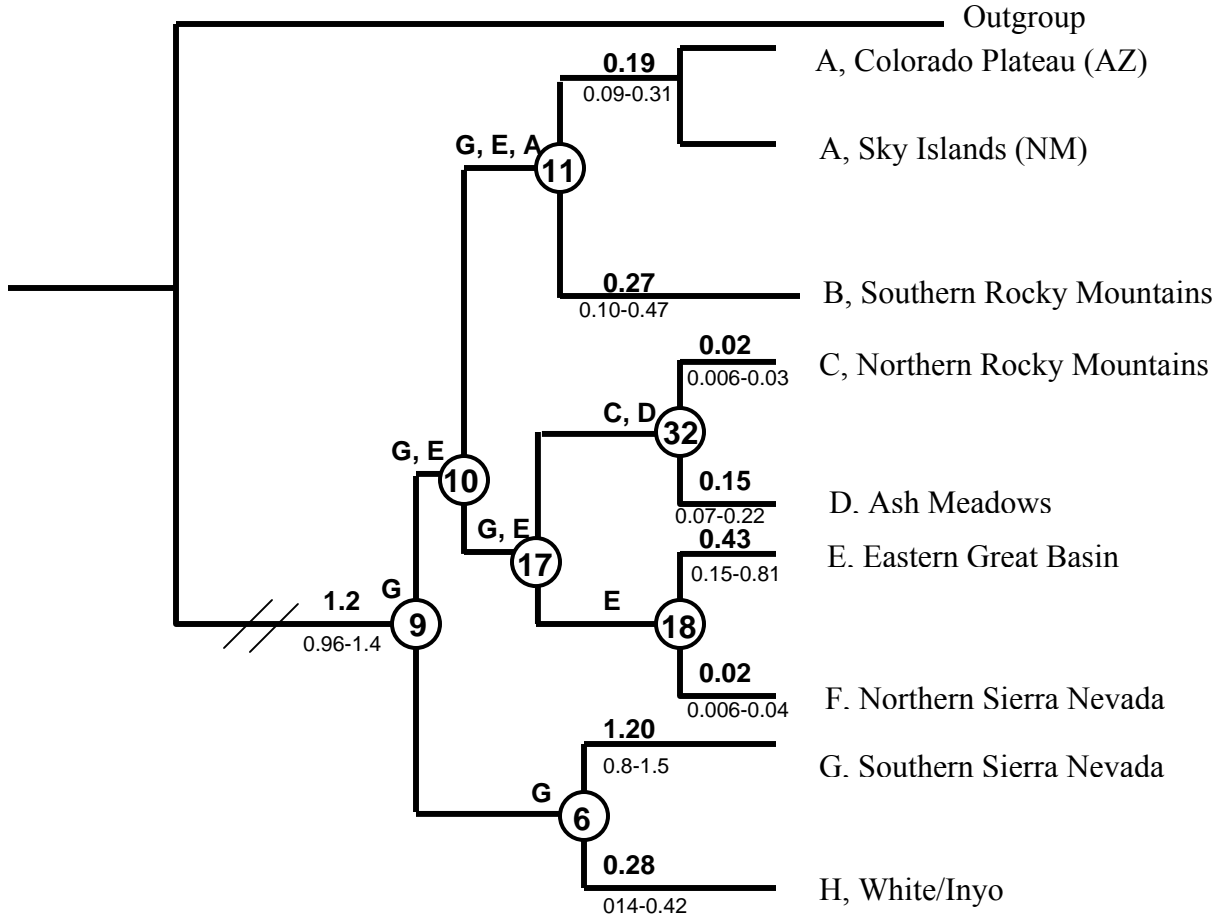
DISCUSSION

Results suggested the southern Sierra Nevada Mountains acted as a source population for post-Pleistocene colonization into neighboring regions such as the eastern Great Basin and northern Rocky Mountains. While my phylogeographic analyses of the mitochondrial DNA suggested that *M*. *montanus* colonized the western United States along two colonization routes, the results from the nuclear DNA indicated a different, albeit complementary scenario. The expected differences in mutation rate, effective population size and coalescence time between mitochondrial and nuclear DNA imply that data from the molecular markers may offer information at disparate temporal scales. This presumed difference allowed me to reconstruct a more holistic view of the colonization history of the montane vole, including a possible colonization scenario and the linkage between a post-Pleistocene source population and expanding lineages. As such, this study supports the assertion made by Brown (1971) that the Sierra Nevada region was an important colonization source for the Great Basin during the middle to late Pleistocene.

Source populations are thought to demonstrate higher levels of haplotype and allelic diversity compared to non-source populations (Rowe et al., 2004; Wright, 1943). For refugia that subsequently functioned as source populations, the increased levels of diversity are thought to be the result of a large refugial population size, rather than diversification in isolation. A large refugial population size may explain the elevated haplotype diversity and polymorphic loci within the Sierra Nevada population (Table 2.2).

Two possible reasons may explain the reduced levels of diversity for the AP5 intron for the Sierra Nevada samples. First, if the effective population size of the Sierra

Nevada was smaller relative to other populations, the small size coupled with the reduced

mutation rate of nuclear introns could result in reduced diversity. It was also possible that

the higher diversity for the eastern Great Basin clade was the result of colonization from

multiple source populations not detected in this study.

My analyses of phylogeographic structure for the cyt*b* and AP5 genes suggest

differing, albeit complementary explanations for the pattern of diversity in montane vole

populations. A phylogeographic analysis of cyt*b* showed that the voles sampled from the

Sierra Nevada and White/Inyo Mountains of southern California and Nevada composed a

distinct clade, which resulted in a split between the western and eastern Great Basin (Fig.

2.3). A similar phylogeographic analysis of the AP5 data demonstrated that samples from

the southern Sierra Nevada were at the center of a haplotype network, linked

hierarchically to additional populations (Fig. 2.4). There were two possible mechanisms

that would link the cyt*b* and AP5 topology: geological variation and disparate

colonization events.

The geological history of the Great Basin offered one possible explanation for the

east-west division from the cyt*b* gene. Beginning in the Miocene, the region was

geologically divided into an eastern paleovalley that drained to the east into the Uinta

Basin and a western paleovalley that drained into the Pacific Ocean (Henry, 2008).

The eastern and western basins were characterized by different bedrock: carbonate rock

underlies the eastern region, which favored ground water flow, and the metapmorphic

rock under the western basin impeded ground water flow (Henry, 2008; Reheis et al.,

2002; Harrill and Prudic, 1998; Plume, 1995). The paleovalley division may not be a

complete explanation for montane vole clade structure. But given the riparian habitat

specificityof the species, the dichotomous ground water conditions of the eastern and western Great Basin may have influenced dispersal route that was available at different times.

Fossil evidence offered a second possible explanation for the phylogeographic structure demonstrated for the montane vole. The east/west division suggested two colonization events were responsible for this topology: one for the Sierra Nevada/White/Inyo group, and another for the second clade composed of the northern Sierra Nevada, eastern Great Basin, Colorado Plateau and Rocky Mountains (Fig. 2.3). This result appears to agree with the colonization scenario suggested by Repenning (1990), in which ancestral *Microtus* used an early coastal route along the western flank of the Sierra Nevada Mountains up to approximately 1.2-1.5 Ma, and a second interior route between the Sierra Nevada and Rocky Mountains around 0.80Ma (Fig. 2.1).

The combined results from the cyt*b* gene and the AP5 gene suggested a possible colonization scenario for the species (Fig. 2.6). The hierarchical connections observed in the AP5 network suggest historical gene flow between populations in the southern Sierra Nevada Mountains and those in the northern Rocky Mountains, the eastern Great Basin and the White/Inyo Mountains. Additionally, the deeper branch lengths, relatively high haplotype diversity and demographic stasis demonstrated for the southern Sierra Nevada clade for the cyt*b* gene appeared to support its role as source populations. The middle Pleistocene divergence indicated for the southern Sierra Nevada and White/Inyo clades suggested that an earlier colonization event produced this topology (Repenning, 1990). A later wave of colonization may have linked vole populations in the southern

Figure 2.6. Colonization scenario suggested by analysis of the nuclear intron AP5. The nuclear and mitochondrial data suggested that vole populations in the southern Sierra Nevada Mountains acted as a source for post-Pleistocene colonization of the White/Inyo Mountains during the middle to late Pleistocene, and for the northern Sierra Nevada Mountains, the eastern Great basin and the northern Rocky Mountains during the late Pleistocene. The relative directions of colonization are indicated as arrows.
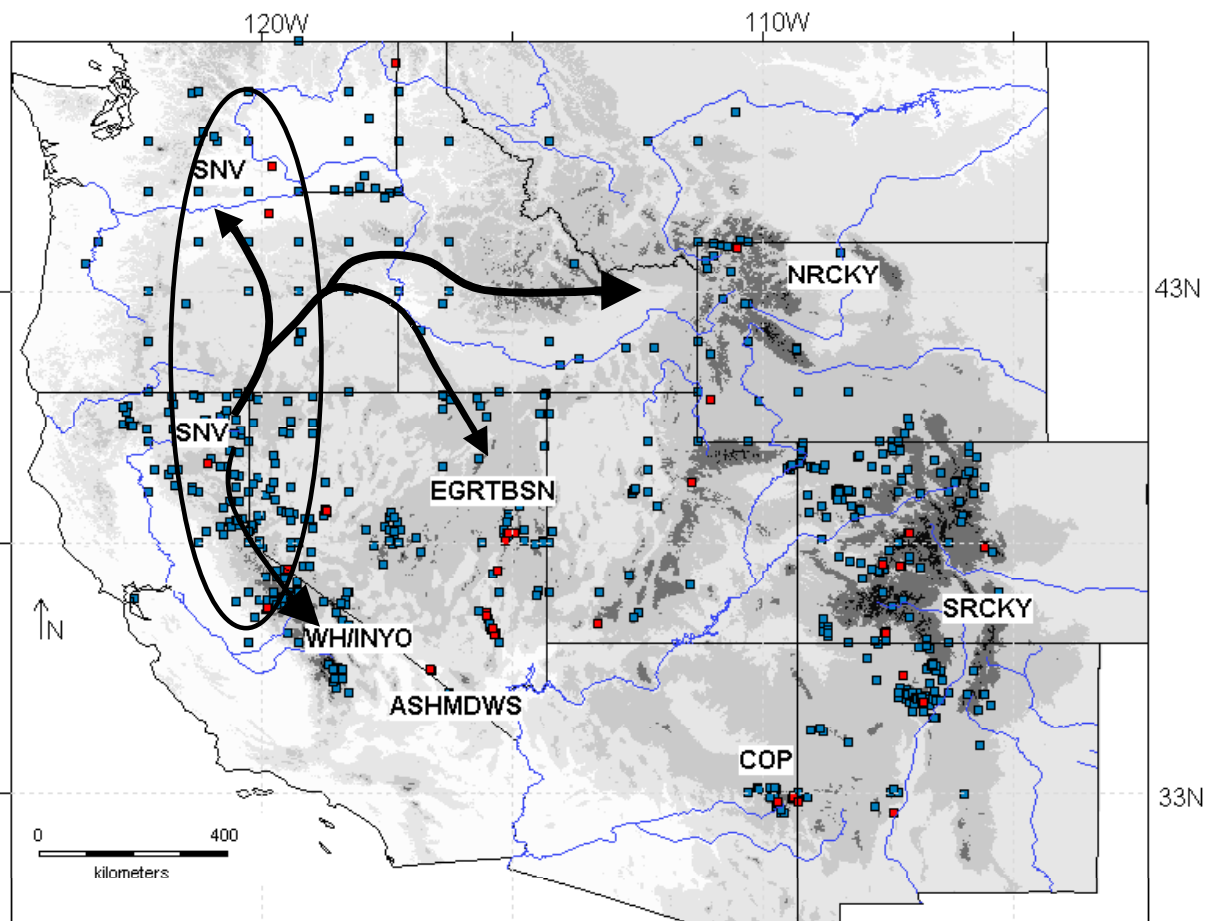
Sierra Nevada clade with adjacent habitat in the northern Rocky Mountains and the eastern Great Basin as suitable habitat became available near the end of the Pleistocene. The preservation of this connection is indicated by the cyt*b* topology: some samples from the eastern Great Basin are nested within the southern Sierra Nevada clade (Fig. 3). The climatic warming that began at the end of the Pleistocene and continued into the late Holocene may have severed the connection between the southern Sierra Nevada and these adjacent habitats, but this conclusion awaits further sampling.

An aspect of Pleistocene refugia often overlooked in studies are their duality: refugia as post-Pleistocene isolates, and those that acted as source populations. This study demonstrated support for the presence of both types of refugia within the montane vole. The relatively long branch lengths and Kimura 2P genetic distances for the Colorado Plateau and White/Inyo clades indicated the dominant influence of isolation. The average distances for the plateau population (Table 2.1b, D = 3.9) and the White/Inyo (D = 3.3) approached that for other recognized species of *Microtus* (Conroy and Cook, 1999), but this average distance was relatively greater than that demonstrated for the putative source population in the southern Sierra Nevada (D = 2.7).

Results of this study demonstrated the potentially significant insights into the colonization history possible due to complementary analyses of different molecular markers. In particular, the expected differences in effective population size and mutation rate of mitochondrial and nuclear DNA allow inferences at more recent and deeper time scales, respectively. This approach suggested great utility for additional studies of Great Basin taxa to reconstruct the origins of diversity in this region.

80

ACKNOWLEDGEMENTS

Chapter III

Elucidating patterns of connectivity between populations of the montane

vole (*Microtus montanus*) in the Great Basin

Peripheral populations are expected to demonstrate reduced genetic diversity, but this expectation has not been consistently observed in biogeographic studies. Genetic data provides a way to examine core/periphery dynamics through assessments of gene flow and genetic diversity. The distribution of the montane vole, *Microtus montanus* in the Great Basin includes several populations along the White River that are separated by large areas of unsuitable habitat. I examined the White River populations as peripheral through analyses of the cyt*b* gene and 6 microsatellite loci from 5 populations. My analyses indicate a general north to south decrease in genetic diversity even while controlling for uneven sampling and missing data. Northern populations also demonstrated large effective population size compared to southern populations. The proximity of the northern populations to water features, combined with the riparian habitat affinity of the species suggested changes in hydrology were influential to the maintenance of population connectivity. The late Pleistocene desiccation of mesic environments is well documented. The retreat of glacial inputs into the White River may have favored a shift from regional to localized gene flow. The localized mesic conditions were maintained by springs. In suggesting a greater role for vicariance in the montane vole, my study offers insight to how riparian ecosystems shaped the distribution of mammal taxa in the Great Basin, and an avenue for future research in this vital region.

INTRODUCTION

The distribution of many small mammals is not continuous; rather, the distribution may be patchy and due to factors such as glacial dynamics (Hawkins and Porter, 2003), specialized niche requirements, or to anthropogenically induced climate and/or habitat change (Epps et al., 2004). Persistence probability is positively related to habitat patch area and negatively related to the level of isolation, with persistence more likely on larger more connected patches (Hanski, 1997; MacArthur and Wilson, 1967). Recent empirical research also supports the relationship between spatial distribution and extinction probability (Lambin et al., 2005; Schwartz et al., 2003).

Presently, there is a lack of consensus regarding the stability of core versus peripheral populations. This controversy revolves around the level of connectedness between species populations and their size. Many peripheral populations demonstrate lower population density and density is expected to be related to extinction risk (Lomolino and Channel, 1995). Therefore, extinctions are more likely at the edges of a species range, i.e. in the peripheral populations (Lomolino and Channel, 1995). On the other hand, peripheral populations may harbor greater genetic diversity, which may decrease extinction risk (Lomolino and Channel, 1995). As a result of the disparities, core populations are assumed to be large, contiguous and either resilient to extinction (Brown, 1984) or extinction prone (Channel and Lomolino, 2000; Lomolino and Channel, 1995; Fisher, 1930). Conversely, peripheral populations are predicted to be small, isolated and prone to extinction (Brown, 1984), or resilient and a source for future evolutionary change (Lomolino and Channel, 1995; Mayr, 1963).

The relationship between distribution and extinction risk is conflated by how core

and peripheral populations are quantified. Traditionally, peripheral populations are defined as those that lie outside the broader, more contiguous range of species (Hall, 1981). Peripheral populations are also quantified as the area outside the centroid of the range as calculated by latitude and longitude (Sagarin et al., 2006). Still other studies define peripheral populations as those separated from other populations by inhospitable geography (Lomolino and Channell, 1995; Brown, 1971). Even this approach is confounded by how the area of suitable habitat is defined (Lawlor, 1998; Grayson and Livingston, 1993).

The application of genetic theory may provide an additional independent solution to the core-periphery debate because it allows the identification of core and peripheral populations through analyses of gene flow. More importantly, the approach offers a framework to test hypotheses concerning genetic diversity as a function of core and peripheral populations. Unlike approaches that rely on satellite telemetry or other passive transmitters, which can be cost-prohibitive or impracticable over large geographic scales, the use of genetic data has proved to be a valuable tool. The use of random, variable repeats found in microsatellite markers are especially useful in detecting gene flow because they tend to be highly polymorphic, selectively neutral and distributed throughout the genome (Goldstein and Pollock, 1997).

The presence of gene flow is suggested by a positive and montonic relationship between geographic and genetic distance (Hutchison and Templeton, 1999). The geographical distance between populations is related to the genetic distance, as measured by pairwise F_{ST} (Hutchinson and Templeton, 1999). According to the stepping stone model, a positive and monotonic relationship between geographical distance and genetic

distance are expected when gene flow and genetic drift are in equilibrium (Fig. 3.1).

The detection of gene flow necessitates the use of molecular markers with a sufficiently rapid mutation rate that can capture variation at shallow times scales; as would be expected when animals disperse to neighboring populations and exchange genes through mating. Several studies have demonstrated the utility of microsatellite markers to the examination of gene flow (reviewed by Selkoe and Tonen, 2006). Microsatellite loci in the nuclear genome are rapidly evolving tandem repeats that may vary between individuals. Tracking this variability can distinguish populations connected by gene flow.

The high topographic and hydrologic variation of the Great Basin leads to a heterogenous landscape of habitat, which makes it an ideal system to examine extinction and colonization dynamics. In a pivotal study, Brown (1971) suggested that the distribution of extant mammals in the Great basin was the primary result of extinction processes, and not immigration. This conclusion was upheld even with additional data (Lawlor, 1998; Grayson and Livingston, 1993), but was based on the assumption that populations, in fact, were isolated. A recent Great Basin study that detected gene flow between mountain populations of the marmot contradicted this earlier work (Floyd et al., 2005). Clearly, additional studies that examine connectivity in the Great Basin can examine the generality of the assumption that the distribution of Great basin mammals is a consequence of extinction. Like mountaintop habitats, riparian meadows in the Great Basin also appear spatially disjunct. The same non-equilibrium dynamics suggested for mountaintop taxa (Brown, 1971) may also be true for riparian obligates at moderate elevations.

Figure 3.1. Expected relationships between FST and geographical distance. 4 dispersal hypotheses are illustrated (Hutchinson and Templeton, 1999): (a) an equilibrium between gene flow and genetic drift at all spatial scales; (b) dominant influence of recent and local colonization at all spatial scales; (c) the dominant influence if genetic drift across all spatial scales through vicariant processes; and (d) the combined influence of gene flow at smaller spatial scales and genetic drift at larger spatial scales. The relative amount of variance around the estimate of FST is indicated with gray shading.
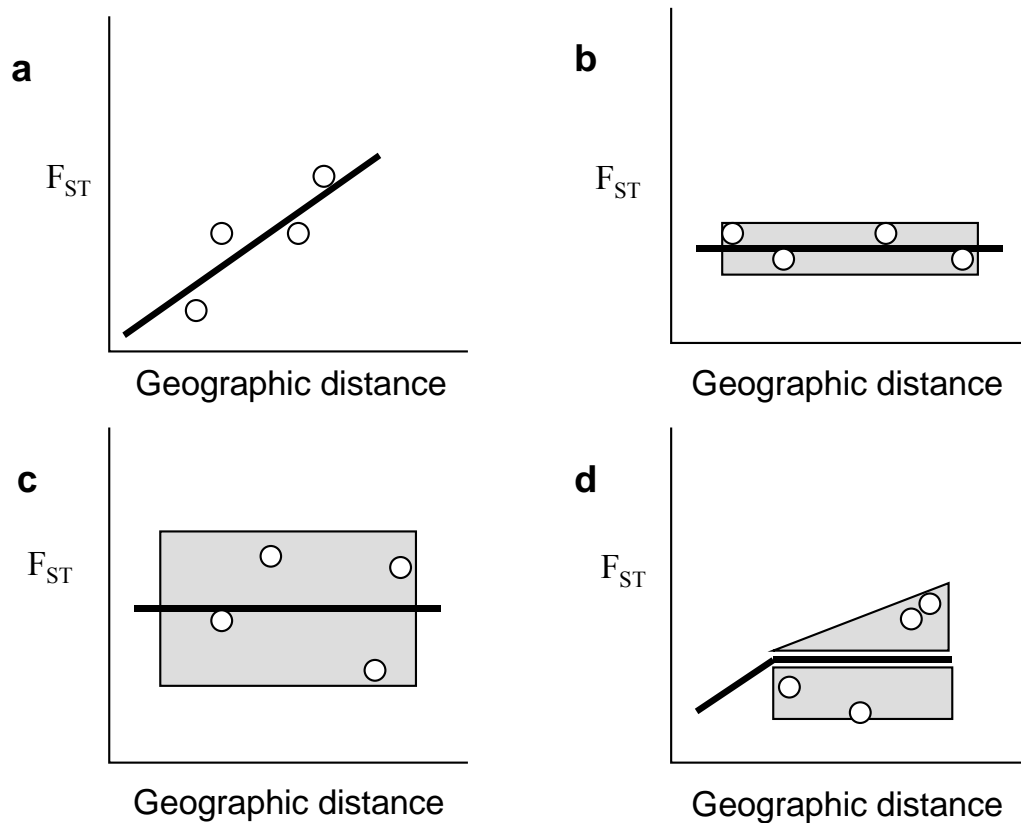
The distribution of *Microtus montanus* in the western United States is characterized by a number of large, more continuous populations at higher latitudes, and several presumed isolated populations along the southern edge of the range (Hall, 1981). Within the Great Basin, the montane vole occupies mesic grasslands between approximately 100-4000m in elevation. Historical records suggest peripheral populations at Ash Meadows, in the Amargosa Valley of southwestern Nevada, and in Pahranagat Valley in the southeast (Hall, 1981). My surveys demonstrated several additional montane vole populations, from south to north: Key Pittman Wildlife Management Area, Kirch Wildlife Management Area and at Steptoe Wildlife Management Area and surrounding streams in Steptoe Valley. Like beads on a string, the vole populations are arranged linearly along the White River in eastern Great Basin, from Pahranagat in the south, to Steptoe in the north (Fig. 3.2), and were separated by 20-126 miles of ephemeral riverbed.

Studies of montane vole habitat associations and field data suggested that the montane vole is a habitat specialist. Unlike many congeners, the species occupies riparian areas, and is usually found within proximity to running water (Findley and Jones, 1962; Anderson, 1959). The White River is classified as an ephemeral drainage and as such, the vole populations found here may be peripheral isolates. In this study, I examined the core and peripheral dynamics of the montane vole through an analysis of molecular variation. If the White River populations are peripheral, they are expected to exhibit: reduced intra-population diversity, bottlenecks and/or increased differentiation among populations and

Figure 3.2 Locations of *Microtus montanus* in the Great Basin and the southern Rocky Mountains. Legend: Steptoe Wildlife Management Area and Steptoe Valley Stptoe); 2, Kirch Wildlife Management Area (Kirch); 3, Key Pittman Wildlife Management Area (KeyPttmn); 4, Pahranagat National Wildlife Refuge and Crystal Springs (Pahranagat); and 5, Stillwater National Wildlife Refuge (Stllwtr). The extinct population from Amargosa Valley is also indicated (AshMdws). The samples examined from the southern Rocky Mountains (SRcky), the northern Rocky Mountains (NRcky), the Colorado Plateau (COP), the White/Inyo Mountains (Wht/Inyo) and the northern and southern Sierra Nevada Mountains (N SNev, S Nevada) are also shown. Map projection is Geographic World Geodetic System 1984.
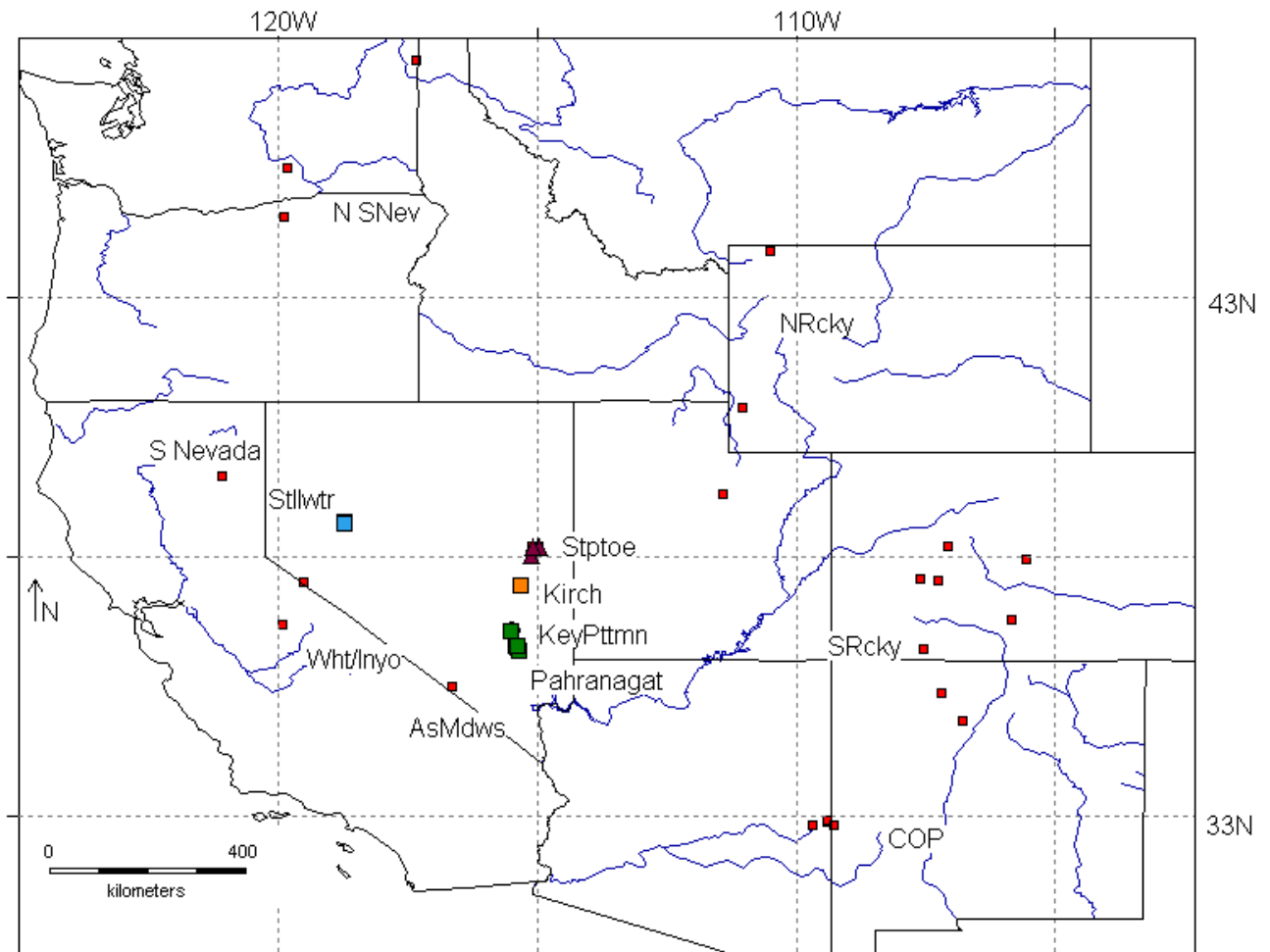
an absence of inter-populations gene flow (Eckstein et al., 2006; Hutchinson, 2002; Hutchison and Templeton, 1999; Lesica and Allendorf, 1995).

MATERIALS AND METHODS

Sampling

I tentatively identified core and peripheral populations a *priori* to organize field surveys (Table 3.1). A recent phylogeographic study indicated the Sierra Nevada as a potential core population, which provided colonists to the northern Rocky Mountains (Crawford and Smith, in prep). Additionally, the riparian habitat in the northern Great Basin is more contiguous (Polhemus and Polhemus, 2000). I therefore selected Steptoe Valley and Stillwater in the northern Great basin as potential core populations. A linear distance of approximately 202 meters separated Steptoe Valley from the next nearest population at Kirch along the White River. I tentatively ascribed a peripheral status to the White River population. The peripheral populations included, from north to south: Kirch Wildlife Management Area, Key Pittman Wildlife Management Area and Pahranagat Valley (Fig. 3.3). I collected 43 tissue samples for genetic analyses through field surveys in the Great Basin.

For comparison, I included montane vole samples from other areas within the Great Basin. Samples from Ash Meadows in the Amargosa Valley of southwestern Nevada were included. My surveys, which totaled over 3800 trap nights, demonstrated the likely extirpation of this population. I obtained a loan of 6 samples from Ash Meadows from the Museum of Vertebrate Zoology (Figs. 3.2, 3.3). To compare the microsatellite variation for montane vole populations in the Great Basin to other regions,

Table 3.1. Summary of data collected from field surveys of Great Basin montane vole populations from 2006-2009. For each location (site), the dates of surveys, the number of trap nights (the number of times traps at each site were available for capture), the number of montane vole captures (#*Microtus*) and the assignment into either core (C) or peripheral (P) status.

Site	Dates	# trap nights	#*Microtus*	Pop. ID
Ash Meadows	11/2006	2000	0	P
	3/2007	2000	0	
	6/2007	1800	0	
	5/2008	5500	0	
	9/2008	4500	0	
	5/2009	1000	0	
Kirch	3/2007	2000	3	P
	5/2007	2000	5	
	8/2008	1000	5	
Steptoe	5/2007	1000	9	C
	3/2008	1000	5	
	8/2008	2000	15	
Stillwater	9/2007	1000	6	C
	10/2007	1000	13	
Pahranagat	9/2006	1000	6	P
	10/2006	1000	2	
	5/2007	1000	3	
	6/2007	1000	0	
	9/2007	1000	2	
	8/2008	1000	0	
	10/2008	1000	5	
Key Pittman	8/2006	1000	2	P
	5/2007	1000	0	
	8/2008	1000	11	
	10/2008	2000	0	

Figure 3.3 The montane vole populations along the White River in the eastern Great Basin. Legend: 1, Steptoe Wildlife Management Area and Steptoe Valley (Stptoe); 2, Kirch Wildlife Management Area (Kirch); 3, Key Pittman Wildlife Management Area (KyPtmn); and 4, Pahranagat National Wildlife Refuge and Crystal Springs (Pahngt). The extinct population from Amargosa Valley is also indicated (AshMdws), which was not included in the microsatellite analyses. The location of the White River drainage is shown (arrow). Map projection is Geographic World Geodetic System 1984.
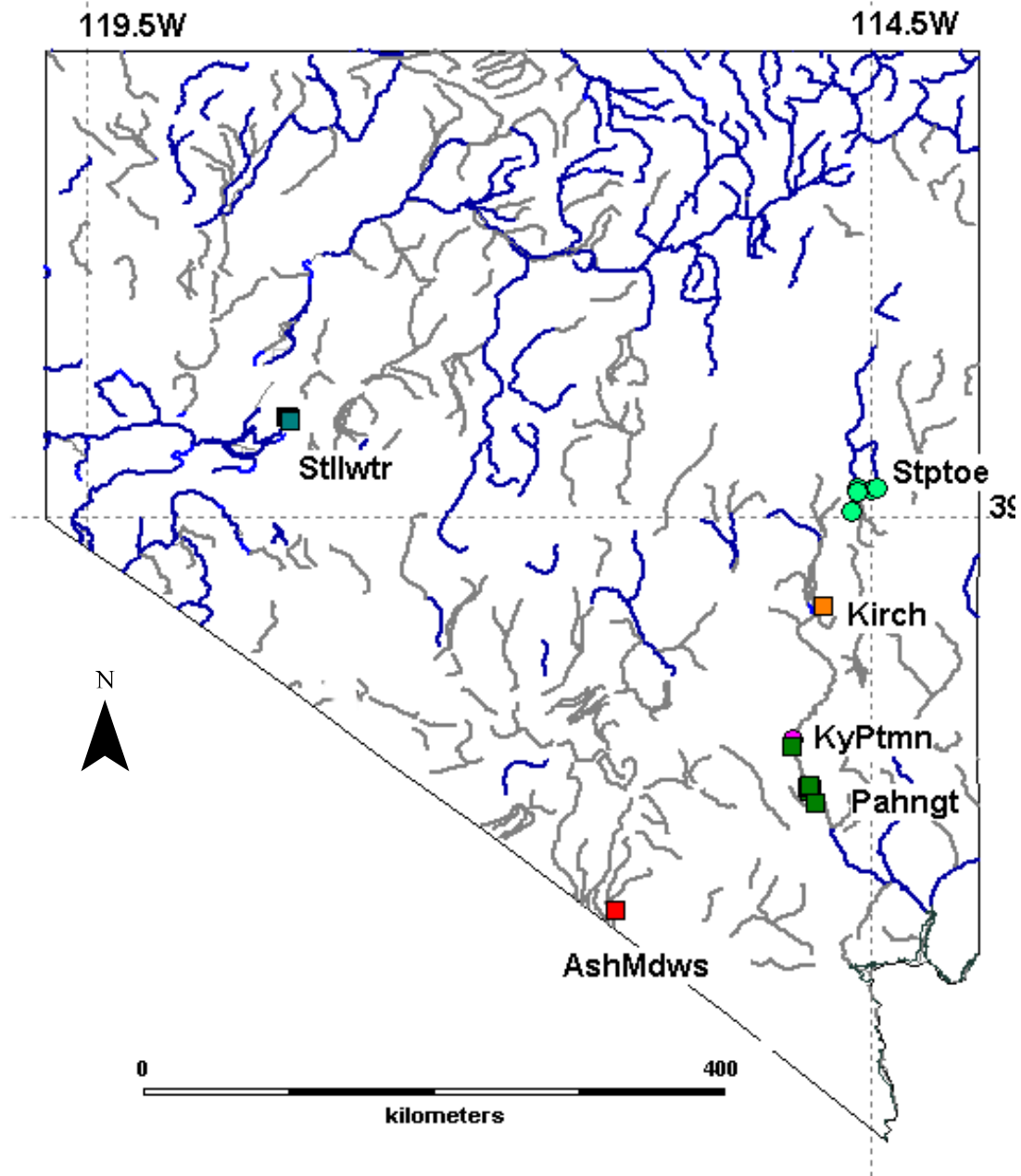
I obtained a loan of tissue samples from the southern Rocky Mountains and the Colorado Plateau from the University of New Mexico's Museum of Southwestern Biology (MSB).

Sample preparation

I analyzed the molecular variation from the mitochondrial DNA (mtDNA) cyt*b* gene from 43 fresh and 14 museum samples (Figs. 3.2, 3.3). DNA was extracted from tissues using a modified salt extraction procedure (Medrano et al., 1990). I amplified and sequenced the cyt*b* gene, which has a mutation rate suited to resolving phylogenetic divergence prior to 2 Ma (Galewski et al., 2006). The amplification used the polymerase chain reaction (PCR), and primer pairs MVZ05-MICRO-O6 and ARVIC07-VOLE14 (Hadly et al., 2004). Reaction conditions for PCR were $95^{o}C$ (30 s), denaturing, $50^{o}C$ (25 s) annealing, and $72^{o}C$ (1 min) extension. PCR products were purified using a QIaquick PCR Purification Kit (Qiagen, Valencia, California). DNA from 6 museum skin samples collected in 1933 from Ash Meadows was extracted by first soaking the samples in 10% 10mM TRIS buffer at $45^{o}C$ for 1 hour, and 5uL of DNA were used per PCR reaction instead of 1uL.

I sequenced the cyt*b* using dye-labeled terminators and cycle sequencing conditions of $96^{o}C$ (10 s) denaturing, $50^{o}C$ (5s) annealing, and $60^{o}C$ (4 min) extension. PCR products were purified using a sodium acetate and ethanol precipitation procedure and read with an ABI 3100 Automated Sequence Analyzer (ABI Applied Biosystems). Sequences were manually aligned using Sequencher v.4.7 (Bromberg, 1995).

Microsatellite variation for between 8-26 animals for 4 presumed isolated populations and 2 core populations was used to characterize gene flow (Table 3.1). Microsatellite amplification from skin samples can be complicated by the failure of

alleles to reliably amplify due to low concentrations or quality of template DNA. Because of complexities associated with allelic dropout (Sefc et al., 2003), I did not attempt to amplify microsatellite loci from museum skin samples collected from Ash Meadows. Consequently, those samples were excluded from microsatellite analyses, leaving a total of 49 samples (Figs. 3.2, 3.3).

The genetic distinctiveness of montane vole populations along the White River was examined through an analysis of 6 nuclear DNA microsatellite loci. I examined nuclear DNA in three separate reaction sets: MSMM2, MSMM 6, MSMM 8 (Ishabashi et al., 1999), MAR76, MAR80, MAR113 (Walser and Heckel, 2008). These microsatellite loci are reported from analyses of *Microtu*s, and researchers suggest the latter 3 loci may be polymorphic for *M. montanus* (Walser and Heckel, 2008). PCR amplification was completed using Qiagen Multiplex Kit. The PCR conditions consisted of a denaturation step of $94\,^{o}C$ (5 min), 30 cycles of $94^{o}C$ (30 sec), $55\,^{o}C$ (1 min) annealing, and a $72\,^{o}C$ primer extension (1 min). A final step of $72^{o}C$ (10 min) was used to complete primer extension. Microsatellite fragments were separated using ABI 3100 sequence Analyzer, and raw data was checked and scored with GeneMapper v.3.7 (ABI Prism, Applied Biosystems).

Patterns of genetic diversity

The rate of mutation in the cyt*b* may detect the incipient stages of population subdivision due to isolation. I conducted a phylogeographic analysis to examine the possibility. I rooted the phylogenetic topology with *M. pennsylvanicus*, which is the suggested sister taxon to *M. montanus* (Conroy and Cook, 2000). The model of evolution that best fit the data was selected with MODELTEST (Posada and Crandall, 1998).   I

93

estimated the genetic relationships between samples using a Bayesian analysis in MRBAYES v.3.1.2 (Huelsenbeck and Ronquist, 2001) using the selected model parameters.

Peripheral populations may demonstrate reduced levels of genetic diversity and increased levels of intra-population differentiation (Eckstein et al., 2006; Hutchinson, 2002; Frankham, 1998). I estimated diversity within the cyt*b* with the number of haplotypes, haplotype diversity (Hd) and the number of mutations (*S*) in ARLEQUIN v. 3.11 (Excoffier et al., 2006). Intra-population differentiation was estimated using the Kimura 2-parameter pairwise distances in ARLEQUIN v. 3.11(Excoffier et al., 2006). I set parameters to allow a maximum of 2% missing data per locus to mitigate for some incomplete sequences. I estimated microsatellite parameters in FSTAT v.2.9.3 (Goudet, 2001). The parameters include allelic richness (AR), gene diversity, number of alleles (A), and the inbreeding coefficient F_{IS} of Weir and Cockerham (1984) per population. The unbiased estimate of expected heterozygity (H_{UE}) is thought to be a more precise estimator of heterozygosity when sample sizes are small (Pruett and Winkler, 2008; Nei, 1987). I estimated H_{UE} using the microsatellite toolkit. I also tested for Hardy-Weinberg equilibrium per population in FSTAT v.2.9.3.

The distribution of alleles within and between populations can correlate the partitioning of genetic diversity with population relationships, and may be a better proxy for the evolutionary history of a population than heterozygosity (Leberg, 2002). I conducted a second analysis of allelic diversity using rarefaction due to uneven sampling across populations. Allelic diversity is summarized in two genetic parameters: allelic richness (A_R) and the number of private alleles, or private allelic richness (P_{AR}). A_R is the

94

number of distinct alleles expected in a randomly drawn population (Szpiech et al., 2008) and P_{AR} is the number of distinct and unique alleles expected from a population when the sample is selected randomly from a collection of populations (Szpiech et al., 2008). The estimation of A_R and P_{AR} is influenced by sample size, and both metrics may exhibit erroneous values when sample sizes vary across populations. Rarefaction controls the effect of sample size in the estimation of both genetic parameters (Kalinowski et al., 2004; Petit et al., 1998; Nei, 1987).

The parameters of A_R and P_{AR} were estimated from microsatellite data using rarefaction in ADZE v.1.0 (Szpiech et al. 2008). I set the maximum sample size to match the smallest sample size across all populations (Key Pittman, N = 8). Because of missing data, I performed several iterations while varying the acceptable percentage of missing data to assess model performance. A final analysis that constrained the maximum percentage of missing data minimized the variance and standard error for A_R and P_{AR}.

I examined the relationship between population size and genetic diversity with the equivalency, $Ne= 2\Theta\mu$, where $\underline{\mu}= \mu\kappa$ is the mutation rate and κ is the generation time. I estimated the diversity parameter theta (Θ) in MIGRATE v.3.0 (Beerli and Felsenstein, 2001). The widely accepted mutation rate for the cyt*b* gene of 0.02 substitutions per site per million years with a range of 0.021-0.029 (Brown et al., 1979) and a published generation time for *Microtus* (Keller, 1985) were also used.

If genetic diversity is a function of population age and isolation, then older populations should demonstrate reduced genetic diversity compared to younger populations. To assess this possibility, I estimated the cyt*b* divergence for each montane vole clade in BEAST v.1.4.6 (Drummond et al., 2006). Following, I estimated the relative

age of montane vole clades by setting the divergence between *Microtus* and *Myodes* using a lognormally distributed prior of 1.3 ± 0.2 Ma based on fossil evidence (Chaline and Graf, 1988; Catzeflis et al., 1989), and an internal lognormal prior representative of divergence within *Microtus* at 0.5 ± 0.2 Ma (Repenning, 1990; Conroy and Cook, 1999; Steppan et al., 2004).

Tests of Neutrality

Reduced genetic diversity may be due to population age or to violations of neutrality such as genetic bottlenecks (Hoffman and Blouin, 2004). To differentiate among these hypotheses, I conducted a mismatch analysis on the cyt*b* data in FSTAT v.2.9.3 (Goudet, 2008). In a population that has undergone a recent expansion or genetic bottleneck, the mismatch distribution is expected to be unimodal and approximate a Poisson curve as haplotypes spread relatively rapidly across the landscape. A population in mutation-drift equilibrium is predicted to exhibit a multi-modal distribution because multiple haplotypes are maintained in stable populations (Harpending, 1994).

The equilibrium between migration and drift may be violated in expanding populations, or populations in decline. I examined equilibrium dynamics through tests of neutrality with Fu's *Fs* (Fu, 1997) and Tajima's *D* (Tajima, 1989) statistics. The *Fs* statistic tests the hypothesis that the number of recent or rare haplotypes is greater than that expected given neutrality. A negative value of *Fs* may be indicative of population expansion and a positive *Fs* may suggest a population bottleneck. The Tajima's *D* is a more conservative test of bottlenecks than Fu's *Fs* (Crawford, 2007). The Tajima's *D* test compares the expected values of segregating sites, *S* and the standardized number of mutations (as represented by the number of pairwise differences), π. At equilibrium, $S = \pi$.

96

Changes in population size violate equilibrium. A positive value for Tajima's *D* indicates evidence of a genetic bottleneck, and a negative value supports population expansion.

I examined additional support for neutrality using an additional bottleneck test. Microsatellite data was examined for population reduction with the *M*-ratio (Garza and Williamson, 2001), where *M* = number of alleles/ range in allele size. A bottleneck is thought to reduce the number of alleles at a faster rate compared to any change in the range of allele sizes. Therefore, a contracting population should exhibit a smaller *M*-ratio compared to a stale population.

Genetic Isolation

Gene flow between populations is expected to yield a linear relationship between geographic and genetic distance (Fig. 1a). When linearity is violated, three additional scenarios are possible. In the first case, the relationship between geographical and genetic distance is dominated by the influence of recent and local gene flow (Fig. 1b): the genetic signature of all populations is similar and levels of F_{ST} are reduced. As time progresses, a vicariant event isolates the populations, increasing the influence of genetic drift, which also increases the variance around F_{ST} (Fig. 1c). A third case depicts the influence of gene flow at local scales and drift at larger spatial scales, which results in a non-linear relationship between F_{ST} and geographic distance and a higher variance about F_{ST} (Fig. 1d).

The correlation between geographic and genetic distance was examined using Mantel test in ARLEQUIN v. 3.11 (Excoffier et al., 2006). Genetic distances were obtained as Kimura 2-parameter distances. I calculated geographic distances using the great circle distance, the shortest distance between any two points on a sphere. Given the

montane vole's propensity for riverine habitats (Findley and Jones, 1962), I calculated the average distance between populations using a straight-line distance along the river. The Stillwater population is outside of the White River drainage. Thus, the geographical distance was measured as the shortest west to east distance to the White River. Spatial analysis functions in Hawth's Tools for ArcGIS were used (Beyer, 2004) to calculate the distance between each point (Km).

Small microsatellite sample sizes can lead to erroneous conclusions about patterns of gene flow (Pruett and Winkler, 2008; Selkoe and Toonen, 2006; Leberg, 2002). Thus, I examined the equilibrium between gene flow and random genetic drift using a Mantel test (Mantel, 1987) on the microsatellite data. Forward simulations of microsatellite data in EASYPOP v.2.0.1 (Balloux, 2001) were conducted to examine the influence of small sample sizes.

Simulation parameters were set using the same number of males and females for each population and a 1-dimensional stepping stone model of migration. This model is appropriate for systems spatially distributed linearly, as along a river (Balloux, 2001). The number of loci was set to 6 and allowed free recombination between loci. I specified the mutation model to allow an equal probability of mutation to the same allelic state and selected an average value for the number of allelic states (n = 500). I ran two sets of simulations. I maintained the samples size as in the empirical data (e.g., Steptoe, n = 12, Kirch, n = 11, Key Pittman, n = 8, Pahranagat n = 12 and Stillwater, n = 4) and varied the number of generations and replications. For the second set, I varied sample size (e.g., n=12, n =15, n =20, n =50, n= 100) as well as the number of generations and replications. Pairwise Fst were estimated from simulation output in GENEPOP v.4.0 (Raymond and

Rousset, 1995). Pairwise Fst values between populations were used in a regression of genetic distance (Fst) on geographic distance in SPSS v. 16.0.1.

Recent gene flow maintains genetic connectivity between populations. The stepping stone model of gene flow proposes that populations in proximity are more likely to be exchange genes, and therefore be more genetically similar. I estimated the spatial extent of genetic similarity using an assignment test in STRUCTURE v. 2.2 (Falush et al., 2003). I used a Bayesian algorithm to calculate the likelihood of the data given the probabilities for the different populations as K (1-6), where K is the number of subpopulations. Genetic clusters are assigned, based on likelihood values, as those that are most similar to one another. I selected the burn-in period by assessing where the likelihood values begin to plateau, i.e. the place in the search algorithm where significant gains in the likelihood of the data given K (Evanno et al., 2005). I set the search parameters on the algorithm to a burn-in of 1,000,000 generations with post-burn-in MCMC replications set to 100,000.

RESULTS

Patterns of genetic diversity for cytb gene

The phylogenetic topology indicated strong support for three main clades within Great Basin populations: a northern clade that consisted of the Steptoe and Kirch samples, a southern clade made up of the Key Pittman and Pahranagat, and a third clade that contained samples from the southern Sierra Nevada (Fig. 3.4). The model of evolution that was the best fit to the cytb data was the HKY +I+ G model (-ln$L$=2112. 3035) with 6 rate parameters. Parameters estimated for this model included empirical

Figure 3.4. Bayesian phylogram of cyt*b* variation for populations of the montane vole. Legend: Steptoe Valley and Kirch (Stpoe/Kirch); Pahanagat Valley, the northern Sierra Nevada Mountains and Key Pittman (Pahrngt/KyPtmn/N SNev); Ash Meadows (Ash Mdws); the Colorado Plateau and southern Rocky Mountains (COP/SRcky); and the White, Inyo and southern Sierra Nevada Mountains (Stllwtr/Wht/Inyo).
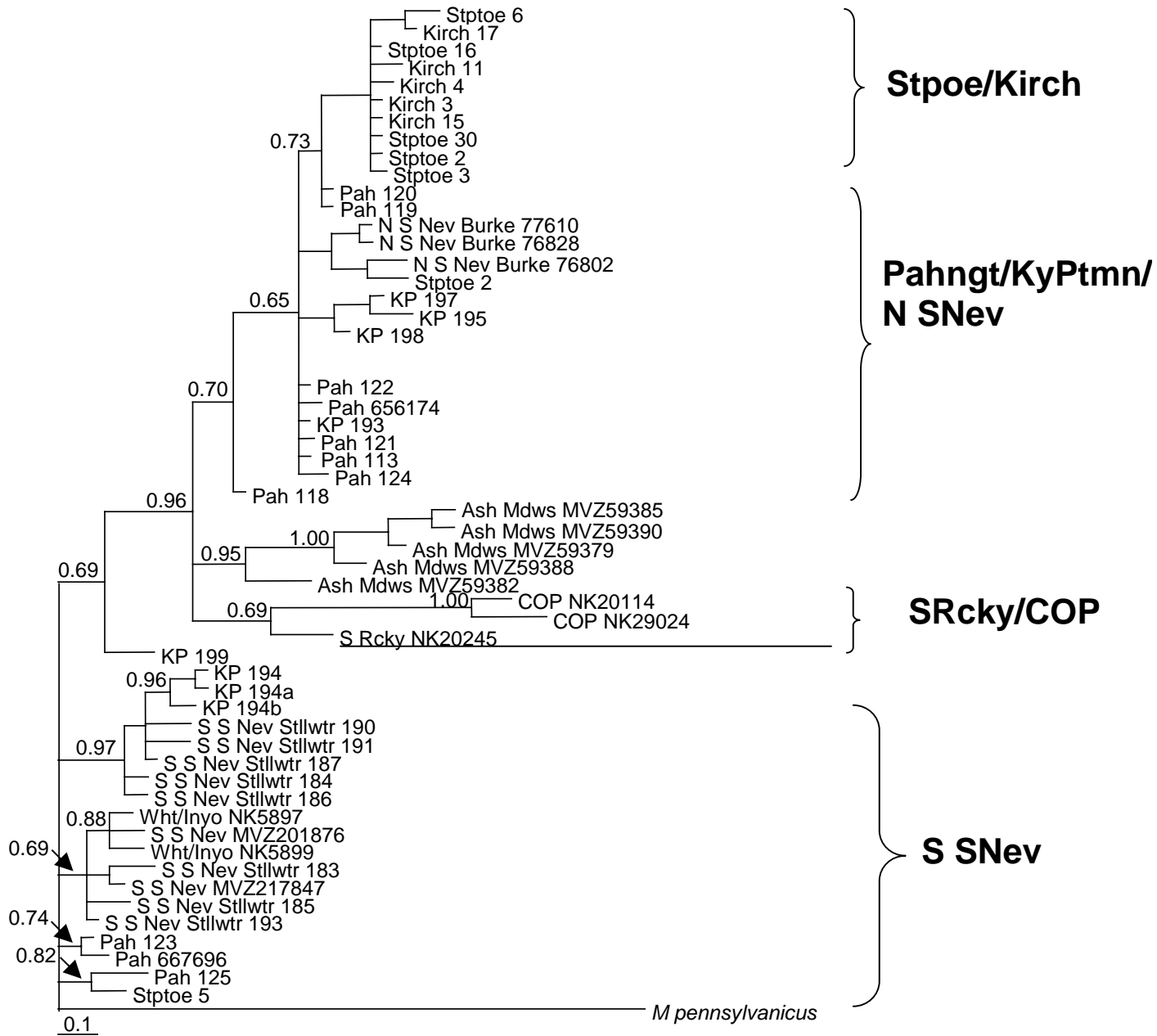
base frequencies (A=0.31, C= 0.28, G= 0.12, and T= 0.29) and a transition/transversion ratio of 1.57. The values for the proportion of invariable sites and the gamma shape parameter were 0.51 and 0.66, respectively.

Overall, my analyses of mitochondrial DNA indicated relatively higher genetic diversity for the northern populations. The northern population of Steptoe exhibited the highest values for genetic diversity parameters and Kirch exhibited the lowest values for parameters such as gene diversity (GD), haplotype diversity (Hd), number of segregating sites (S) and the number of polymorphic loci (P) (Table 3.2).

Genetic distances between populations were similar. The Kimura 2-parameter distance indicated slightly greater intra-clade F_{ST} for the southern clade of Key Pittman, but overall Fst were similar for all White River populations (Table 3.2). Estimates of effective population size demonstrated the highest range of N_e for Pahranagat (θ = 0.0296, N_e = [21,142-14,581]) and the lowest for Kirch (θ = 0.0024, Ne = [1182-1632]).

A reduction in genetic diversity may be due to long-term divergence *in situ*, the stochastic influence of genetic drift or recent population expansion. My analyses indicated that the oldest populations were also the most genetically diverse. The two northern clades of Steptoe and Stillwater demonstrated the earliest divergence times, with the divergence for Stillwater preceding all other clades. The populations in the southern White River valley exhibited more recent late Pleistocene divergence estimates (Table 3.3).

Table 3.2. Genetic diversity parameters for cyt*b* in Great Basin populations of the montane vole. Legend: (N), sample size; (S), number of mutations; (h), number of haplotypes (h); (Hd), haplotype diversity and Kimura 2-parameter intra-clade distance (D).

Population	N	S	h	Hd (± S.D.)	D
Steptoe	7	37	9	1.00 ± 0.09	2.72
Kirch	5	8	6	0.90 ± 0.03	1.02
KeyPittman	8	7	8	0.96 ± 0.01	16.96
Pahranagat	11	11	10	0.85 ± 0.01	2.0
AshMdws	5	2	5	0.80 ± 0.03	37.10
Stillwater	8	19	8	1.00 ± 0.07	1.07

Table 3.3. Divergence estimates for primary clades within the montane vole. Legend: Mean divergence (T_{MRCA}), effective sample size (ESS) and range of divergence estimate (highest posterior density). An ESS value of ≥ 100 is considered appropriate for robust parameter estimation (Drummond et al., 2006).

Clade	Mean divergence (Ma)	ESS	Range (Ma)
Eastern Great Basin	0.43	125.21	0.15-0.81
Steptoe	0.30	613.17	0.15-0.50
Kirch	0.17	450.45	0.090-0.255
Key Pittman	0.11	364.23	0.06-0.165
Pahranagat	0.14	250.72	0.08-0.22
Stillwater	1.16	460.32	0.80-1.5
AshMdws	0.15	240.32	0.07-0.222
AZ/NM	0.19	243.32	0.09-0.31

Patterns of diversity for microsateliite markers

Patterns of variation in microsatellite loci demonstrated relatively higher polymorphism for the northern populations. My analyses demonstrated that all microsatellite loci are polymorphic and exhibited a range of FIS values of 0.011 to 0.875 and a range for Θ of 8.1to 2.53 (Table 3.4). The total number of alleles per locus ranged from 5-13. The observed heterozygosity (HO) over all loci ranges from 0.097 to 0.721 and the mean HO is 0.456. The locality at Steptoe exhibited the highest mean values for allelic richness, AR (2.6), number of alleles, A (5.3), gene diversity (GD = 0.75), and among the highest for expected heterozygosity (HEXP = 0.63). My estimates of unbiased expected heterozygosity (HUE) indicated that Pahranagat and Stillwater demonstrated the highest HUE and Kirch the lowest (Table 3.5).

Rarefaction analyses also demonstrated a north to south decrease in private allelic richness (PAR). Steptoe exhibited the highest PAR (0.859) and Pahranagat exhibited the lowest value (0.437, Table 3.6). My rarefaction analysis for allelic richness (AR) also indicated a north to south decrease, and Kirch exhibited the lowest value (Table 3.6).

Tests of Neutrality

I examined the relationship between genetic diversity and demographic dynamics with tests of neutrality. My mismatch analyses demonstrated non-significant raggedness indices for southern clades of Kirch, Key Pittman and Pahranagat Valley, which indicated a unimodal haplotype distribution and population expansion. The null hypothesis of demographic expansion was also supported with the Fu's *Fs* test for the southern populations. The population at Ash Meadows demonstrated a large, positive value for

Table 3.4. Summary of *F*-statistics (Weir and Cockerham, 1984) for polymorphic microsatellite loci of the montane vole, *Microtus montanus*. Legend: the inbreeding coefficient (F_{IS}), a measure of genetic differentiation (F_{ST}), the observed proportion of heterozygotes (H_O), and the within sample gene diversity (H_S) (Nei, 1987).

Locus	F_{IS}	F_{IT}	F_{ST}	H_O	H_S
M2	0.127	0.212	0.097	0.721	0.837
M6	0.011	0.123	0.011	0.279	0.281
M8	0.220	0.418	0.253	0.459	0.655
MAR76	0.392	0.460	0.115	0.510	0.818
MAR80	0.875	0.897	0.177	0.097	0.434
MAR113	0.014	0.093	0.081	0.706	0.845
Mean	0.276	0.364	0.141	0.456	0.632

Table 3.5. Summary statistics across six microsatellite loci of the montane vole, *Microtus montanus.* Legend: sample size (N), the average number of alleles per population (A), allelic richness (A_R), and the number of polymorphic loci (P). The unbiased estimate of expected heterozygosity (H_{ue}) and the expected hetrozygosity (H_{EXP}), gene diversity (GD), the inbreeding coefficient (F_{IS}) (Weir and Cockerham, 1984), and bottleneck index, *M* (Garza and Williamson, 2001). Populations that exhibit deviation from Hardy-Weinberg equilibrium are indicated by significant values for F_{IS}.

Population	N	A	A_R	P	H_{ue}	H_{EXP}	GD	FIS	M
Steptoe	12	5.3	2.6	2	0.63 ± 0.09	0.63	0.75	-0.09 (ns)	0.43
Kirch	11	4.1	2.2	1	0.58 ± 0.10	0.592	0.59	0.33***	0.35
Key Pittman	8	4.8	2.4	5	0.60 ± 0.11	0.60	0.61	0.20 (ns)	0.39
Pahranagat	12	4.8	2.5	4	0.65 ± 0.09	0.65	0.66	0.22 (ns)	0.38
Stillwater	4	4.3	2.6	4	0.65 ± 0.14	0.65	0.58	0.08 (ns)	0.41

*** significant at P≤ 0.001, ns, not significant

Table 3.6. Results from rarefaction analysis of allelic richness (AR) and private allelic richness (PAR). The mean, variance and standard error are given for each population are shown in order from north to south; western Great Basin clade of Stillwater shown last.

Genetic Parameter	Population	Mean	Variance	Std Error
PAR	Steptoe	0.859	0.342	0.235
	Kirch	0.697	0.013	0.109
	Key Pittman	0.467	0.134	0.149
	Pahranagat	0.439	0.170	0.168
	Stillwater	0.762	0.233	0.194
AR	Steptoe	3.00	1.41	0.450
	Kirch	2.58	1.06	0.380
	Key Pittman	2.81	1.92	0.520
	Pahranagat	2.76	1.11	0.410
	Stillwater	2.62	0.63	0.300

Tajima's D, which indicated support for a genetic bottleneck (Table 3.7).

A north to south decrease in the microsatellite data was also observed, with the highest *M*-ratio for Steptoe and the lowest value for Kirch. The population at Kirch also demonstrated the highest F_{IS}, which significantly deviated from equilibrium (Table 3.5).

Genetic Isolation

I estimated pairwise F_{ST} values and plotted those against geographical distances (Ln+1 transformed) among locations for both the cyt*b* gene and microsatellite data to analyze the influence of geographic distance on microsatellite variation within *M. montanus*. The Mantel test produced a non-significant correlation between genetic and great circle geographic distance and did not support isolation by distance over the range of the study area for cyt*b* data (y= -0.183x + 1.13, r^2=0.26, *P*=0.94, Fig. 5). I also calculated the distance between populations as a function of inter-population distance along the White River. Incorporation of a riverine distance increased the amount of variation explained by the data, but was not significant (cyt*b*, y=0.187x + 1.12, r^2=0.29, *P*=0.06).

I estimated pairwise F_{ST} values and plotted those against geographical distances (Ln+1 transformed) among locations for both the cyt*b* gene and microsatellite data to analyze the influence of geographic distance on microsatellite variation within *M. montanus*. The Mantel test produced a non-significant correlation between genetic and great circle geographic distance and did not support isolation by distance over the range of the study area for microsatellite loci (y = 0.0122x + 0.0623; r^2 = 0.03, df = 5, *P* = 0.213). I also calculated the distance between populations as a function of inter-population distance along the White River. Incorporation of a riverine distance increased

Table 3.7. Tests of neutrality for the cyt*b* gene in 6 Great Basin montane vole populations. Parameters calculated include a measure of the fit of the data to a model of population expansion (raggedness index, *r*), Fu's *Fs* test (*Fs*), Tajima's *D* test statistic (*D*), and associated *P*-values for each parameter. Under a model of population expansion, *P*-values for the raggedness index are expected to be non-significant, and *Fs* and *D* to be significantly negative. A large positive value for Tajima's *D* suggests the influence of a genetic bottleneck.

Population	*r* (p)	Fu's *Fs* (p)	Tajimas's *D* (p)
Steptoe	0.06 (0.76)	-2.04 (0.09)	-0.01 (0.54)
Kirch	0.09 (0.96)	-9.21 (0.00)	-1.48 (0.14)
Key Pittman	0.05 (0.89)	-9.04 (0.00)	-0.52 (0.32)
Pahranagat	0.07 (0.63)	-5.23 (0.01)	-0.94 (0.17)
Stillwater	0.27 (0.70)	-2.15 (0.05)	+0.43 (0.74)
Ash Meadows	0.10 (0.74)	-1.74 (0.07)	+1.24 (0.96)

the amount of microsatellite variation explained by the data, but was not significant (y = 4.23x + 0.639, r^2 = 0.42, df = 5, P = 0.10).

My analyses using coalescent simulations raised the possibility that the current sample size is insufficient to provide reliable estimates of gene flow between montane vole populations. For the first set of simulations, I maintained the same population sample sizes as in the empirical data, but varied the number of generations (ngen) and replications. All analyses exhibited non-significant correlations between genetic and geographic distance, but the explained variation improved with increased replications (ngen =50; y=-1.51x + 4.76, r^2 = 0.02, P = 0.32; ngen = 500; y = 6.83x + 3.64, r^2 = 0.22, P = 0.17) and compared to the empirical data (y = 0.0122x + 0.06, r^2 = 0.03, P = 0.213). In the second set of simulations, the sample size and the number of generations and replications per simulation were varied. Explained variation improved with an increase in sample size (sn) and generation time (sn =12, ngen= 50; y=-4.95x + 4.89, r^2 = 0.05, P = 0.35, sn=50, ngen=100; y= 23.84x + 3.58, r^2 = 0.20, P = 0.19). I uncovered a significant correlation between genetic and geographic distance with generations ≥ 500 (sn=50, ngen=500, y= 0.0076x – 0.0164, r^2 = 0.39, P = 0.043). For all simulations, the riverine distance derived from the distance between populations along the White River appeared to explain more variation in the data compared to the great circle distance.

I also examined the genetic similarity of montane vole populations using an assignment test. I calculated population structure using the relationship lnP (D), where lnP (D) is the likelihood of every *K* (likelihood for replication 1, -794.9 and likelihood for replication 2, -798.1) identified 5 clusters as the most likely structure for the data. The

identified clusters included a Steptoe cluster comprised of Steptoe Valley; a Kirch

cluster; a Pahranagat cluster composed of samples from Pahranagat Valley and Crystal

Springs; a Key Pittman cluster; and a Stillwater cluster comprised of all Stillwater

samples (Table 3.8).

Table 3.8. Results of assignment tests. Separate cluster assignments are indicated for all populations. The simulation parameters include k=1-6, a burn-in period of 1,000,000 generations and 50,000 MCMC replications. Bold indicates significant likelihood values.

Population	Cluster1	Cluster2	Cluster3	Cluster4	Cluster5
Steptoe	0.145	0.017	**0.646**	0 017	0.176
Kirch	**0.355**	0.159	0.221	0.012	0.252
Key Pittman	0.033	0.137	0.014	**0.687**	0.129
Pahranagat	0.069	**0.722**	0.085	0.090	0.033
Stillwater	0.029	0.047	0.045	0.020	**0.859**

DISCUSSION

The scenario suggested by the mitochondrial DNA is one of range expansion from the northern Rockies into the southern reaches of the White River. Conversely, the microsatellite data depicted a scenario of fragmentation and isolation of populations in the southern Great Basin. Combining the data resolved the apparent disparity, to reveal a possible history of the montane vole in the eastern Great Basin.

The north to south reduction in cyt*b* diversity was compelling evidence for core and peripheral populations of the montane vole in the eastern Great Basin (Table 3.2). Peripheral populations are suggested to exhibit reduced genetic variation due to smaller effective population size and/or isolation (Schwartz et al., 2003). However, in this case, the reduced genetic diversity was due to recent expansion from a source. The northern clades of Steptoe and Stillwater demonstrated a higher number of haplotypes, higher haplotype diversity and number of mutations (*S*). Higher diversity is expected in ancestral populations, especially when the effective population size was large. Steptoe and Stillwater both demonstrated a large effective population size that was slightly lower than Pahranagat.

Populations that recently colonized an area are expected to be younger than the source population from which expansion occurred (Rogers, 1995). The two northern clades represented by the southern Sierra Nevada and Steptoe Valley demonstrated the earliest divergence in the late Pleistocene. Further, the divergence estimates for the northern clades did not overlap with those of the southern clades, which were estimated in the late Pleistocene to early Holocene. My mismatch analyses indicated unimodal haplotype distributions for the southern clades of Kirch, Key Pittman and Pahranagat

(Table 3.7). The Poisson distribution expected from relatively rapid expansion across the landscape has significant theoretical support (Harpending, 1994; Rogers, 1995). My result demonstrated support for the correlation between genetic diversity and relative clade age.

Overall, the cyt*b* phylogeny suggested late Pleistocene connectivity between a source population/s in the northern Rocky Mountains and those to the south in the White River valley. Despite this connectivity, the phylogeographic analysis suggested an incipient breakdown of connectivity between the populations. The partitioning of genetic variation yielded strong support for a northern clade that consisted of populations from Kirch and Steptoe and a southern clade comprised of the Key Pittman and Pahanagat populations. The topology suggested a barrier separated the northern and southern White River populations beginning after the late Pleistocene. Two potential barriers included hydrological changes along the river since the end of the Pleistocene and the geohydrological dynamics of the region.

The distribution of montane vole clades may be due to the hydrological history of the White River. During the Pleistocene, the White River received regular glacial runoff from the Bonneville Basin (Metzger et al., 1973). The runoff likely supported a number of riparian meadows along the rivers' approximate 322km length, acting as a dispersal corridor for the montane vole. The retreat of the glaciers precipitated a disruption of population connectivity. Hydrological changes that began by the end of the late Pleistocene continued into the Holocene. Throughout the Holocene, the White River became increasingly desiccated, which made successful regional dispersal by voles more unlikely. The shift in environmental conditions from mesic to xeric may have

precipitated a concomitant shift in vole connectivity. The regional gene flow that likely predominated during amenable periods of the Pleistocene was replaced by more localized gene flow within disjunct patches of spring-fed riparian habitat. The scenario of post-Pleistocene desiccation is suggested for other taxa. A similar scenario of lake and marsh desiccation was suggested to explain the contraction and extinction of other Great Basin populations (Grayson, 1993). In addition, the Caliente-Enterpise Accomodation Zone, a region of volcanic and tectonic intrusion perpendicular to the middle extent of the White River, (Reheis, et al., 2002) is a suspected topographic and biological barrier late Pleistocene and Holocene (Polhemus and Polhemus, 2002).

Of course, the topology could be explained by factors other than changes in hydrological patterns. If the expansion occurred over a long period of time, the polymorphisms ancestral to the source population may have been distributed as a decreasing curve; with populations closer to the source retained more of the ancestral polymorphisms compared to those farther from the source. The partitioning of variation into northern and southern clades may have resulted from incomplete sampling. The addition of genetic data from montane vole populations not yet uncovered in field surveys may erase the topological pattern.

The north to south reduction in genetic diversity was also demonstrated for the microsatellite data. While cyt*b* data indicated historical connectivity and the potential for more recent fragmentation into northern and southern clades, the microsatellite analyses suggested a greater potential for recent fragmentation and isolation of White River populations. The level of genetic diversity, as measured by allelic richness (AR) and private allelic richness (PAR) was higher in the northern clades even while controlling for

missing data and uneven sample sizes among populations. The north to south decline in

genetic parameters was also demonstrated for the unbiased estimator of expected

heterozygosity (H$_{UE}$). The greater than expected H$_{UE}$ for Pahranagat may be the result of

large historical effective population size (Ne), which was among the largest Ne

recovered. In addition to a reduction in genetic diversity, the reduction in the *M*-ratio for

southern populations was the likely result of genetic bottlenecks (*M*-ratio, Garza and

Williamson, 2001). The reduction in genetic diversity parameters was especially severe

in the population at Kirch, including a significant deviation from Hardy Weinberg

equilibrium through an analysis of inbreeding coefficient (F$_{IS}$). Peripheral populations are

expected to be smaller, and therefore exhibit reduced genetic diversity compared to

larger, core populations (Frankham, 1997). Patterns of microsatellite variation indicated

here support the reduced genetic diversity of peripheral populations observed in other

studies (e.g., Huang et al., 2005; Schwartz et al., 2003).

I suggest that the putative peripheral populations of the montane vole along the

White River were the result of more recent climate and hydrological influences. Current

abiotic conditions differ between the northern end and the southern end of the White

River valley. The southern end of the Great Basin is more xeric, and most streams and

rivers in the southern end of the Great Basin are ephemeral, including the White River

(Reheis et al., 2002; Harrill and Prudic, 1998). The southern basin is also significantly

warmer compared to the northern basin (Polhemus and Polhemus, 2002). The greater

density of perennial streams in the northern Great Basin was likely due to reduced

evaporation and greater inputs from the Rocky Mountains. As a consequence, the

northern basin may maintain significant riparian habitats, and perhaps larger and more

contiguous vole populations, despite post-Pleistocene climate change (Harrill and Prudic, 1998).

The isolation by distance test has been used successfully to distinguish peripheral populations (e.g., Vignieri, 2005; Hutchinson, 2003; Hutchinson and Templeton, 1999). However, my examination of peripheral populations using the approach was inconclusive due to limited sample size for some populations. The influence of limited sampling was supported by simulations, which suggested that a sample size ≥ 50 would be needed for reliable estimation of gene flow. An earlier study recommended a minimum of 20 samples per population because this is the minimum number at which the genetic parameters between high and low diversity populations converge on their true values (Pruett and Winkler, 2008). My difficulty in obtaining an adequate sample size may be related to the discontinuous distribution of the montane vole, as noted for other organisms (Pruett and Winkler, 2008.

Despite potential shortcomings related to sampling, my study offers additional insight into the dynamics that shape the distribution of core and peripheral populations in the Great Basin. I gained this insight through the analysis of two molecular markers. Rather than a statistical impediment, the differences in the presumptive mutation rate between the mitochondrial DNA and microsatellite markers can inform population dynamics at different time scales. While the cyt*b* indicated historical connectivity maintained by glacial runoff along the White River basin, the microsatellite data suggested significant changes in population connectivity due to the continued desiccation of the Great Basin since the end of the Pleistocene. Taken together, the results appear to support the theoretical expectations for the genetic diversity of core and peripheral

populations (Sagarin and Gaines, 2002; Frankham, 1997; Brown, 1984; MacArthur and Wilson, 1967), which suggest that peripheral, isolated populations are smaller in size and exhibit lower genetic diversity, including a propensity for genetic bottlenecks (Karron, 1987) compared to connected (i.e. core) populations (Frankham, 1997).

The addition of genetic data may improve the equilibrium theory (Heaney, 2000; MacArthur and Wilson, 1967) and may also resolve the core/periphery debate in mammalian extinction dynamics. Genetic data is a recent contribution to biogeographical study (Parker and Marwith, 2007) but is already showing great promise as an independent source of information about species history (Riddle et al., 2008). My analyses of mitochondrial and nuclear variation for the montane vole represented this new integrative biogeography. In suggesting a greater role for vicariance in the distribution of genetic variation in the montane vole, my study offers insight into the evolutionary dynamics that shaped the distribution of mammal taxa in the Great Basin, and an avenue for future research in this vital region.

ACKNOWLEDGMENTS

Chapter IV

Ecological and genetic correlates of extirpation in Great Basin populations of the

montane vole (*Microtus montanus*)

A multi-disciplinary approach is a potentially productive way to evaluate extinction risk because ecological and evolutionary factors may influence extinction over different time scales. Here, I analyze the ecological and genetic characteristics of *Microtus montanus*, the montane vole at Ash Meadows to examine the parameters influential to its recent extirpation. For comparison, I also examined the ecological and genetic variation from an additional 8 vole populations in the region. I summarized the genetic diversity parameters within and between vole populations through an analysis of the cytochrome b (cyt*b*) gene from 97 montane vole samples for 9 populations. The effect of *Mu*s captures on extirpation risk was also examined. I selected 30 ecological variables related to the species habitat requirement to examine the relationship between abiotic parameters on extirpation risk. The effect of independent variables on extirpation was assessed using a logistic regression and classification tree approach. My analyses demonstrated that Ash Meadows was significantly smaller, warmer and drier compared to other sites. The site also had the highest density of *Mus* captures. The population at Ash Meadows demonstrated a modest reduction in molecular diversity and a late Pleistocene population bottleneck. Results indicated that extirpation at Ash Meadows was the result of ecological change at different times; the genetic response to one of these disturbances was indicated. This study underscores the potential to gain valuable understanding of extinction processes using a multi-disciplinary approach.

INTRODUCTION

The complex way that ecological, genetic and biotic parameters interact to influence extinction risk is unclear. Recent studies suggest a causal link between extinction risk and parameters such as climate variability, spatial distribution, habitat area, climate suitability and autoecological characteristics (Raxworthy et al., 2008; Frey et al., 2007; Duncan and Forsyth, 2006; Fagan et al., 2005; Beever et al., 2003; Burbridge et al., 2002; Davis and Callahan, 1992; Lomolino and Channell, 1995; McDonald and Brown, 1992; Brown, 1971). It is also suggested that endangered species populations may exhibit significant genetic decline compared to non-endangered conspecific populations (Spielman et al., 2006).

Small, isolated populations are expected to be at greater risk of extinction due to random fluctuations in population size, genetic drift and the influence of stochastic environmental events (Frankham, 1995; Lande, 1988). An inverse correlation between genetic diversity and extinction risk is expected (Frankham, 1995). Moreover, the genetics of endangered populations are cyclical: a decrease in diversity leads to reduced fitness and reproductive potential, reducing population size, leading to further reductions in diversity (Spielman et al., 2006; Frankham, 2002). Despite the theoretical expectations between genetic diversity and extinction risk, some researchers suggest that environmental and/or demographic factors are a more immediate cause of extinction (Lande, 1988). Clearly there is a lack consensus regarding the potential interactions between demographic, ecological and genetic parameters on extinction risk.

A variety of techniques typically have been used to assess the vulnerability of a population to extirpation, such as population viability analysis, incidence-function

121

analysis or qualitative ranking (Lacy, 2000; Hanski, 1997; Mace and Lande, 1991). A multi-data approach to extinction studies has been suggested because different variables may influence extinction over different time scales (Dunham et al., 1999). Detecting the temporal variation in extinction dynamics may be possible even within one approach. For example, the mutation rate of different genes, or the rate of change expected in different genetic parameters may reflect the genetic response of a population at different times in its history. A recent study employed a classification tree approach on both ecological and historical data to evaluate the vulnerability of birds on oceanic islands (Boyer, 2008), which uncovered two temporally spaced extinction events. Classification trees are robust to data that vary in type, scale and distribution, and hence may be particularly useful when combining ecological and genetic data. An integrated method that combines genetic and ecological data using a classification tree approach is emerging (Gebremedhin et al., 2009).

Mountains of the Great Basin were a focus of several studies of the extinction dynamics in mammals. In a seminal study, Brown (1971) concluded that the current distribution of small, isolated populations of mammals were primarily the result of extinction processes, although this result was not supported by subsequent genetic analyses (Floyd et al., 2005). The Great Basin is an ideal forum for the examination of extinction dynamics because of its diverse small mammal community and its topography, which may increase the isolation of less vagile taxa.

Here, I examined the patterns of ecological and molecular diversity within and between 9 populations of the montane vole, *Microtus montanus* to characterize extirpation risk. The montane vole may be susceptible to extirpation because of its highly

122

specialized ecology and patchy distribution. Although *M. montanus* is distributed throughout western North America in mountains from the Cascades and Sierra Nevada to the southern Rocky Mountains and Colorado Plateau (Hall, 1981, Fig. 4.1), within this range it tends to be found only in riparian habitats. The disjunct distribution is especially noteworthy along the southern edge of the range, which includes the Great Basin. In the Great Basin, the montane vole occupies moderate elevations (1000m-2500m) in wet meadows (Hall, 1946).

One of these sites, Ash Meadows in the Amargosa Valley provides a unique window into extinction dynamics. The species was considered extirpated after repeated surveys during 2007-2209 totaling over 15,000 trap nights did not recover any montane voles. Yet this is a documented historical locality for the species. The last montane voles were trapped in 1933; these animals (n = 13) are housed at the Museum of Vertebrate Zoology (MVZ) at the University of California at Berkeley.

My study incorporated a variety of data including information from field surveys, ecological and genetic and data to assess extirpation risk in the montane vole. To represent the range of variation within populations of the montane vole, I included additional data from populations in the Colorado Plateau, Rocky Mountains and Sierra Nevada Mountains. I selected field survey sites within the Great Basin to represent a spatially contiguous sample across the region. I surveyed the historical vole locations of Ash Meadows in southwestern Nevada, Stillwater/Fallon region in western Nevada and Fish Lake Valley, in northwestern Nevada. For Fish Lake and Ash Springs, surveys were limited by time and logistical constraints. In addition, I surveyed additional sites along the White River in the eastern great basin for an analysis of population connectivity

Figure 4.1. Locations of *Microtus montanus* in the Great Basin examined in this study.
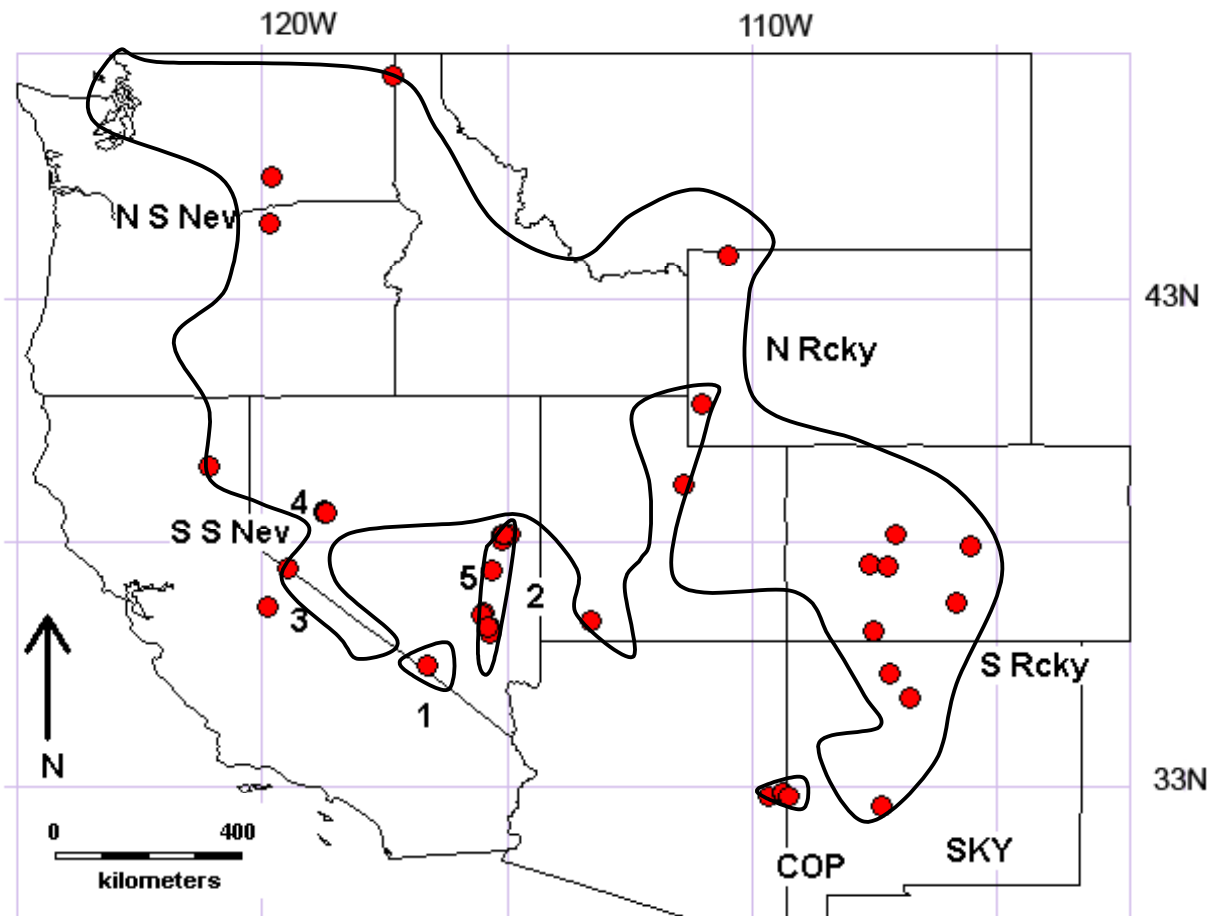
(Crawford and Smith, in prep). The sites included Steptoe Valley in northeastern Nevada, Kirch Wildlife Management Area, Key Pittman Wildlife Management Area and Ash Springs in east-central Nevada, and Pahranagat/Crystal springs in southeastern Nevada (Fig. 4.2).

A change in ecological conditions that leads to the contraction and/or disappearance of suitable habitat may lead to the extinction of animal populations. It has been suggested that the montane vole is a riparian obligate (Findley and Jones, 1962; Anderson, 1959). Therefore, the ecological parameters that I selected included those that quantify the condition of the riparian habitat for each vole location, such as the distance to the nearest water feature, the density of river per unit area, mean annual temperature and annual precipitation.

Extinction may be inversely correlated with genetic diversity (Frankham, 2002; Frankham, 1995). I quantified the molecular variation from the cytochrome b gene (cyt$b$) of the mitochondrial DNA from both extirpated and extant populations to assess this possibility. The genetic parameters that I examined included the number of mutations (S), nucleotide diversity (π), number of haplotypes (h), haplotype diversity (Hd) and the average number of pairwise differences (k).

To examine extirpation factors for the Ash Meadows population, I conducted two analyses. For the analyses of abiotic parameters on extinction, I examined a dataset of 1050 montane vole locations, which was divided into 11 geographical populations, 3 extinct populations (Ash Meadows, Ash Springs and Fish Lake Valley) and 8 extant populations. For the biotic data, I conducted analyses on 64 montane vole samples, representative of 6 geographic populations, and further subdivided into 1 extinct

Figure 4.2. Locations of montane vole populations sampled from the White River Basin.
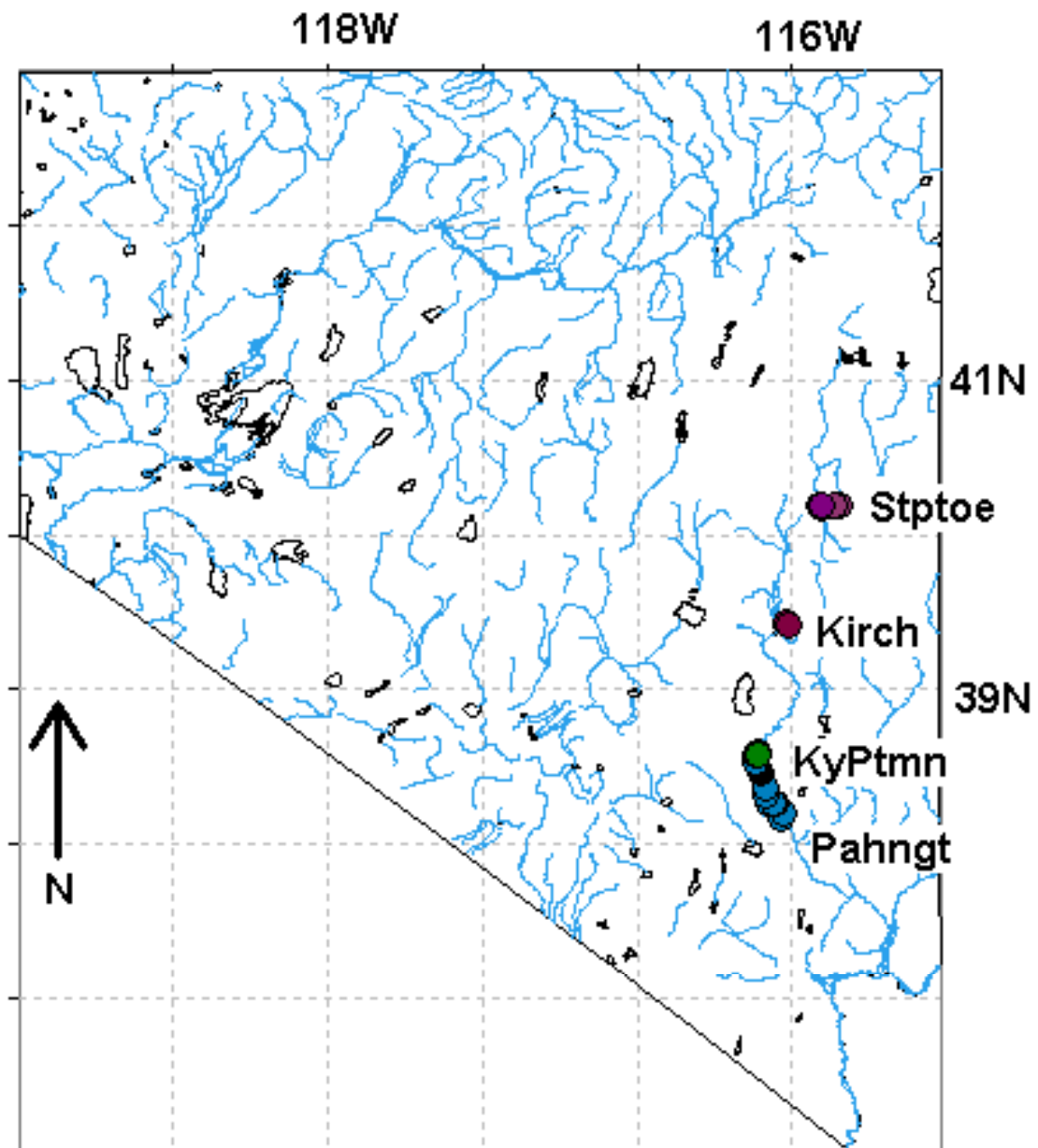
population (Ash Meadows) and 5 extant populations.

MATERIALS AND METHODS

I conducted field surveys to collect demographic data for each population and to collect tissue for genetic analyses (Table 1). Surveys were standardized using the guidelines for live trapping mammals (American Society of Mammalogists, 1987). Each population was surveyed a minimum of 3 times during the spring, summer and/or fall seasons from 2006-2009, using a minimum of 1000 trap nights each time; total trap nights was 41,300. Traps were set as lines of 20-25 traps perpendicular to the nearest stream or river for a total of 200- 250 traps per location. Traps were set 1.5m apart with the trap lines set approximately 10m apart. To minimize the influence of overnight moisture and/or solar radiation on the captured animals, I covered each trap with a thick layer of vegetation where necessary. Because the montane vole appears to prefer areas of dense herbaceous cover, traps were placed in shady locations wherever possible. Traps were checked twice a day: once in the early morning, and again before noon. Traps were closed between 10am-12pm, and opened again an hour before dark. Standard data was collected from each animal, including species, sex, relative age and reproductive conditions. I also collected the latitude, longitude and elevation for each capture for subsequent statistical analyses. A 2mm plug of ear tissue was collected for genetic analyses. In some cases, DNA was obtained from vouchers.

I recovered montane vole from all trapping surveys except the locations of Ash Meadows, Fish Lake valley and Ash Springs. My field surveys from 2007-2009 demonstrated strong support for the extirpation of the montane vole at Ash Meadows in the Amargosa River valley in southwestern Nevada. The trapping surveys conducted at

Ash Meadows include my surveys of over 4000 trap nights, an additional 1800 trap nights through a collaboration with the Nevada Division of Wildlife and a final set of 10,500 trap nights was conducted by Wildlife West, Inc.. No surveys produced montane vole captures (Table 4.1). The refuge is composed of over 22,000 acres of habitat, and approximately 1/3 is spring-fed grasslands. Ash Meadows has been the site of aggressive development attempts from the early 1900's until the refuge was established in 1984. Potential disturbances to the population of montane vole were the diversion of spring-fed streams for agriculture and the removal of peat in the 1970's.

Surveys also indicated tentative support for the extirpation of the species at Ash Springs and Fish Lake Valley. The Nevada Division of Wildlife acquired Ash Springs in the spring of 2009 after the property was privately held since the mid 1900's. The site is named after the spring, and is a historical locality for the montane vole from surveys in the early 1900's. The hot springs in the area surrounding Ash springs are a popular destination for recreation. Despite the chain link fence surrounding Ash Springs, the surveys conducted in 2009 indicated significant disturbance to the habitat, with much of the riparian vegetation community absent or in significant decline.

My 2007 survey at Fish Lake Valley did not recover any montane voles. Much of the property described as a historical locality for the species is under private ownership since the mid 1900's. I conducted a road survey of the property and noted that a fire had destroyed mush of the riparian community, and the spring was dry.

Table 4.1. Summary of data collected from field surveys of Great Basin montane vole populations from 2006-2009. For each location (site), the dates of surveys, the number of trap nights (the number of times traps at each site were available for capture), the number of montane vole captures (#*Microtus*) and the percentage of *Mus musculus* (%*Mus*) captured (#*Mus*/total # captures) are provided. na= data not available

Site	Dates	# trap nights	#*Microtus*	%*Mus*
Ash Meadows	11/2006	2000	0	0.80
	3/2007	2000	0	0.65
	6/2007	1800	0	0.55
	5/2008	5500	0	0.75
	9/2008	4500	0	na
	5/2009	1000	0	na
Fish Lake Valley	3/2007	1000	0	0.35
Kirch	3/2007	2000	3	0
	5/2007	2000	5	0.08
	8/2008	1000	5	0.15
Steptoe	5/2007	1000	9	0.05
	3/2008	1000	5	0.07
	8/2008	2000	15	0.10
Stillwater	9/2007	1000	6	0.07
	10/2007	1000	13	0.05
Pahranagat	9/2006	1000	6	na
	10/2006	1000	2	0.10
	5/2007	1000	3	0.25
	6/2007	1000	0	0.20
	9/2007	1000	2	0.05
	8/2008	1000	0	0.08
	10/2008	1000	5	0.07
Key Pittman	8/2006	1000	2	0.15
	5/2007	1000	0	0.15
	8/2008	1000	11	0.05
	10/2008	2000	0	0.60
Ash Springs	3/2009	1500	0	0.55
	5/2009	1000	0	0.50

Genetic samples

Isolated populations may be at greater risk of extinction due to ecological, demographic and/or genetic factors. I examined the relationship between genetic diversity and extirpation risk for the Ash Meadows population though analyses of the cyt*b* gene. In addition to the *M. montanus* tissue samples collected during field surveys, historical samples for 6 animals at Ash Meadows were obtained from the Museum of Vertebrate Zoology. The genetic dataset consisted of 64 samples from 9 populations.

I extracted nucleic acids from 64 *M. montanus* samples using a modified salt extraction procedure (Medrano et al., 1993). For museum skins, an additional processing step was added and samples were soaked in 10% 10mM TRIS buffer at 45_oC for 1 hour prior to extraction. I chose the cyt*b* gene because its relatively high mutation rate is useful for resolving nodes younger than 2 Ma (Galewski et al., 2006). The cytochrome *b* (cyt*b*) was amplified using the polymerase chain reaction (PCR), and primer pairs MVZ05-MICRO-06 and ARVIC07-VOLE14 (Hadly et al., 2004). Reaction conditions for PCR were 95^oC (30 s), denaturing, 50^oC (25 s) annealing, and 72^oC (1 min) extension. PCR products were purified using a QIaquick PCR Purification Kit (Qiagen, Valencia, California). Automated sequencing used dye-labeled terminators and cycle sequencing conditions of 96^oC (10 s) denaturing, 50^oC (5s) annealing, and 60^oC (4 min) extension. PCR products were sequenced using the same primer sets as above, and purified using a sodium acetate and ethanol precipitation procedure. Skin samples were amplified and sequenced three separate times to examine the potential influence of contamination. Sequenced samples were read using an ABI 3100 Automated Sequence Analyzer.

Sequences were manually aligned using Sequencher v.4.7 (Bromberg, 1995). I checked

the validity of each *Microtus* sequence by comparing each sequence against published

sequences of *M. montanus* or related *Microtus* on GenBank.

Biotic data

To examine the relationship between correlates of genetic diversity and extinction

risk, I choose to use the number of haplotypes per population (h), haplotype diversity

(Hd), the number of mutations (*S*), nucleotide diversity (π), theta (Θ) and the mean

number of pairwise differences (*k*). Each of these parameters is thought to change at

different rates and as such may provide information about processes acting on different

spatial scales. I estimated the parameters in DnaSP v.4.90 (Rozas et al., 2003) on

extinction risk. The influence of population size was assessed by estimating theta (Θ) in

MIGRATE v.3.0 (Beerli and Felsenstein, 2001), which is used to calculate effective

population size (N_e). Typically, Θ for haploid data is presumed to be equal to $\Theta = 2N_e\mu$,,

where μ is the mutation rate of μ=2.0% substitutions/lineage/million years (Brown et al.,

1979).

Environmental disturbances such as climate shifts are suspected to induce

demographic changes in isolated populations. The changes are potentially detectable

through tests of neutrality on molecular parameters. Isolated populations may be at

greater risk of extinction following a bottleneck, which reduces population size and can

substantially reduce genetic diversity through random genetic drift on small populations.

I conducted tests of neutrality in ARLEQUIN v.3.11 (Excoffier et al., 2005). To examine

this, the potential for equilibrium between mutation and drift was assessed with Fu's *Fs*

(Fu, 1997) and Tajima's D (Tajima, 1989) test statistics. Tajima's D examines the differences between the number of segregating sites (S) and the relative number of pairwise nucleotide differences (π): at equilibrium, $S = \pi$. The test detects a population bottleneck (i.e. population contraction) when $S < \pi$, which yields a positive value for D. Conversely, a population expansion is due to $\pi > S$, indicated by a negative D. I conducted a second test of neutrality with the Fu's F_s test, which tests the difference between the observed (k_O) and expected number of alleles (k) in a sample (Fu, 1997).

Assuming neutrality of the mitochondrial DNA, a population expansion event will increase k_O relative to K, which results in a negative Fs. A positive F_s when $k_O < k$ supports a genetic bottleneck event. An excess of low frequency alleles, which may indicate population expansion and result in a significantly negative Fu's F_s test, may also be due to purifying selection or selective sweeps. To distinguish between population expansion and selection, I conducted a third test of neutrality using a Fu and Li's D^* test (Fu and Li, 1993) in DnaSP (Rozas et al., 2003), which may detect selection when Fu's F_s and/or Tajima's D are also significantly negative.

Earlier work suggested the density of *Mus musculus* was positively correlated with degree of habitat disturbance (Stacey and Post, 2009). I examined the influence of *Mus* density on extirpation risk by first calculating the percentage of *Mus* captured at each site by dividing the average number of *Mus* captured by the total number of captures.

Abiotic data

I examined the contribution of 30 ecological parameters on the extinction likelihood of the montane vole in the Great Basin (Table 4.2). The variables were

selected based on the suggested habitat specificity reported for the montane vole for

mesic grasslands (Anderson, 1959). I vetted a dataset of 3954 locality records from the

Global Biodiversity Information Facility (http://data.gbif.org/datasets/resource) and

Arctos Multi-Institution Multi-Collection Museum Database

(www.arctos.database.museum). The locality records were proofed and corrected for

errors due to missing latitude and/or longitude data. Duplicate records or records without

locality information were excluded. I also eliminated data with a coordinate precision

below 15km as calculated in Biogeomancer (www.biogeomancer.org). This left a dataset

of 1050 locality records. I categorized each location into one of 11 populations based on

geographical location, as reported from a phylogeographic analysis (Crawford and Smith,

in prep). I further subdivided the populations into groups: known extinct or questionable

status populations (category = 0, n = 3) and extant populations (category = 1, n = 8).

I extracted values for 19 bioclimatic data layers with 1 Km resolution for the

montane vole locations in Nevada using DIVA-GIS v. 5.2 (Hijimans et al., 2005a,

2005b). I also selected 5 water features because surveys suggested the species is a

riparian obligate in the Great Basin (Table 4. 2). To examine vole locations as a function

of distance from water features, the feature datasets of: all streams, perennial streams

only, wetland, and pluvial lake feature datasets were converted into grids. Each grid

represented the distance between each vole location and the feature as calculated with

spatial analyst functions in ARCGIS v. 9.3. I imported the distance grids into DIVA-GIS

v.5.2 (Hijimans et al., 2005b) to extract the numerical value from each grid for each vole

location. I also calculated the density of perennial streams in the study area in ARCGIS v.

9.3. The influence of habitat area on extinction was estimated by calculating the area and

Table 4.2. Data type, resolution and source for each of the 30 variables used in

analyses of extinction parameters for the montane vole.

Data Coverage	Resolution	Data Source
Land cover	1Km	Global Land Cover Facility
Soil	1Km	Natural Resources Conservation Service SSURGO v.2.1
Geological formation	1:24,000.	Keck Earth Sciences and Mining Research Information Center
Pluvial lakes	1:24,000	Great Basin Center for Geothermal Energy
Perennial stream	1:2,000,000	National Atlas (www.nationalatlas.gov)
Wetlands	1:24,000	National Wetlands Inventory (www.fs.gov)
BIO1	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO2	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO3	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO4	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO5	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO6	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO7	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO8	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO9	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO10	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO11	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO12	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO13	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO14	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO15	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO16	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO17	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO18	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
BIO19	$1km^2$ WORLDCLIM data	(Hijimans et al., 2005a)
Water Table	1:24,000	Great Basin Center for Geothermal Energy
Area	ArcGIS	
Perimeter	ArcGIS	
MUS	Field Survey data	
DistNeighbor	Calculated as straight-line distance in ArcGIS	

perimeter for the largest polygon of suitable habitat at each site using spatial analysis

functions in Hawth's Tools (Beyer, 2004). A minimum convex polygon was constructed

around each of 11 vole populations (i.e. Pahranagat, Key Pittman) in ARCGIS v. 9.3 and

the perimeter and area measured. The influence of land cover, soil and geology was

estimated by extracting the value for each feature for each vole location in DIVA-GIS

v.5.2 (Hijimans et al., 2005b).

Statistical analyses

The strength of each independent variable on extinction was assessed in a

MANOVA. Non-significant variables were excluded from subsequent analyses. A

regression analysis was conducted to further characterize the influence of each predictor

on the dependent variable extirpation risk (STATUS). Finally, the influence of each

independent variable was assessed through a classification tree analysis in MATLAB

v.7.4.0.287 (The MathWorks, Inc.). The final dataset was composed of potentially

crosscorrelated numeric and categorical (geological formation, land cover, soil) data. The

classification tree algorithm in MATLAB v.7.4.0.287 is well suited to these types of data.

The method searches for the best fit of the data to the dependent variable using a

bifurcation approach. The classification approach predicts the response variable at each

node based on a true/false criterion. A "true" answer proceeds to the left of the node, and

a "false" answer moves to the right. Finally, a final model is calculated by iteratively

fitting a multiple of trees to the data. The tree produced by this method tends to have

many bifurcating branches, which can lead to overfitting. I pruned the subtrees from the

model with a minimum error algorithm that yielded a final tree.

RESULTS

A MANOVA of genetic variables on STATUS showed that Ash Meadows demonstrated the lowest number of segregating sites, S (df = 2, F = 17.805, P = 0.000) . The population at Ash Meadows was also characterized by low mean number of pairwise differences (k), lowest number of haplotypes, reduced haplotype diversity, but modest nucleotide diversity (Table 4.3). Of the genetic parameters examined, S was the most significant predictor of STATUS (df = 2, F = 10.811, P = 0.000). My examination of neutrality demonstrated that Ash Meadows demonstrated significantly positive values for Tajima's D and Fu's F_s test statistics, suggestive of a population bottleneck (Table 4.4). The effective population size (N_e) indicated the N_e for Ash Meadows (N_e = 1,910,000) was only greater than the population at Kirch (N_e = 240,000) and less than the average (N_e = 3,862,000). The average *Mus* density was the highest at Ash Meadows (df = 5, F = 4020, P < 0.0001). When I included the data for *Mus musculus* captures, the variable MUS was identified in the logistic regression (df = 95, MS = 2.255, F = 246.9, P = 0.000, R_2 = 0.72) as the most significant predictor of STATUS. My analyses using a classification tree approach were congruent with the logistic regression: there is a 68% probability that extirpated populations exhibit a MUS ≥ 60% (Fig. 4.4).

An analysis of 33 independent ecological variables on extinction yielded 29 significant variables (Wilks' Lambda = 0.082, F = 328, df = 14, P = 0.000). A regression that added the independent variables randomly demonstrated that the distance to perennial streams was the most significant predictor of STATUS, followed by steam distance, wetland distance and the annual mean temperature (bio1) (df = 26, F = 370.69, P = 0.000, r_2 = 0.71). Ash Meadows was the farthest from perennial streams compared to

Table 4.3. Genetic diversity parameters estimated from analysis of the cyt*b* gene. The 64 samples represent 9 populations subdivided into extirpated (STATUS = 0) or extant (STATUS = 1). Parameters estimated include the average number of pairwise differences (k), haplotype diversity (Hd), the number of haplotypes (h), the number of segregating sites (S) and nucleotide diversity (π). Sample size (N) for each population is also provided.

Population	N	Status	k	Hd (S.D.)	h	S	π (S.D.)
Ash Meadows	6	0	2.4	0.80 ± 0.03	4	4	0.014 ± 0.01
S SNev	12	1	6.89	1.00 ± 0.002	5	19	0.010 ± 0.004
N SNev	4	1	4	0.67 ± 0.09	2	21	0.004 ± 0.004
Phngt	11	1	5.93	0.85 ± 0.01	7	11	0.22 ± 0.001
KyPtmn	8	1	2.11	0.96 ± 0.01	8	7	0.014 ± 0.001
Stptoe	6	1	14	1.00 ± 0.09	6	37	0.024 ± 0.002
Kirch	6	1	1.31	0.90 ± 0.03	4	8	0.001 ± 0.001
COP	7	1	24	0.95 ± 0.01	6	43	0.017 ± 0.002
SRcky	5	1	18	1.00 ± 0.03	5	19	0.016 ± 0.003

Table 4.4. Results from tests of neutrality for the cyt*b* gene from 62 montane vole samples. Test of neutrality included Fu's *Fs*, Tajimas' *D* and Fu and Li's *D** test statistics and associated p-values. Results indicated that the population at Ash Meadows demonstrated significantly positive values for Fu's *Fs* and *D* (see text for details). A large, positive value for *D* is suggestive of a population bottleneck.

Population	*Fs* (p)	*D* (p)	*D** (p)
Ash Meadows	0.95 (0.39)	+1.64 (0.96)	-0.07 (0.53)
Steptoe	0.28 (0.46)	-1.05 (0.12)	0.04 (0.46)
Kirch	-0.78 (0.22)	-1.48 (0.14)	0.01 (0.49)
Key Pittman	-2.73 (0.04)	1.34 (0.93)	-0.03 (0.44)
Pahranagat	-0.10 (0.21)	-1.46 (0.05)	-0.08 (0.42)
Stillwater	-1.58 (0.09)	0.22 (0.58)	-0.04 (0.46)

other populations (d.f. = 2, F = 2231.1, P = 0.000). This result is likely due to the higher density of perennial streams found in the northern Great Basin compared to the southern basin (Fig. 4.3). Ash Meadows was also significantly warmer (df = 2, F = 14.54, P = 0.000) and smaller (d.f. = 2, F = 197.01, P = 0.000) compared to other populations. A discriminant function analysis (DFA) correctly classified 69.6% of extinct records and 100% of extant records. I calculated two functions from the DFA; an axis related to perennial stream distance and a second axis correlated to temperature.

A stepwise regression of 29 predictor variables on extinction was conducted to reduce the number of independent variables. The regression model explained approximately 61% of the variation in extinction (Table 4.5). The most significant predictor of extinction was the distance to perennial streams, which explained approximately 38% of the variation in the data, followed by stream and wetland distance, predicting 11 and 2% of the variation, respectively (Table 4.5). My DFA of the same data correctly predicted 86.6% of the original cases for extinct and extant classes. I calculated a regression tree with 29 predictor variables identified from the stepwise regression. The regression tree depicts a model for which distance to the nearest perennial stream distance to nearest stream, and distance to the nearest wetland are the most significant predictors of STATUS (Fig. 4.5).

When a second set of analyses was conducted with the reduced dataset of 64 records, results were similar. A stepwise regression of the reduced set of predictor variables on extinction explained 71% of the variation in the data, and the most significant predictors were distance to the nearest perennial stream, stream distance, wetland distance, mean annual temperature (bio1) and precipitation seasonality.

Figure 4.3 Density of perennial streams in the Great Basin.
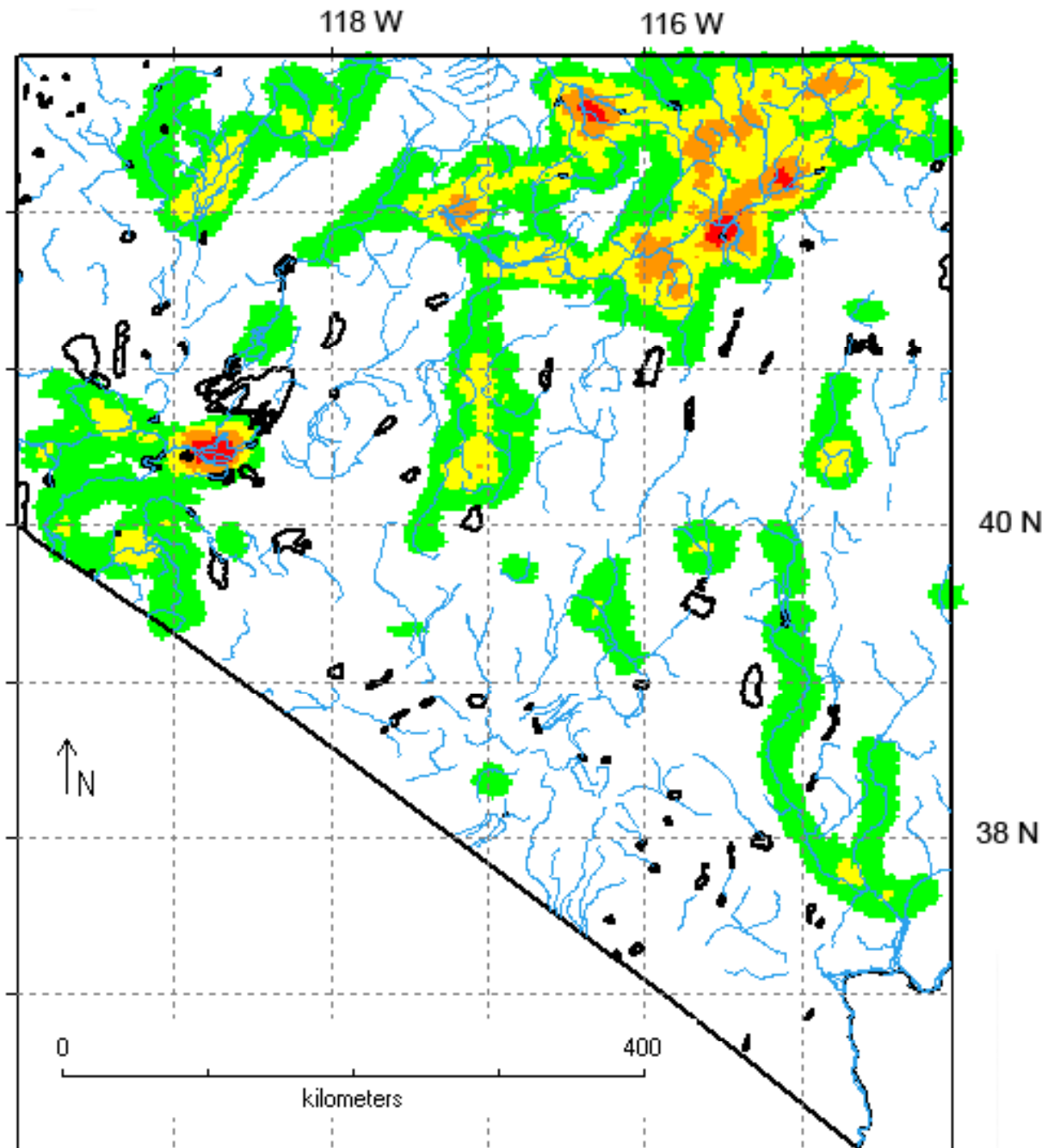
Table 4.5. Regression analysis of 29 ecological parameters on extinction likelihood (d.f. = 29, F = 246.98, P = 0.000). The distance to the closest perennial stream was the most predictive of extinction in the species. The final model contained 8 variables that explained over 60% of the variation in the data (df = 18, F = 519.911, P = 0.000, r^2 = 0.605).

Model Parameters	r	r^2 Change
Distance to perennial streams	0.372	0.372
Distance to closest stream	0.478	0.106
Distance to nearest wetland	0.502	0.024
Bio8	0.516	0.01
Bio1	0.531	0.015
Bio	0.552	0.021
Bio19	0.598	0.046
Perimeter	0.605	0.007

DISCUSSION

Results of this study indicated that extirpation of montane vole at Ash Meadows was the result of ecological change over two time scales. My analyses suggested the combined and cumulative influences of post-Pleistocene climate change and recent habitat alteration to extirpation (Fig. 4.6). Reductions in several genetic parameters suggested the genetic response of the Ash Meadows population to ecological change at the level of the cyt*b* gene. In this case, the genetic data was a key indicator of population vulnerability. While recent habitat change appears to be the most proximate cause for the extirpation, the population was probably at risk, as indicated by the demographic and genetic response to post-Pleistocene climate change.

In unglaciated areas, the cooler and wetter environmental conditions that predominated during the early to middle Pleistocene were probably favorable for dispersal by the montane vole. The montane vole demonstrates an affinity for cool, mesic grassland habitat (Findley and Jones, 1962; Anderson, 1959) and its historical occurrence at Ash Meadows, the driest ad warmest site, suggested that the species colonized the area during more amenable environmental periods.

My examination of ecological parameters indicated that significantly warmer and drier conditions now prevail at Ash Meadows. The significantly greater distance between vole locations and water features (perennial streams and wetlands) at Ash Meadows were indicative of these conditions. The predominately warm and dry conditions at Ash Meadows were the likely consequence of post-Pleistocene climate change, which also precipitated the contraction of riparian habitat throughout the Great Basin (Benson et al., 1996; Brown, 1978). In fact, significant environmental changes beginning in the middle

Pleistocene were reported for the Great Basin (Benson et al., 1996).

The predominance of warm springs and higher temperatures in the southern basin is suggested to promote higher invertebrate diversity (Polhemus and Polhemus, 2002), but the opposite may be true for the montane vole. Habitat contraction that began in the middle Pleistocene would have caused the contraction of vole habitat with concomitant reduction in effective population size. The decrease in population size was detected from analysis if the cyt*b*. The genetic bottleneck inferred from the data seemed intuitive, given the evidence for habitat contraction demonstrated in the ecological data.

In addition, population reductions also cause reductions in genetic diversity through stochastic nature of random genetic drift, birth and mortality. My examination of genetic diversity within and between montane vole populations indicated that Ash Meadows population demonstrated significantly reduced values for some genetic parameters, but modest values for others (Table 4.3). The significant reduction in the number of mutations, the average pairwise difference and haplotype diversity distinguished Ash Meadows. The number of segregating sites is expected to be more sensitive to changes in population size compared to other genetic metrics (Tajima, 1989), and changes in S may represent the initial genetic response of a population to disturbance. The difference in S is likely not due to a shorter coalescent time because estimates of divergence for Ash Meadows population are similar to other populations that demonstrated larger values of S.

Surprisingly, the population at Ash Meadows did not demonstrate the lowest value for nucleotide diversity (π). The expected inverse relationship between diversity and extinction risk is built on the assumption that severe reductions in genetic diversity

are a symptom of endangered populations (Spielman et al., 2006; Frankham, 1995; Slatkin, 1985), although challenged by reports of high genetic variability in endangered populations (Genovart et al., 2007; Rieman and McIntyre, 1993). Nucleotide diversity for Ash Meadows was slightly lower than that reported for Steptoe, one of the northern populations and greater for most of the populations along the southern end of the White River (Kirch, Key Pittman, Pahranagat). The modest value of nucleotide diversity in Ash Meadows suggested two potentially significant findings regarding genetic data. First, the results demonstrate the utility of assessing multiple genetic parameters to assess extinction risk, because different parameters change at different rates. A related issue is that the rate of change in nucleotide diversity over time may be greater than the time elapsed since the original habitat perturbation, which I estimated to be during the middle to late Pleistocene. If so, a significant reduction in nucleotide diversity would not be apparent. Interestingly, the range of nucleotide diversity estimated for Great Basin populations of the montane vole (0.011-0.022) are significantly lower than that reported for *Microtus longicaudus* (0.12-0.44, Conroy and Cook, 2000) or for *M. oeconomus* (0.043-0.628, Galbreath and Cook, 2004), despite similar sample sizes and geographical range.

In addition to post-Pleistocene climate change and the concomitant contraction of riparian habitats, recent habitat disturbance was also influential to the extirpation of the montane vole at Ash Meadows. Relatively rapid ecological change was the likely proximate cause for extirpation. The area had maintained a permanent human settlement since the 1880's (Young and Sparks, 1985), and was the site of ranching, agricultural, residential development. Prior to transfer to the U.S. Fish and Wildlife Service, the

impoundment and channeling of groundwater for residential and agricultural use since the early 1900's has altered much of the original habitat at Ash Meadows (Polhemus and Polhemus, 2002). The majority of the area's peat was mined in the 1970's, approximately ten years prior to the last undocumented sighting of a montane vole by refuge staff. The combined result of human activities has been an estimated 50% reduction in the extent of riparian habitat at Ash Meadows between 1948 and 2004 (Trammell et al., 2008). Hardest hit were the riparian grasslands that support the montane vole.

A genetic response to recent habitat change was not detected through analyses of the cyt*b*. The lack of a genetic signature in endangered populations has undermined the use of genetic data in conservation studies. In a seminal paper, Lande (1988) discussed that the potentially significant impacts of environmental and demographic stochasticity are neglected in studies of extinction, in favor of elucidating a genetic cause. While genetic variability does appear linked to endangered and extinct populations (Spielman et al., 2006), other studies support the influence of random environmental fluctuations (Pimm et al., 1995), or a combination of demographic and environmental factors (Bull et al., 2006). In this case, a significant reduction in genetic diversity was not demonstrated for the Ash Meadows population because the mutation rate and effective population size of the cyt*b* may not be amenable to tracking recent changes.

The use of the cytochrome *b* gene of the mitochondrial DNA presents two potential difficulties. The gene is assumed to be selectively neutral (but see Messier and Stewart, 1997), and not represent the actual genetic response of a population to perturbation. Future analyses incorporating genes under selection, such as nuclear genes may add additional insight. Secondly, although the mutation rate of mitochondrial DNA

is relatively high compared to most nuclear genes, it may be unable to detect an environmental change (Moore, 1995). While it would have been desirable to use genetic markers that mutate at a relatively faster rate (microsatellite marker, e.g.), use of museum skins precluded this. Museum skins are prone to allelic dropout, in which the failure to amplify an allele due to poor DNA quality may be misinterpreted as a homozygous individual (Sefc et al., 2003).

In addition to difficulties associate with the use of cyt*b*, other factors confound my interpretation of extirpation in the montane vole. First, sample sizes used in analyses were low. The extirpated population from Ash Meadows was represented by 6 museum samples. I have not located DNA samples for Ash Springs and Fish Lake Valley. Hence, those localities were excluded from the genetic analyses. I note that this highlighted the importance of voucher specimens and museum collections: the only information that exists for Ash Meadows were the 13 animals collected in 1933.

Despite the studies that suggest a minimal contribution of genetic data to studies of extinction dynamics, my analyses suggested the potential of a multi-data approach. While the genetic markers used here were constrained by the nature of the samples, additional studies that incorporate information from multiple molecular markers may gain insights from the expected differences in mutation rate, effective population size and selection. In fact, a multidisciplinary approach is recommended because of the variation in the rate of change in data from sources such as nuclear and mitochondrial DNA, climate and ecology (Dunham et al., 1999).

Extirpation of the montane vole at Ash Meadows was likely the result to ecological change over disparate time scales. The initial reduction in genetic diversity

may have been a response to climate-induced contraction of suitable habitat that began in the middle Pleistocene. The extirpation of the Ash Meadows population occurred despite modest levels of diversity reported for most genetic parameters. Overall, my analyses appear to suggest that the rate of environmental change exceeded the adaptive ability of the species. Despite the limitations related to sample size and molecular markers, this study adds potentially valuable insight into the role of biotic and abiotic parameters on the long-term viability of Great Basin mammal populations.

Figure 4.4 Classification tree of genetic and demographic variables in extinction.

$MUS < 0.60$, P = 0.285 $MUS > 0.60$, P = 0.678
P > 1, P = 0.889   P < 1, P = 0.111
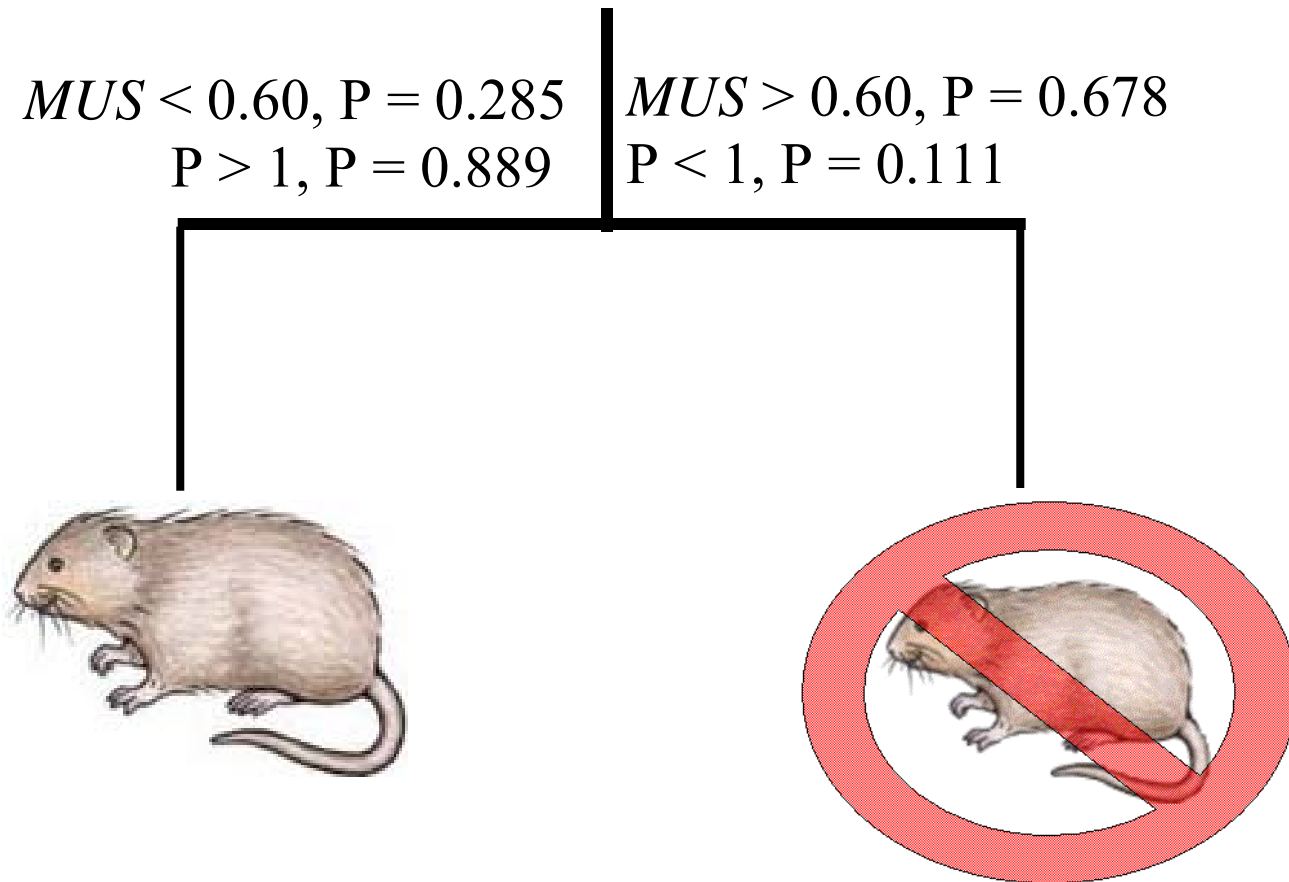
Figure 4.5 Classification tree of ecological and genetic variables on extinction.
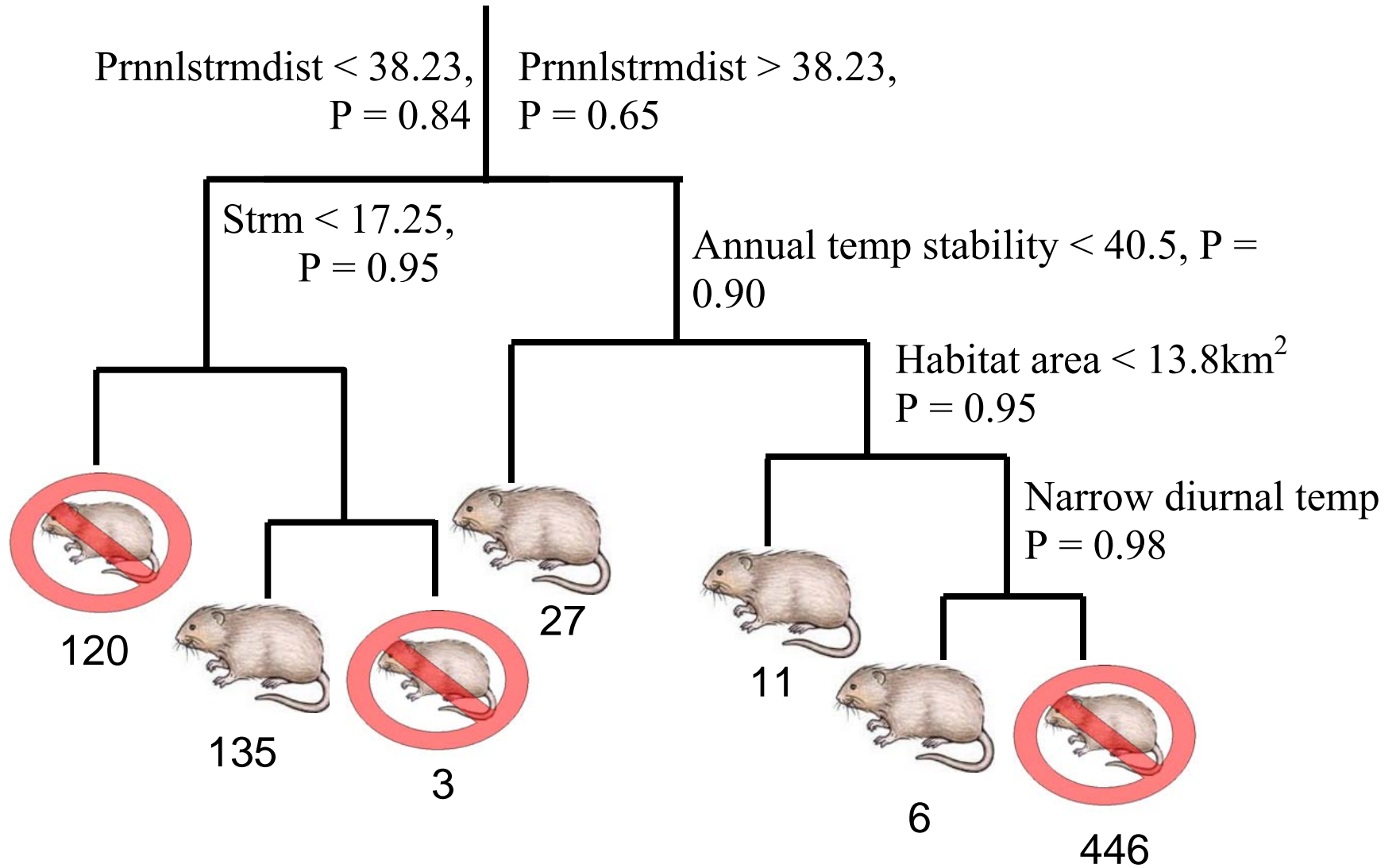
Figure 4.6 Extinction scenario for the Ash Meadows population of the montane vole. (a) Population size may have been reduced following a bottleneck due to late Pleistocene climate change. (b) More recent anthropogenic habitat change further reduced population size, leading to its extirpation.

ACKNOWLEDGEMENTS

CONCLUSION

The incorporation of multiple, independent lines of evidence, a trademark of integrative biogeography, may minimize the need for inferential explanations. Early attempts to explain the distribution of animal taxa were mostly inferential because researchers were limited by the availability of additional evidence, beyond the data concerning where a particular species was found. For example, in the absence of additional evidence, Croziat inferred that the distribution of *Glossopteris* was due to icerafting, which carried *Glossopteris* seeds across the oceans in the southern hemisphere. It was a plausible scenario, given the limited availability of data.

The inferential nature of explanations characterized biogeography into the twentieth century. Even early studies in phylogeography (Avise, 2000) were primarily inferential in nature because of limitations in the availability of paleoecological, geological or paleoclimate data, among others. The additional accumulation of data and the development of spatial and statistical approaches have influenced the biogeographical approach. Recent advances in spatial analysis have allowed the development of modern climate (e.g. WORLDCLIM, Hijimans et al., 2005a) and paleoclimate (PMIP2, www.pmip2.org) data, which was possible due to the accumulated climate data from ice cores, packrat middens, tree cores and pollen analyses. The development of advanced simulation approaches provide a way to combine genetic data from modern and fossil organisms in a single analysis in an effort to reconstruct a more temporally cohesive species history (Drummond et al., 2005; Sheperd et al., 2005; Hadly et al., 2004). Modern integrative biogeography will likely benefit from the current wealth of data and

approaches. This may be especially relevant to the analysis of current distributions, given the current level of habitat alteration and extinctions, which are the unprecedented consequence of both post-Pleistocene and anthropogenic climate change.

The primary motivation for my dissertation research at the University of New Mexico was to examine the interplay between ecology, climate and genetic factors to the distribution of long-term viability of species populations. I chose high elevation rodent taxa because the limited dispersal and habitat requirements suggest they may be more susceptible to habitat change. In addition, rodent taxa comprise over half of the species diversity within Class Mammalia, and an understanding of the dynamics that influence rodent taxa can provide important insights into the evolution of this diverse group and other mammalian taxa as well. The results of my separate studies all highlight the utility of using an integrated approach: whether the approach necessitates the combination of genetic and ecological data or nuclear with mitochondrial data. In all 4 studies, this approach facilitated a holistic explanation for the distribution of *Microtus* voles.

APPENDIX A

Museum catalog numbers and localities for 49 *Microtus* specimens from the Museum of Southwestern Biology (NK) and Michigan State University Museum MSUM) used in the analysis of cytochrome *b* gene sequence variation. One *Microtus californicus* sequence from GenBank (Conroy et al., 2000) was used as an outgroup.

     Microtus mexicanus fulviventer-Mexico: Oaxaca, 15.9 mi. (by road) Guelato De Juarez, NK9688, NK9690, NK9695
     Microtus mexicanus guadalupensis- New Mexico: Jicarilla Mountains, 0.9 mi. SE Jicarilla, NK20767, NK20769, Sacramento Mountains, Eagle Creek, 1.5 mi. W. Hwy 95, NK20716, Torrance Co., 1.7 mi. S., 4.6 mi. W. Manzano, Red Canyon Campground, NK9292, NK9296, Union Co., Sierra Grande, 3 mi. S., 3 mi. W. Des Moines, NK9219
     Microtus mexicanus mexicanus-Mexico: Veracruz, 1.5 mi. S. Altotonga, NK 9706, NK9707, NK9708
     Microtus mexicanus mogollonensis- Arizona: Apache Co., NK20108 (originally identified as *M. montanus*), Coconino Co., Coconino National Forest, Mogollon Rim, Lee Johnson Spring, NK20054, Navajo Co., White Mountains, Stinky Creek, NK31126, Yavapai Co., Prescott National Forest, Bradshaw Mountains, Turkey Creek, NK20053, Bradshaw Mountains, Godwin, NK20940, NK20941, NK29042, NK20944, NK20943, Santa Maria Mountains, Pine Creek, NK20929, NK20930; New Mexico: Bernalillo Co., Residence of Bill Degenhardt, Juan Thomas, NK28499, 5 mi. S., 2 mi. E of Cedro, N Monzano Springs, NK28483, Catron Co., Apache Creek, NK133563, NK136225, Elk Mountains, NK29061, NK29062, Engineer Spring Canyon, 3 mi. W., 0.75 mi. N. Luna, NK20464, Cibola Co., El Morro National Monument, Inscription Rock Pool, NK34977, Ojo Redondo Campground, Zuni Mountains, NK20002, 6 mi. N., 14 mi. E. Grants, NK9182, McKinley Co., Zuni Mountains, McGaffey Lake, NK20287 (originally identified as *M. montanus*), another location NK121579 (originally identified as *M. montanus*), Socorro Co., Magdalena Mountains, Meadow East W1, NK19605, Magdalena Mountains, South Baldy Meadow, NK19643, another location NK62451 (originally identified as *M. montanus*), Valencia Co., 6.7 mi. S., 14 mi. W. Grants, Agua Fria Creek, NK2703
     Microtus mexicanus madrensis-Mexico: Durango, 18 mi. SSW Tepehuanes, pine-oak sacaton, MSUM10540, MSUM10542 and also MSUM10547, MSUM16367, MSUM16368, MSUM33286, MSUM12712
     Microtus mexicanus navajo- Utah: Navajo Mountain, NK20983
     Microtus mexicanus subsimus-Mexico: Coahuila, 10 mi. E. San Antonio de las Alazanas, NK9541, NK9524, NK9543, NK9544
     Outgroup *Microtus pennsylvanicus*-New Mexico: Lincoln Co., Gallinas Mountains, NK9876 (originally identified as *M. mexicanus*), Colorado: Gunnison Co., NK56510 (originally identified as *M. montanus*)

APPENDIX B

Museum catalog numbers and localities for 26 *Microtus* specimens from the Museum of Southwestern Biology (NK) and Michigan State University Museum MSUM) used in the analysis of acid phosphatase V (AP5) nuclear intron variation. Two *Microtus californicus* sequences (Conroy and Neuwald, 2008) from GenBank were used as an outgroup.

   Microtus mexicanus fulviventer-Mexico: Oaxaca, 15.9 mi. (by road) Guelato De Juarez, NK9688, NK9690, NK9695
   Microtus mexicanus hualpaiensis- Arizona: Mohave Co., Hualpai Mountains, Camp Levi, NK20958, NK20957, Wheeler Wash, NK20959 Musil Mountains, near Grand Wash Cliffs, NK20960
   Microtus mexicanus mexicanus-Mexico: Veracruz, 1.5 mi. S. Altotonga, NK 9706, NK9707
   Microtus mexicanus mogollonensis- Arizona: Apache Co., Bradshaw Mountains, Goodwin, NK20941, 15.9 mi. S., 10 mi. E. Springerville, Terry Flat, NK20129, another location NK20108 (originally identified as *M. montanus*), Coconino Co., Coconino National forest, 3.5 mi. S., 4.5 mi. E. Mormon Lake, NK20048, Kaibab National Forest, S. rim Grand Canyon, W Skinner Tank, NK20020, Colorado: Gunnison Co., NK56518; New Mexico: Catron Co., NK29002 (originally identified as *M. montanus*), Cibola Co., El Morro National Monument, Inscription Rock Pool, NK34977, McKinley Co., NK121579 (originally identified as *M. montanus*), Socorro Co., Magdalena Mountains, Meadow East W1, NK19605
   Microtus mexicanus subsimus-Mexico: Coahuila, 10 mi. E. San Antonio de las Alazanas, NK9523, NK9524, NK 9541, NK9542, NK9544, Districto Federal, 17.5 Km S., 7 Km W Toluca, Nevado De Toluca, NK9635
    Outgroup *Microtus californicus*- MVZ198769, MVZ00277

APPENDIX C

Raw data for analyses of the cytochrome b gene in 51samples of the Mexican vole, *Microtus mexicanus*.

NK29061

ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTGTTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TCATAGCAACAGCATTTATAGGTTATGTACTACCATGAGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATATCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTGCCATTTATTATTACCGCCCTTGTA
TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA

NK09296

ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATTGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAACTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGAGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTGCCATTTATTATTACCGCCCTTGTA
TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC

TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCA
GGCATAATCGAAAATAACATCTTAAACCTAGACTAA
NK19605
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGAGGACAAATATCATTC
TGAGGGGCTGCA--

GCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGACTTTCTAGGAGT
CCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCT
CGGAGACCCAGACAATTATACCCCCGCAAATCCACTCAACACTCCACCACCC
ATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCC
TAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCTTTATCCTAGCCC
TCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCCTTCCGTCCAATT
ACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCATCCTCACATGAAT
CGGCGGTCAACCAGTTGAATACCCATTTGTCATCATTGGACAAACAGCTTCA
ATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCAGGCATAATCGA
AAATAACATCTTAGACCTAGACTAA
NK19643
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTACCGCCCTTGTA
TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TAAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCCGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCACC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT

CCTCACATGAATCGGCGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA
NK163897
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATCTGCCTATTCCTTCATGTAGGCCGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTGCCATTTATTATTACCGCCCTTGTA
TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA
NK28499
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCATTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGAGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTGCCATTTATTATTACCGCCCTTGTA
TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TAAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTATCAAAGA
CTTTCTAGGAATCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT

CCTCACATGAATCGGTGGCCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA
NK9219
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATCTGCCTATTCCTTCATGTAGGCCGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTGCCATTTATTATTACCGCCCTTGTA
TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA
NK121579

CAACCACTCATTCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAA
ACTTCGGATCCCTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTA
TTCCTAGCTATACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGC
ACACATCTGCCGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCC
AACGGCGCTTCCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAAT
TTACTACGGCTCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTA
CTATTCGCCGTTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACA
AATATCATTCTGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCAT
ATGTCGGTACAACACTAGTAGAGTGGATTTGAGGTGGTTTCTCAGTAGACAA
AGCCACTCTCACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTAC
CGCCCTTGTATTAGTCCACCTACTATTCCTACACGAAA------------------------------
----

TCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCTTTAT
CCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCCTTCC
GTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCATCCTC
ACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGACAAA

CAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCAGGC
ATAATCGAAAATAACATCTTAGACCTAGACTAA
NK20767
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTA---------
NK20464
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAGCATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATATCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTACCGCCCTTGTA
TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGCCAACCAGTTGAATACCCATTCGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA
NK9296
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATTGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAACTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGAGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTGCCATTTATTATTACCGCCCTTGTA

TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCA
GGCATAATCGAAAATAACATCTTAAACCTAGACTAA
NK9182
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTACGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCCATCCCATATGTCGGTA
CAACACTAGTAGAGTGGATTTGAGGTGGTTTCTCAGTAGACAAAGCCACTCT
CACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTACCGCTCTTGT
ATTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGT
TTGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAG
ACTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTT
TCCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAA
CACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCT
TACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATC
TTTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGC
CTTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCA
TCCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGA
CAAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGC
AGGCATAATCGAAAATAACATCTTAGACCTAGACTAA
NK2703
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGTGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTACCGCCCTTGTA

TTACTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTTT
GAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGAC
TTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTTC
CCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAACA
CTTCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTTA
CGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCTT
TATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCCT
TCCGTCCAATTACACAAACCATTTACTGAATCTTAGTAGCTGATCTTCTCATC
CTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGACA
AACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCAG
GCATAATCGAAAATAACATCTTAGACCTAGACTAA
NK29062

ATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCCCT
ATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTGTTCCTAGCTATAC
ACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGCCG
AGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTTCC
ATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGCTC
CTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCGTC
ATAGCAACAGCATTTATAGGTTATGTACTACCATGAGGACAAATATCATTCTG
AGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATATCGGTACA
ACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTCA
CACGATTCTTTGCTTTCCACTTCATCCTGCCATTTATTATTACCGCCCTTGTAT
TAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTTT
GAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGAC
TTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTTC
CCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAACA
CTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTTA
CGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCTT
TATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCCT
TCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCATC
CTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGACA
AACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCAG
GCATAATCGAAAATAACATCTTAGACCTAGACTAA
NK20002
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGTGGTTTCTCAGTAGACAAAGCCACTCTC

ACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTACCGCCCTTGTA
TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA

NK20769
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGAGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTGCCATTTATTATTACCGCCCTTGTA
TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAAACCAACCGGTT
TAAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGCCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA

NK20983
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCAACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGTGGTTTCTCAGTAGACAAAGCCACTCTC

ACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTACCGCCCTTGTA
TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACAATCAAAGA
TCTA-----------
NK136226
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTCTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGAGGACAAATATCATTC
TGAGGGGCTGCA----------------
NK8710
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGTGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTACCGCCCTTGTA
TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATTT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA
NK20287

CAACCACTCATTCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAA
ACTTCGGATCCCTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTA
TTCCTAGCTATACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGC
ACACATCTGCCGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCC
AACGGCGCTTCCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAAT
TTACTACGGCTCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTA
CTATTCGCCGTTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACA

AATATCATTCTGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCAT
ATGTCGGTACAACACTAGTAGAGTGGATTTGAGGTGGTTTCTCAGTAGACAA
AGCCACTCTCACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTAC
CGCCCTTGTATTAGTCCACCTACTATTCCTACACGAAA------------------------------
----
GCAGACAAATTCCCTTTCCACCCCTACTACACTGTCAAAGACTTTCTAGGAGT
CCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCT
CGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAACACTCCACCACAC
ATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCC
TAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCTTTATCCTAGCCC
TCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCCTTCCGTCCAATT
ACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCATCCTCACATGAAT
CGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGACAAACAGCTTCAA
TCGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGAAGGCATAATCGAA
AATAACATCTTAGACCTAGACTAA
NK20716
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCTTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCGT
TATAGCAACAGCATTTATAGGTTATGTACTACCATGAGGACAAATATCATTCT
GAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTGCCATTTATTATTACCGCCCTTGTA
TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCCCCAAATAAACTCGGC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCAATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA
NK133563
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTTTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG

165

TTATAGCAACAGCATTTATAGGTTATG-------
TAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTCGGAGACCCAGACA
ATTATACCCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATG
GTATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCTAACAAACTAGGCG
GAGTCCTAGCACTAATCCTATCTATTTTTATCCTAGCCCTCATACCTTTCCTCC
ATACCTCAAAACAACGAACACTCGCCTTCCGTCCAATTACACAAACCATATA
CTGAATCTTAGTAGCTGATCTTCTCATCCTCACATGAATCGGTGGTCAACCAG
TTGAATACCCATTTGTCATCATTGGACAAACAGCTTCAATTGCCTACTTCGCT
ATCATCGTAATCTTCATACCAATTGCAGGCATAATCGAAAATAACATCTTAGA
CCTAGACTAA
NK20053

ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTCTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGGAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAGAGCCACCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCATTAATTATTACCGCCCTTGTA
CTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGCCAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCTATTGCA
GGCATAATCGAAAATAACATTTTAGACCTAGACTAA
NK20108

CGAGACGTAAACTACGGATGATTAATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGTGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTACCGCCCTTGTA
TTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATCCCTTTCCACCCTTACTACACAGTCAAAGA
CTTCCTAGGGATTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTTGGAGACCCTGATAATTACACTCCTGCAAATCCACTCAACA

166

CTCCACCACACATTAAGCCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTA
CGATCCATCCCTAACAAGCTTGGCGGCGTACTAGCATTAATCTTATCAATCCT
AATCCTAGCTCTCATGCCCTTCCTCCATACCTCAAAACAACGATCACTCACCT
TTCGCCCAATTACCCAAACAATATACTGAATCCTAGTAGCGGACCTCCTCATC
CTTACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGAC
AAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCA
GGTATAATCGAAAACAACATCCTAGACCTAGATTAA
NK20942
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTCTGCCTAATTATCCAAATTCTCACAGGATTATTCTTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCA
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTACCGCCCTTGTA
CTAGTCCACCTACTATTCCTACACGAAACAGGRTCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTACTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCTATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA
NK20929
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTAGGATCC
CTATTAGGCCTCTGCCTAATTATCCAAATTCTCACAGGATTATTCTTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTACCGCCCTTGTA
CTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC

ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTTCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTACTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCTATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA

NK20941

ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTCTGCCTAATTATCCAAATTCTCACAGGATTATTCTTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCA
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTACCGCCCTTGTA
CTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCTATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA

NK20930

ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTAGGATCC
CTATTAGGCCTCTGCCTAATTATCCAAATTCTCACAGGATTATTCTTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTACCGCCCTTGTA
CTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC

ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTTCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTACTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCTATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA
NK20940
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTCTGCCTAATTATCCAAATTCTCACAGGATTATTCTTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTCGCCA
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAAGCCACTCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCATTTATTATTACCGCCCTTGTA
CTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGACAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC
ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCTATTGCA
GGCATAATCGAAAATAACATCTTAGACCTAGACTAA
NK20053
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATTATCAACCACTCAT
TCATCGACCTCCCGACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTATTAGGCCTCTGCCTAATTATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCACACATCTGC
CGAGACGTAAACTACGGATGACTTATCCGATACTTACATGCCAACGGCGCTT
CCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGAATTTACTACGGC
TCCTACAACATAATCGAAACATGGAACATAGGAATCATCCTACTATTCGCCG
TTATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGACAAATATCATTC
TGAGGCGCCACAGTAATTACAAACCTACTATCAGCTATCCCATATGTCGGTAC
AACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGACAAGAGCCACCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCATTAATTATTACCGCCCTTGTA
CTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGTT
TGAACTCAGACGCAGCCAAAATTCCTTTCCACCCCTACTACACTGTCAAAGA
CTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTCGGAGACCCAGACAATTATACCCCTGCAAATCCACTCAAC

ACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTACGCCATCTT
ACGATCCATCCCTAACAAACTAGGCGGAGTCCTAGCACTAATCCTATCTATCT
TTATCCTAGCCCTCATACCTTTCCTCCATACCTCAAAACAACGAACACTCGCC
TTCCGTCCAATTACACAAACCATATACTGAATCTTAGTAGCTGATCTTCTCAT
CCTCACATGAATCGGTGGTCAACCAGTTGAATACCCATTTGTCATCATTGGAC
AAACAGCTTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCTATTGCA
GGCATAATCGAAAATAACATTTTAGACCTAGACTAA
MSUM10542

GTAGCCCACATCTGCCGAGACGTAAATTATGGATGACTCATCCGATATCTAC
ATGCCAACGGAGCTTCCATATTCTTCATTTGCCTGTTCATACACGTAGGACGA
GGAATTTATTATGGCTCCTATAACATAATCGAAACATGAAACATAGGAATCA
TCCTACTATTCGCCGTAATAGCAACAGCATTTATAGGTTATGTATTACCATGA
GGACAGATATCATTCTGAGGAGCCACAGTAATTACAAACCTACTATCAGCTA
TCCCATATATCGGTACAACACTAGTAGAGTGAATTTGAGGCGGCTTCTCAGTA
GACAAAGCCACTCTCACACGATTCTTTGCTTTCCACTTCATCCTTCCCTTTATT
ATTACTGCCCTCGTATTAGTTCACCTACTATTCCTACACGAAACAGGATCAAA
CAACCCAACCGGCTTAAACTCAGACGCAGACAAAATCCCATTCCACCCCTAC
TACACAATCAAAGATCT-----------------------
MSUM10540

CCTCCGTAGCCCACATCTGCCGAGACGTAAATTATGGATGACTCATCCGATAT
CTACATGCCAACGGAGCTTCCATATTCTTCATTTGCCTGTTCATTCACGTAGG
ACGAGGAATTTATTACGGCTCTTATAACATAATCGAAACATGAAACATAGGA
ATCATCCTACTATTCGCCGTAATAGCAACAGCATTCATAGGTTATGTACTACC
ATGAGGACAGATATCATTCTGAGGAGCCACAGTAATTACAAACCTGCTATCA
GCTATCCCATATATCGGTACAACACTAGTAGAGTGAATCTGAGGTGGCTTCTC
AGTAGACAAAGCCACTCTCACACGATTCTTTGCTTTCCACTTCATCCTTCCCTT
TATTATTACCGCCCTCGTATTAGTTCACCTACTATTCCTACACGAAACAGGAT
CAAACAACCCAACCGG-
MSUM10547

CACCACAATAGCAACAGCATTCATAGGTTATGTACTACCATGAGGACAGATA
TCATTCTGAGGAGCCACAGTAATTACAAACCTGCTATCAGCTATCCCATATAT
CGGTACAACACTAGTAGAGTGAATCTGAGGTGGCTTCTCAGTAGACAAAGCC
ACTCTCACACGATTCTTTGCTTTCCACTTCATCCTACCCTTTATTATTACCGCC
CTCGTATTAGTTCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAA
CCGGTTTAAACTCAGACGCAGACAAAATCCCGTTCCACCCTTACTACACAGT
CAAAGACTTTCTAGGCGTCCTATTCCTATTGATAGCTTTCATAATTTTAACTTT
ATTTTTCCCAGATATTCTCGGAGACCCAGATAATTACACCCCTGCAAATCCAC

TCAACACTCCACCACACATTAAACCAGAATGATATTTCCTATTAA----------------
----
MSUM16368

CCTCCGTAGCCCACATCTGCCGAGACGTAAACTACGGATGACTTATCCGATA
CTTACATGCCAACGGCGCTTCCATATTCTTCATCTGCCTATTCCTTCATGTAGG
ACGAGGAATTTACTACGGCTCCTACAACATAATCGAAACATGAAACATAGGA
ATCATCCTACTATTCGCCGTTATAGCAACAGCATTTATAGGTTATGTACTACC
ATGAGGACAAATATCATTCTGAGGCGCCACAGTAATTACAAACCTACTATCA
GCTATCCCATATGTCGGTACAACACTAGTAGAGTGGATTTGAGGCGGTTTCTC
AGTAGACAAAGCCACTCTCACACGATTCTTTGCTTTCCACTTCATCCTGCCAT
TTATTATTACCGCCCTTGTATTAGTCCACCTACTATTCCTACACGAAACAGGA
TCAAACAACCCAACCGGTTTAAACTCAGACGCAGACAAAATCCCATTCCACC
CCTACTACACAGTCAAAGACTTTCTAGGAGTCCTATTCCTATTGATAGCTTTC
ATAATTTTAACTTTATTTTTCCCAGATATTCTCGGAGACCCAGATAATTATAC
CCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGATATTTCC
TATTAG---

MSUM33286

GCACACATCTGCCGAGACGTAAACTACGGATGACTTATCCGATACTTACATG
CCAACGGAGCTTCCATATTCTTCATTTGCCTATTCCTTCATGTAGGACGAGGA
ATTTACTACGGCTCCTACAACATAATCGAAACATGAAACATAGGAATCATCC
TACTATTCGCCGTAATAGCAACAGCATTTATAGGTTATGTACTACCATGGGGA
CAAATATCATTCTGAGGAGCCACAGTAATTACAAACCTACTATCAGCTATCCC
ATATATCGGTACAACACTAGTAGAGTGGATTTGAGGCGGTTTCTCAGTAGAC
AAAGCCACTCTCACACGATTCTTCGCTTTCCACTTCATCCTGCCCTTTATTATT
ACCGCCCTAGTATTAGTCCACCTACTATTCCTACACGAAACAGGATCAAACA
ACCCAACCGGTTTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTA
CACAGTCAAAGACTTTCTAGGAGTCCTATTACTATTGATAGCTTTCATAATTT
TAACTTTATTTTTCCCAGATATTCTCGGAGACCCAGATAATTATACTCCTGCA
AATCCACTCAACACTCCACCACACATTAAACCAGAATGATATTTCCTATTTG--

MSUM12712

TCCTCCGTAGCCCACATCTGCCGAGACGTAAACTACGGATGACTAATCCGAT
ACTTACATGCCAACGGAGCTTCCATATTCTTCATTTGCCTATTCCTTCATGTAG
GACGAGGAATTTACTACGGCTCCTACAACATAATCGAAACATGAAACATAGG
AATCATCCTACTATTCGCCGTTATAGCAACAGCATTTATAGGTTATGTACTAC
CATGAGGACAAATATCATTCTGAGGCGCCACAGTAATTACAAACCTACTATC
AGCTATCCCATATGTCGGTACAACACTAGTAGAGTGGATTTGAGGAGGTTTCT
CAGTAGATAAAGCCACTCTAACACGATTCTTCGCCTTCCACTTCATCCTGCCT
TTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACACGAAACAGG
ATCAAACAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCAC
CCTTACTACACAGTCAAAGACTTTCTAGGAGTTCTCATCCTATTGATAGCTTT

CATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCAGATAATTATAC
CCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGATATTTCC
TATT--------------------
MSUM16367
CATGCCAACGGAGCTTCCATATTCTTCATTTGCCTATTCCTTCATGTAGGACG
AGGAATTTACTACGGCTCCTACAACATAATCGAAACATGAAACATAGGAATC
ATCCTACTATTCGCCGTAATAGCAACAGCATTCATAGGTTATGTACTCCCATG
AGGACAGATATCATTCTGAGGGGCCACAGTAATTACAAACCTACTATCAGCT
ATCCCCTATATCGGCACAACACTAGTAGAGTGAATCTGAGGAGGTTTCTCAG
TAGACAAAGCCACTCTCACACGATTCTTTGCTTTCCACTTCATCCTGCCTTTTA
TTATTACCGCCCTAGTATTAGTCCACCTTCTATTCCTACACGAAACAGGATCA
AACAACCCAACCGGTTTAAACTCAGACGCAGACAAAATCCCATTCCACCCCT
ACTACACAGTCAAAGACTTTCTAGGAGTCCTAATCCTATTGATAGCTTTCATA
ATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACCCCT
GCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGATATTTCCTATT
TGG
NK09524
ATGACAATCATCCGAAAAAAACACCCGCTAATCAAAATCATCAACCACTCAT
TCATCGACCTCCCAACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTGCTAGGCCTTTGCCTAATCATCCAAATCCTCACAGGATTATTCCTAGCCAT
ACACTACACATCAGATACAGCAACGGCATTCTCATCAGTAGCTCACATCTGC
CGAGATGTAAACTATGGATGACTCATCCGATATATGCATGCCAACGGAGCCT
CCATATTCTTCATCTGCCTGTTCCTACACGTAGGACGAGGAATTTATTACGGC
TCCTATAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTTGCCGT
AATAGCAACAGCATTCATAGGTTATGTACTACCATGAGGACAAATATCATTC
TGAGGGGCCACAGTAATTACAAACCTACTATCAGCTATCCCATACATCGGTA
CAACACTAGTAGAGTGGATATGAGGCGGCTTCTCAGTAGACAAAGCTACCCT
CACACGATTTTTTGCTTTCCACTTCATCCTACCCTTTATCATTACTGCCCTMGT
ATTAGTTCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGC
TTAAACTCKGATGCAGACAAAATCCCATTCCACCCCTACTACACTTTCAAAG
ACTTTTTAGGCGTCCTATTACTATTAATAGCTTTCATAATTTTAACTTTATTTT
TCCCAGATATTCTCGGAGACCCAGATAATTACACCCCTGCAAATCCACTCAA
CACTCCACCACACATTAAGCCAGAATGGTACTTCCTCTTCGCCTACGCCATCT
TACGATCCATCCCTAATAAACTAGGCGGTGTTCTAGCACTAATCCTATCAATC
TTAATCCTAGCTCTCATACCACTACTCCACACCTCAAAACAACAAGCACTCAC
TTTCCGTCCAATCACACAAACCATATACTGAATCCTAGTAGCCGATCTCCTAA
TTCTCACATGAATCGGTGGCCAACCAGTCGAATACCCATTCATCATCATCGGA
CAAACAGCCTCAATTGCCTACTTCGCTATCATCGTAATTTTCATACCAATCGC
AGGCATAATCGAAAACAACATCTTAAACCTAGACTAA
NK09541
ATGACAATCATCCGAAAAAAACACCCGCTAATCAAAATCATCAACCACTCAT
TCATCGACCTCCCAACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTGCTAGGCCTTTGCCTAATCATCCAAATCCTCACAGGATTATTCCTAGCCAT
ACACTACACATCAGATACAGCAACGGCATTCTCATCAGTAGCTCACATCTGC
CGAGATGTAAACTATGGATGACTCATCCGATATATGCATGCCAACGGAGCTT
CCATATTCTTCATCTGCCTGTTCCTACACGTAGGACGAGGAATTTATTACGGC

TCCTATAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTTGCCGT
AATAGCAACAGCATTCATAGGTTATGTACTACCATGAGGACAAATATCATTC
TGAGGGGCCACAGTAATTACAAACCTACTATCAGCTATCCCATACATCGGTA
CAACACTAGTAGAGTGGATCTGAGGCGGCTTCTCAGTAGACAAAGCTACCCT
CACACGATTTTTTGCTTTCCACTTCATCCTACCCTTTATCATTACTGCCCTAGT
ATTAGTTCACCTACTATTCCTACACGAAACAGGATCAAACAACCCAACCGGC
TTAAACTCGGATGCAGACAAAATCCCATTCCACCCCTACTACACATTCAAAG
ACTTTTTAGGCGTCCTATCTATATTAATAGCTTTCATAATTTTAACTTTATTTT
TCCCAGATATTCTCGGAGACCCAGATAATTACACCCCTGCAAATCCACTCAA
CACTCCACCACACATTAAGCCAGAATGGTACTTCCTCTTCGCCTACGCCATCT
TACGATCCATCCCTAATAAACTAGGCGGTGTTCTAGCACTAATCCTATCAATC
TTAATCCTAGCTCTCATACCACTACTCCACACCTCAAAACAACAAGCACTCAC
TTTCCGTCCAATCACACAAACCATATACTGAATCCTAGTAGCCGATCTCCTAA
TTCTCACATGAATCGGTGGCCAACCAGTCGAATACCCATTCATCATCATCGGA
CAAACAGCCTCAATTGCCTACTTCGCTATCATCGTAATTTTCATACCAATCGC
AGGCATAATCGAAACAACATCTTAAACCTAGACTAA
NK09543
ATGACAATCATCCGAAAAAAACACCCGCTAATCAAAATCATCAACCACTCAT
TCATCGACCTCCCAACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTGCTAGGCCTTTGCCTAATCATCCAAATCCTCACAGGATTATTCCTAGCCAT
ACACTACACATCAGATACAGCAACGGCATTCTCATCAGTAGCTCACATCTGC
CGAGATGTAAACTATGGATGACTCATCCGATATATGCATGCCAACGGAGCTT
CCATATTCTTCATCTGCCTATTCCTACACGTAGGACGAGGAATTTATTACGGC
TCCTATAACATAATCGAAACATGAAACATAGGAATCATCTTACTATTTGCCGT
AATAGCAACAGCATTCATAGGTTATGTACTACCATGAGGACAAATATCATTC
NK9544
ATGACAATCATCCGAAAAAAACACCCGCTAATCAAAATCATCAACCACTCAT
TCATCGACCTCCCAACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTGCTAGGCCTTTGCCTAATCATCCAAATCCTCACAGGATTATTCCTAGCCAT
ACACTACACATCAGATACAGCAACGGCATTCTCATCAGTAGCTCACATCTGC
CGAGATGTAAACTATGGATGACTCATCCGATATATGCATGCCAACGGAGCCT
CCATATTCTTCATCTGCCTGTTCCTACACGTAGGACGAGGAATTTATTACGGC
TCCTATAACATAATCGAAACATGAAACATAGGAATCATCCTACTATTTGCCGT
AATAGCAACAGCATTCATAGGTTATGTACTACCATGAGGACAAATA
NK9708
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATCATTAACCACTCAT
TCATCGATCTCCCAACCCCATCAAACATCTCATCATGATGAAACTTCGGATCT
CTGCTAGGCCTCTGCCTAGTCATCCAAATTCTCACAGGATTATTCCTAGCCAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCTCACATCTGC
CGAGATGTAAACTATGGATGACTCATCCGATATATGCATGCCAACGGAGCTT
CCATATTCTTCATCTGCCTGTTCCTACACGTAGGACGAGGGATTTATTACGGC
TCCTATAACATAATCGAAACATGAAACATAGGAATCATCCTGCTATTCGCCGT
AATAGCAACAGCATTCATAGGTTATGTATTACCATGAGGACAGATATCATTTT
GAGGAGCCACAGTAATTACAAACCTGCTATCAGCTATCCCATATATCGGTAC
AACACTAGTAGAATGAATCTGAGGTGGCTTCTCAGTAGACAAAGCCACCCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCCTTTATCATTACTGCCCTCGTA

TTAGTTCACCTGTTATTTCTACACGAAACAGGATCGAACAACCCAACCGGHTT
AAACTCAGATGCAGACAAAATCCCGTTCCACCCCTACTACACAGTCAAAGAC
TTTTTAGGCGTCCTATTACTATTGATAGCTTTCATCAATTTAACTTTATTTTTC
CCCGATATTCTCGGAGCCCCAGATAATTACACCCCTGCAAATCCACTCAACA
CCACACCACGCATTAAGCCAGAATGGTATTTCCTCTTCGCCTACGCCATCTTA
CGATCCATCCCTAACAAACTAGGCGGTGTTCTAGCACTAATCCTATCAATCTT
AATCCTAGCTCTCATACCACTACTTCACACCTCAAAACAACGAGCACTCACTT
TCCGTCCAATCACACAAACTATATACTGAATCCTAGTAGCAGATCTCCTCATT
CTCACATGAATCGGCGGCCAACCAGTCGAATATCCATTCATCATAATCGGAC
AAGCAGCCTCAATTGCCTACTTCACCATCATCGTAATTCTTATACCAATCGCA
GGTATAATCGAAACAACATCTTAAACCTAGACTAA
NK9707
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATCATTAACCACTCAT
TCATCGATCTCCCAACCCCATCAAACATCTCATCATGATGAAACTTCGGATCT
CTGCTAGGCCTCTGCCTAGTCATCCAAATTCTCACAGGATTATTCCTAGCCAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCTCACATCTGC
CGAGATGTAAACTATGGATGACTCATCCGATATATGCATGCCAACGGAGCTT
CCATATTCTTCATCTGCCTGTTCCTACACGTAGGACGAGGGATTTATTACGGC
TCCTATAACATAATCGAAACATGAAACATAGGAATCATCCTGCTATTCGCCGT
AATAGCAACAGCATTCATAGGTTATGTATTACCATGAGGACAAATATCATTCT
NK9706
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATCATTAACCACTCAT
TCATCGATCTCCCAACCCCATCAAACATCTCATCATGATGAAACTTCGGATCT
CTGCTAGGCCTCTGCCTAGTCATCCAAATTCTCACAGGATTATTCCTAGCCAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCTCACATCTGC
CGAGATGTAAACTATGGATGACTCATCCGATATATGCATGCCAACGGAGCTT
CCATATTCTTCATCTGCCTGTTCCTACACGTAGGACGAGGGATTTATTACGGC
TCCTATAACATAATCGAAACATGAAACATAGGAATCATCCTGCTATTCGCCGT
AATAGCAACAGCATTCATAGGTTATGTATTACCATGAGGACAGATATCATTTT
GAGGAGCCACAGTAATTACAAACCTGCTATCAGCTATCCCATATATCGGTAC
AACACTAGTAGAATGAATCTGAGGTGGCTTCTCAGTAGACAAAGCCACCCTC
ACACGATTCTTTGCTTTCCACTTCATCCTACCCTTTATCATCACTGCCCTCGTA
TTAGTTCACCTGTTATTTCTACACGAAACCGGATCGAACAACCCAACCGGHTT
AAACTCAGATGCAGACAAAATCCCGTTCCACCCCTACTACACAGTCAAAGAC
TTTTTAGGCGTCCTATTACTATTGATAGCTTTCATCAATTTAACTTTATTTTTC
CCCGATATTCTCGGAGCCCCAGATAATTACACCCCTGCAAATCCACTCAACA
CCACACCACGCATTAAGCCAGAATGGTATTTCCTCTTCGCCTACGCCATCTTA
-
GATCCTTCCCTACCAAACTAGGCGGTGTTCTAGCACTAATCCTATCAATCTTA
ATCCTAGTCCTCATACCACTACTTCACACCTCAAAACAACGAGCACTCACTTT
CCGTCCAATCACACAAACTATATACTGAATCCTAGTAGCAGATCTCCTCATTC
TCACATGAATCGGCGGCCAACCAGTTGAATACCCATTTATCATAATCGGACA
AACAGCCTCAATTGCCTACTTCACCATCATCGTAATTCTTATACCAATCGCAG
GTATAATCGAAACAACATCTTTAAACCTAGACTA
NK9690
ATGACAATTATCCGAAAAAAACACCCATTAATCAAAATCATTAACCACTCAT

TCATCGACCTCCCAACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTGCTAGGCCTTTGCCTACTAATACAAATCCTCACAGGGTTGTTCCTAGCCAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCCCACATCTGC
CGAGATGTAAACTATGGGTGACTCATCCGATATATACATGCCAACGGAGCTT
CCATATTCTTCATCTGCCTGTTCCTACACGTAGGACGAGGAGTTTATTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGAATCATCCTGCTATTCGCCG
TAATAGCAACAGCATTCATAGGTTATGTATTACCATGAGGACAGATATCATTC
TGAGGGGCCACAGTAATTACAAACCTACTATCAGCTATCCCATACATCGGTA
CAACACTAGTAGAGTGGATCTGAGGAGGCTTCTCAGTAGACAAAGCCACTCT
CACACGATTCTTCGCTTTTCACTTCATCCTTCCCTTTATCATTACTGCCCTCGT
ATTAGTTCACCTTCTATTCCTACACGAAACAGGATCAAACAACCCAACCGGC
TTAAACTCGGATGCAGACAAAATCCCATTCCACCCCTACTACACAGTCAAAG
ACTTTTTAGGAGTCCTACTACTATTGATAGTTTTCATAATTTTAACTTTATTTT
TCCCAGATATTCTCGGAGACCCAGATAATTACACCCCTGCAAACCCACTCAA
CACTCCACCACACATCAAGCCAGAATGGTACTTCCTCTTCGCCTACGCCATCT
TACGATCCATCCCTAACAAACTAGGCGGTGTTCTAGCACTAATCCTATCAATC
TTAATCCTAGCCCTCATGCCACTACTCCACACCTCAAAACAACGAGCACTCAC
TTTTCGTCCAGTCACACAAACCATATACTGAATCCTAGTAGCCGATCTCCTCA
TTCTCACATGAATCGGAGGTCAACCAGTAGAGTACCCATTCATCATCATCGG
ACAAACAGCCTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCGATCG
CAGGCATAATCGAAAACAACATCTTAAACCTAGACTAA

NK9695
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATCATTAACCACTCAT
TCATCGACCTCCCAACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTACTAGGCCTTTGCCTAATAATCCAAATTCTCACAGGATTATTCCTAGCTAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCTCACATCTGC
CGAGATGTGAACTATGGGTGACTCATCCGATACATGCATGCCAACGGAGCTT
CCATATTCTTCATCTGCCTGTTCCTACACGTAGGACGAGGCGTCTATTACGGC
TCCTACAACATAATCGAAACATGAAACATAGGGATCATCCTGCTATTCGCCG
TAATAGCAACAGCATTATAGGTTATGTATTACCATGAGGGCAAATATCATTCT
GAGGGGCCACAGTAATTACAAACCTACTATCAGCTATCCCATACATCGGTAC
AACACTAGTAGAATGGATCTGAGGCGGCTTCTCAGTAGACAAAGCCACCCTC
ACACGATTCTTCGCTTTTCACTTCATCCTTCCCTTTATCATTACTGCCCTCGTA
TTAGTTCACCTTCTATTCCTACACGAAACAGGATCAAACAACCCAACCGGCTT
AAACTCGGATGCAGACAAAATCCCATTCCACCCCTACTACACAGTCAAAGAC
TTTTTAGGAGTCCTACTACTATTGATAGTTTTCATAATTTTAACTTTATTTTTC
CCAGATATTCTCGGAGACCCAGATAATTACACCCCTGCAAACCCACTCAACA
CTCCACCACACATCAAGCCAGAATGGTACTTCCTCTTCGCCTACGCCATCTTA
CGATCCATCCCTAACAAACTAGGCGGTGTTCTAGCACTAATCCTATCAATCTT
AATCCTAGCCCTCATGCCACTACTCCACACCTCAAAACAACGAGCACTCACTT
TTCGTCCAGTCACACAAACCATATACTGAATCCTAGTAGCCGATCTCCTCATT
CTCACATGAATCGGAGGTCAACCAGTAGAGTACCCATTCATCATCATCGGAC
AAACAGCCTCAATTGCCTACTTCGCTATCATCGTAATCTTCATACCGATCGCA
GGCATAATCGAAAACAACATCTTAAACCTAGACTAA

NK9688
ATGACAATCATCCGAAAAAAACACCCACTAATCAAAATCATTAACCACTCAT

TCATCGACCTCCCAACCCCATCAAACATCTCATCATGATGAAACTTCGGATCC
CTGCTAGGCCTTTGCCTACTAATACAAATCCTCACAGGGTTGTTCCTAGCCAT
ACACTACACATCAGATACAGCAACAGCATTCTCATCAGTAGCCCACATCTGC
CGAGATGTAAACTATGGTGACTCATCCGATATATACATGCCAACGGAGCTTC
CATATTCTTCATCTGCCTGTTCCTACACGTAGGACGAGGAGTTTATTACGGCT
CCTACAACATAATCGAAACATGAAACATAGGAATCATCCTGCTATTCGCCGT
AATAGCAACAGCATTCATAGGTTATGTATTACCATGAGGACAGATATCATTCT
GAGGGGCCACAGTAATTACAAACCTACTATCAGCTATCCCATACATCGGTAC
AACACTAGTAGAGTGGATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTC
ACACGATTCTTTGCTTTCCACTTCATCCTGCCCTTTATCATTACTGCCCTCGTA
TTAGTTCACCTTCTATTCCTACACGAAACAGGATCAAACAACCCAACCGGCTT
AAACTCGGATGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGAC
TTTTTAGGAGTCCTATTACTAATGATAGCTTTCATAATTTTAACTTTATTTTTC
CCAGATATTCTCGGAGACCCAGATAATTATACCCCTGCAAATCCACTCAACA
CTCCACCACACATTAAGCCAGAATGGTACTTCCTCTTCGCCTACGCCATCTTA
CGATCTATCCCTAACAAACTAGGCGGCGTTCTAGCACTAATCCTATCAATCTT
AATCCTAGCTCTCATACCTCTACTCCACACCTCAAAACAACGAACACTCACTT
TCCGTCCAATCACACAAACTATATACTGAATTCTAGTAGCCGATCTCCTCATT
CTCACATGAATCGGTGGACAACCAGTCGAATACCCGTTCATCGTCATCGGAC
AAACAGCCTCAATTGCCTACTTCGCTATCATCGTAATTTTCATACCAATCGCA
GGAATAATCGAAAACAACATCCTAAACCTAGACTAA

APPENDIX D

Raw data for analyses of the nuclear intron acid phosphatase V (AP5) from 26 samples of
the Mexican vole, *Microtus mexicanus*.

NK20959
TGCAGAGACGTTGCCAAGGTGGTCATGGTTTCCAGCTAGCACGTACCAAGGA
ATATTGCGGAGGGCACGATCAGAGAACACGTCCTCAAAGGTTTCCTGTGACA
GACAGAAAGGGCATCCCACCTCTGGGGTCAACAGCCACCCTGCCGCTGCCAC
CAAATCAGAGACCAAGGACAAAGGAACCACTGGATGTCCAAGGTAGCTACC
TGGAGGAGCCCTTCTGTTTCTCTCCGCCCTCCCCAGAGGAAGGCCCCAGGTCG
GAGTGACCTCCCCTCGGGACCCTCTAACCCTGTTGTTGAGCACATACCTGGAA
CCTCTTGTCATTGGCATCGTGCACTCCAGTGAAGTAGAAGTTGTCCCCTAGGG
ACATGATGAAGTCAGCACCCATAATCTGCACGGTTCTGGCGATTTCTTTAGCA
TTGGCCATTTCCCTGGCTGTGTGGAATGGGGC
NK20257 -------------GCCAAGGTGGTCATGGTTTCCAGCTAGCACGTACCAA---
TATTGCGGAGGGCACGATCAGAGAACACGTCCTCAAAGGTTTCCTGTGACAG
ACAGAAAGGGCATCCCACCTCTGGGGTCAACAGCCACCCTGCCGCTGCCACC
AAATCAGAGACCAAGGACAAAGGAACCACTGGATGTCCAAGGTAGCTACCT
GGAGGAGCCCTTCTGTTTCTCTCCGCCCTCCCCAGAGGAAGGCCCCAGGTCG
GAGTGACCTCCCCTCGGGACCCTCTAACCCTGTTGTTGAGCACATACCTGGAA
CCTCTTGTCATTGGCATCGTGCACTCCAGTGAAGTAGAAGTTGTCCCCTAGGG
ACATGATGAAGTCAGCACCCATAATCTGCACGGTTCTGGCGATTTCTTTAGCA
TTGGCCATTTCCCGGGCTGTGTGGAATGGGGC
NK20958
TGCAGAGACGTTGCCAAGGTGGTCATGGTTTCCAGCTAGCACGTACCAAGGA
ATATTGCGGAGGGCACGATCAGAGAACACGTCCTCAAAGGTTTCCTGTGACA
GACAGAAAGGGCATCCCACCTCTGGGGTCAACAGCCACCCTGCCGCTGCCAC
CAAATCAGAGACCAAGGACAAAGGAACCACTGGATGTCCAAGGTAGCTACC
TGGAGGAGCCCTTCTGTTTCTCTCCGCCCTCCCCAGAGGAAGGCCCCAGGTCG
GAGTGACCTCCCCTCGGGACCCTCTAACCCTGTTGTTGAGCACATACCTGGAA
CCTCTTGTCATTGGCATCGTGCACTCCAGTGAAGTAGAAGTTGTCCCCTAGGG
ACATGATGAAGTCAGCACCCATAATCTGCACGGTTCTGGCGATTTCTTTAGCA
TTGGCCATTTCCCGGGCTGTGTGGAATGGGGC
NK20960
TGCAGAGACGTTGCCAAGGTGGTCATGGTTTCCAGCTAGCACGTACCAAGGA
ATATTGCGGAGGGCACGATCAGAGAACACGTCCTCAAAGGTTTCCTGTGACA
GACAGAAAGGGCATCCCACCTCTGGGGTCAACAGCCACCCTGCCGCTGCCAC
CAAATC--
GGCCCCAGGTCGGAGTGACCTCCCCTCGGGACCCTCTAACCCTGTTGTTGAGC
ACATACCTGGAACCTCTTGTCACTGGCATCGTGCACTCCAGTGAAGTAGAAG
TTGTCC---
NK56518 ---
GTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTT
CACTGGAGTGCACGATGCCAGTGACAAGAGGTTCCAGGTATGTGCTCAACAA

CAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCTCTG
GGGAGAGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACAT
CCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCT
GTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGG
ATGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTGGTACGTGCTAGCTGGA
AACCATGACCACCTTGGC------------
NK29002   ----------------------------
CCAATGCTAAAGAAATCGCCAGAACCGTGCAGATTATGGGTGCTGACTTCAT
CATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCACGATGCCAATGAC
AAGAGGTTCCAGGTAGAGGGTCCCGAGGGGAGGTCGCTCTGACCCGGGGCCT
TCCTCTGGGGAGGGCGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTT
GGACATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGG
GTGGCTGTTGACCCTAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCT
TTGAGGACGTGTTCTCTGACCGTGCCCTCCGCAATATTCCCTGGTACGTGCTA
GCTGGAAACCATGACCACCTTGGCAACGTCTCTGCA--------------------
NK9690
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAA
CCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTAC
TTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAAC
AACAGGGTTAGAGGGTCCAGAGGGGAGGTCACTCCGACCTGGGGCCTTCCTC
TGGAGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGAC
ATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGG
CTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGA
GGACGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTGGTACGTGCTAGCTG
GAAACCATGACCACCTTGGCAACGTCTCTGCA
NK9706
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAA
CCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTAC
TTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAAC
AACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCTC
TGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGAC
ATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGG
CTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGA
GGACGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTGGTAC--------------------
----
NK9707
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAA
CCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTAC
TTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAAC
AACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCTC
TGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGAC
ATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGG
CTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGA
GGACGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTGGTACGTGCTAGCTG
GAAACCATGACCACCTTGGCAACGTCTCTGCA

NK9523
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAA
CCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTAC
TTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAAC
AACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCTC
TGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGAC
ATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGG
CTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGA
GGACGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTGGTACGTGCTAGCTG
GAAACCATGACCACCTTGGCAACGTCTCTGCA
NK20941
GCCCCATTCCACACAGCCAGGGAAATGGCCAATGCTAAAGAAATCGCCAGA
ACCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTA
CTTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAA
CAACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCT
CTGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGA
CATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTG
GCTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTG
AGGACGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTGGTACGTGCTAGCT
GGAAACCATGACCACCTTGGCAACGTCTCTGCC
NK9542   --
CGCCAGAACCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGAC
AACTTCTACTTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTAT
GTGCTCAACAACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGG
GCCTTCCTCTGGAGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTA
CCTTGGACATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGG
CAGGGTGGCTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGA
AACCTTTGAGGACGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTGGTAC---

NK9541
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAA
CCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTAC
TTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAAC
AACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCTC
TGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGAC
ATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGG
CTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGA
GGACGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTG---------------------------
----
NK20020   --
GGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGC
ACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAACAACAGGGTTAGAG
GGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCTCTGGGGAGGGCGGA
GAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATCCAGTGGTTCCT
TTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCCAGA

GGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTG
ATCGTGCCCTCCGCAATATTCCTTGGTACGT--

NK20129

GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAA
CCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTAC
TTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAAC
AACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCTC
TGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGAC
ATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGG
CTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGA
GGACGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTGGTAA----------------------
----

NK9635

GCCCCATTCCACACAGCCAGGGAAATGGCCAATGCTAAAGAAATCGCCAGA
ACCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTA
CTTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAA
CAACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCT
CTGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGA
CATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTG
GCTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTG
AGGACGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTGGTACGTGCTAGCT
GGAAACCATGACCACCTTGGCAACGTCTCTGCA

NK20048

GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAA
CCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTAC
TTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAAC
AACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCTC
TGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGAC
ATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGG
CTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGA
GGACGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTGGTACGTGCTAGCTG
GAAACCATGACCACCTTGGCAACGTCTCTGCA

NK34977

GCCCCATTCCACACAGCCAGGGAAATGGCCAATGCTAAAGAAATCGCCAGA
ACCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTA
CTTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAA
CAACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCT
CTGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGA
CATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTG
GCTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTG
AGGACGTGTTCTCTGATCGTGCCCTCCGCAATATT------------------------------------
----

NK121579

GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAA
CCGTGCAGATTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTAC
TTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTAGGGGTCCAAG

GGAGTCGCTCCGACCCGAGGCCTTCCTCTGGGGAGGGCAGAGAAAAACAGA
AGGGCTCCTCCAGGTAGCTACCTTGGACATCCAGTGGTTCCTTTGTCCTTGGT
CTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCTAGAGGTGGGATGC
CCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGTGCCCT
CCGCAATATTCCTTGGTACG---
NK19605
GCCCCATTCCACACAGCCAGGGAAATGGCCAATGCTAAAGAAATCGCCAGA
ACCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTA
CTTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAA
CAACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCT
CTGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGA
CATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTG
GCTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTG
AGGACGTGTTCTCTGATCGTGCCCTCCGCA---

NK9695
GCCCCATTCCACACAGCCAGGGAAATGGCCAATGCTAAAGAAATCGCCAGA
ACCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTA
CTTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAA
CAACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCT
CTGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGA
CATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTG
GCTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTG
AGGACGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTGGTACGTGCTAGCT
GGAAACCATGACCACCTTGGCAACGTCTCTGCC
NK20108
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAA
CCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTAC
TTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAAC
AACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCTC
TGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGAC
ATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGG
CTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGA
GGACGTGTTCTCTGATCGTGCCCTCCGCAATATCCTG----------------------------------
----
NK9544   -----------------------------CAATGCTAAA-G-A-CGCC-
GAACCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCT-
GGGGACAACTTCTACTTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCC
AGGTATGTGCTCAACAACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGA
CCTGGGGCCTTCCTCTGGGGAGAGCGGAGAGAAACAGAAGGGCTCCTCCAG
GTAGCTACCTTGGACATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGG
CAGCGGCAGGGTGGCTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTC
ACAGGAAACCTTTGAGGATGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTT
GGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCTCTGCA
NK9524   -----------------------------CAATGCTAAAGAAATC-GC--GA-
CCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTGGGGACAACTTCTACT

181

TCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATGTGCTCAACA
ACAGGGTTAGAGGGTCCCGAGGGGAGGTCACTCCGACCTGGGGCCTTCCTCT
GGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACA
TCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGC
TGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAG
GACGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTGGTACGTGCTAGCTGG
AAACCATGACCACCTTGGCAACGTCTCTGCA
MVZ198769
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAA
CCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTAC
TTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATATGCTCAAC
AACAGGGTTAGAGGGTCCAGAGGGGAGGTCACTCCGACCTGGGGCCTTCCTC
TGGGGAGGGCAGAGAGAAACATAAGGGCTCCTCCAGGTAGCAACCTTGGAC
ATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGG
CTGTTGACCCTAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGA
GGACGTGTTCTCTGACCGTGCCCTCCGCAATATTCCTTGGTACGTGCTAGCTG
GAAACCATGACCACCTTGGCAACGTCTCTGCA
MVZ00277
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAA
CCGTGCAGATTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTAC
TTCACTGGAGTGCACGATGCCAATGACAAGAGGTTCCAGGTATATGCTCAAC
AACAGGGTTAGAGGGTCCAGAGGGGAGGTCACTCCGACCTGGGGCCTTCCTC
TGGGGAGGGCAGAGAGAAACAGAAGGGCTCCTCCAGGTAGCAACCTTGGAC
ATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGG
CTGTTGACCCTAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGA
GGACGTGTTCTCTGACCGTGCCCTCCGCAATATTCCTTGGTACGTGCTAGCTG
GAAACCATGACCACCTTGGCAACGTCTCTGCA

APPENDIX E

An ecological niche model that was parameterized with a minimum training presence depicts the distribution of suitable habitat (darker colors) for the Mexican vole, *Microtus mexicanus* based on climate data parameterized for the Last Glacial Maximum (circa 0.21Ma).
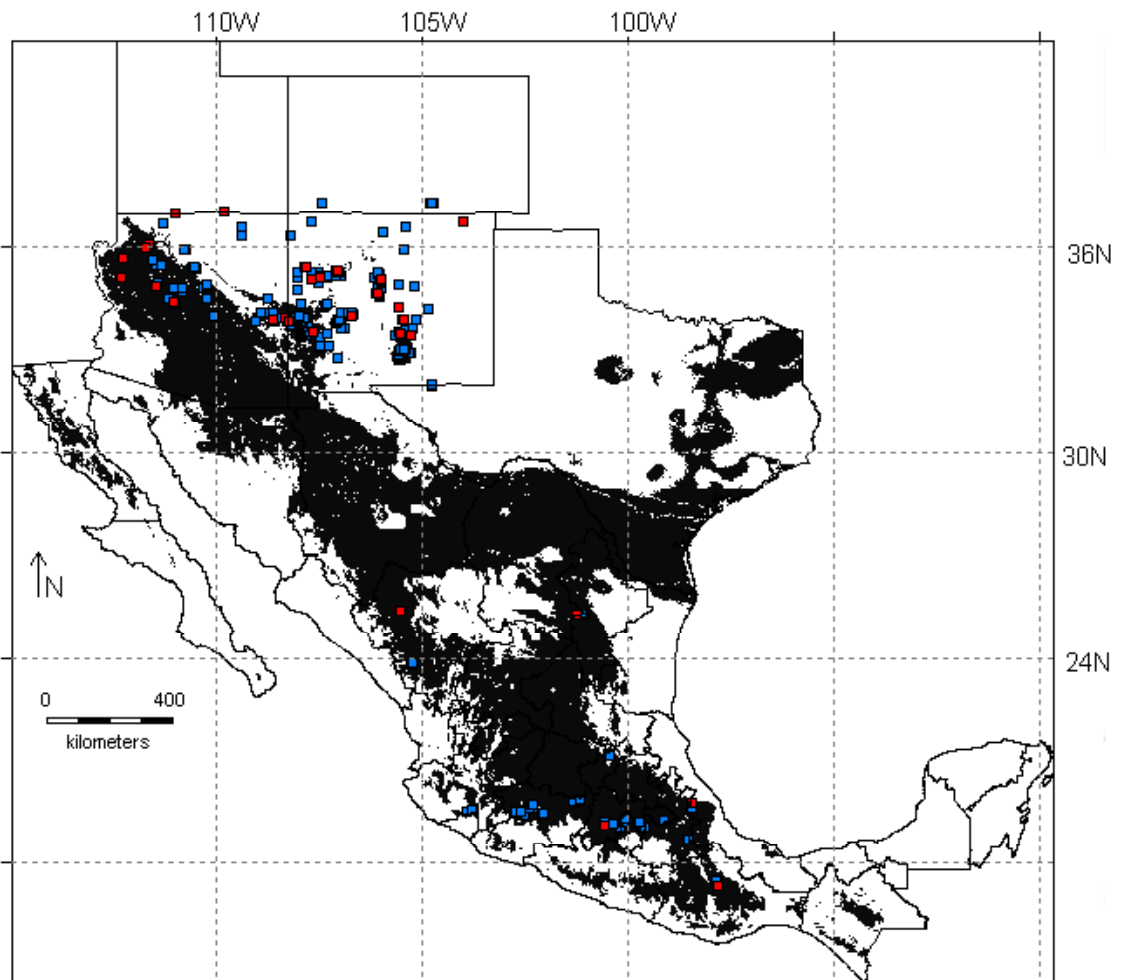
An ecological niche model that was parameterized with an equal training presence
depicts the distribution of suitable habitat (darker colors) for the Mexican vole, *Microtus
mexicanus* based on climate data parameterized for the Last Glacial Maximum
(circa 0.21Ma).
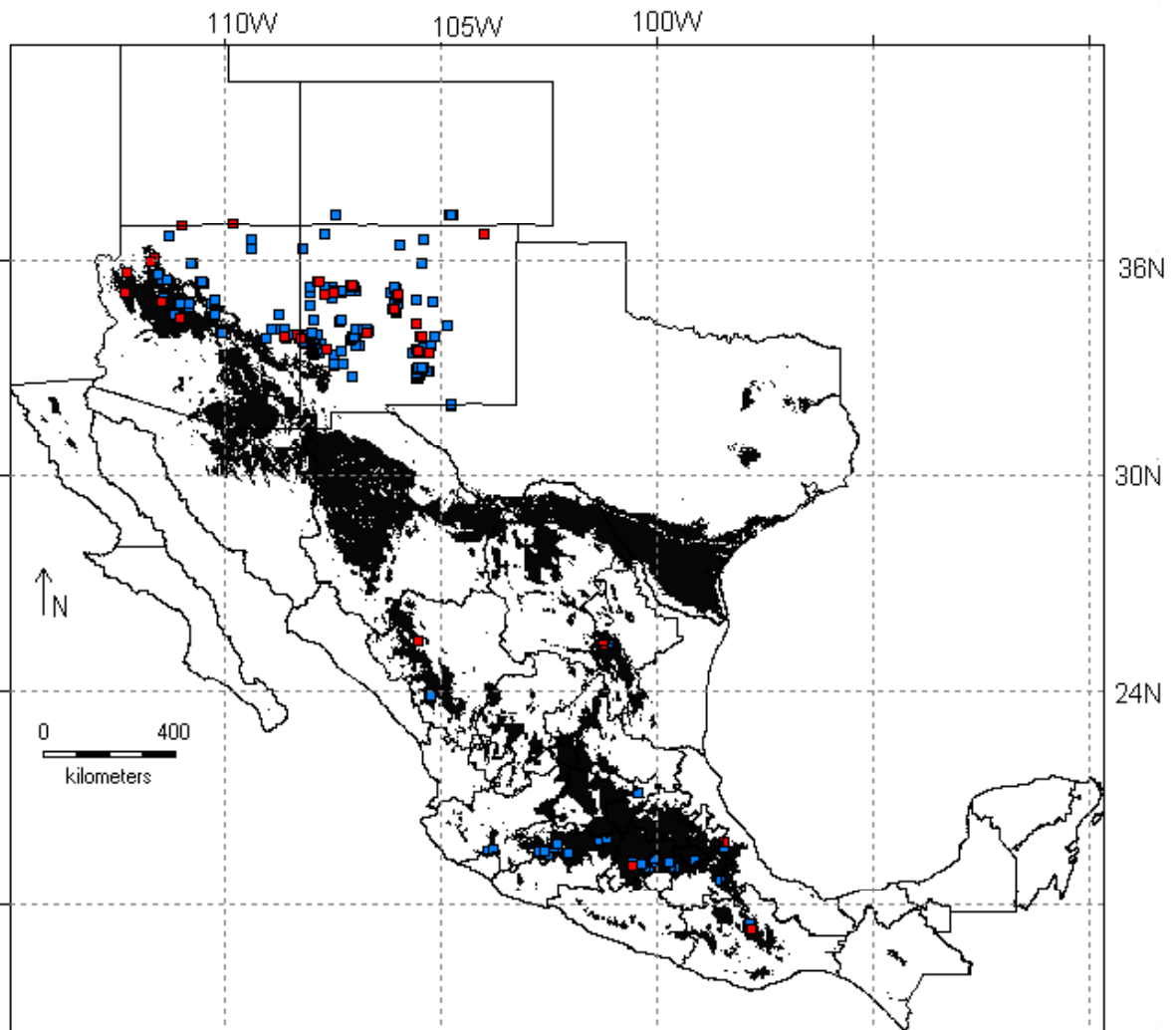
Appendix F
Microtus montanus tissue samples sequenced in the phylogeographic analysis.

Subspecies	Locality	n	Museum ID
amosus	UT: Wasatch Co.	1	MSB NK55568
arizonensis	AZ: Apache Co.	4	MSB NK20111, 20114, 20128, 20107
arizonensis	NM: Catron Co.	2	MSB NK 1916, 29024
canescens	WY: Chelan Co., Jump-Off Ridge	1	Burke 74028
codiensis	WY: Park Co.	4	MSB NK0934, 50940, 51005-51006
dutcheri	CA: Inyo Co.	1	MVZ 219347
fucosus	NV: Lincoln Co., Pahranagat Valley NWR	8	
fusus	CO: Gunnison Co.	2	MSB NK20258, 20260,
micropus	NV: Nye Co., Kirch WMA	4	
micropus	NV: White Pine Co., Steptoe WMA	4	
montanus	CA: Lassen Co., Merril Creek	1	MVZ 217847
montanus	CA: Mono Co.,	3	MSB NK5897, 5899, 5900
nanus	CO: Pitkin Co.	3	MSB NK20244-20246
nevadensis	NV: Nye Co., Ash Meadows	5	MVZ 59379, 59382, 59385, 59388, 59390
rivularis	UT: Salt Lake Co., Camp Williams, Wood Hollow	1	MSB NK 55041
undosus	NV: Churchill Co., Stillwater NWR	4	
nanus	WA: Pend Oreille Co., Deemer Creek	1	Burke 76610
fusus	WA: Grant Co., Beverly	1	Burke 76802
nanus	WA: Kitattas Co., Vantage	1	Burke76828
nanus			
nanus	UT: Washngton Co.,	1	MSB NK36857
yosemite	CA: Mariposa Co., Peregoy Meadow	1	MVZ 201876

Appendix G

The 88 *M. montanu*s samples used in the analysis of cyt*b* variation.

Stllwater183
CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAACAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCTTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
Stllwater184
CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATTCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTTCATCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTAATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
NK001916
CTGAGGGGCTACGGTAATTACAAATCTGCTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATTTGAGGAGGCTTTTCAGTAGATAAAGCCACTCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGATTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCTGATAATTACACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAGCCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGCGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCCCTC
CTCCATACCTCAAAACAACGATCACTCACCTTTCGCCCAATTACCCAGACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
AF119280
CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGTGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCTTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT

186

CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGAATAATCGAAAACAACATCCTAGACCTAGATTAA
NK005897
CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGCAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACACCATCCTAGACCTAGATTAA
Stllwater190
CTGAGGGGCTACAGTTATTACAAATCTACTATCAGCCATTCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTTCATCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
Stllwater186
CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATTCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTTCATCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCTTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
PAH00122
CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATCATCGTAATCTTCATACCAATCG
CAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
PAH00124   ---

GCAGACAAAATCCCATTTCACCCTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTA
TTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATA
CTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCT
ATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCC
TGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCA
ATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGG
CCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCA
TCATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
NK020258
         CTGAGGGGCTACAGTAATTACAAATCTGCTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTTTCAGTAGATAAAGCCACTCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAACAC
TCCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCC
TAACAAGCTTGGCGGTGTACTAGCGTTAATCTTATCAATCCTAATCCTAGCTCTCATGCCTCT
TCTCCATACCTCAAAACAACGAGCACTCACCTTTCGCCCAATTACTCAGACAATATACTGAGT
CCTAGTAGCGGACCTCCTCATCCTTACATGAATCGGAGGACAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCG
CAGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
NK020246
         CTGAGGGGCTACAGTAATTACAAATCTGCTATCAGCTATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTTTCAGTAGATAAAGCCACTCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCTGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACTCAAACAATATACTGAATC
CTAGTAGCGGACCTCTTTATCCTTACATGAATCGGAGGACAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
STEP0015
         CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
PAH00013
         CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACT

CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATCATCGTAATCTTCATACCAATCG
CAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
Stllwater185
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAATGAATCTGAGGGGGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
PAH00121   --

CATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCC
AACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACT
TTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTT
GGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGA
ATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGC
TAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAAC
GAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTT
ATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGC
CTCGATTGCCTACTTCGCCATCATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAA
CATCCTAGACCTAGATTAA
NK020107
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATTTGAGGGGGGCTTTTCAGTAGATAAAGCCACTCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGATTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCCCTC
CTCCATACCTCAAAACAACGAGCACTCACCTTTCGCCCAATTACGCAGACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
PAH10120
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCG
CAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA

PAH00119

CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCG
CAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
NK020128

CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTTTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGATTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAGCCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGATCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
STEP0006

CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTCACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATAAA
KIRCH017

CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTCACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATAAA
SHLNBRG3

CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA

CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
NK20114
        CTGAGGGGCTACAGTAATTACAAATCTGCTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATTTGAGGAGGCTTTTCAGTAGATAAAGCCACTCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CCTACTACACAATCAAAGACTTCCTAGGGATTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAGCCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGCGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCCCTC
CTCCATACCTCAAAACAACGATCACTCACCTTTCGCCCAATTACCCAGACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
NK20111   --

TTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACACGAAACAG
GATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCCTACTAC
ACAATCAAAGACTTCCTAGGGATTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTT
CCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACTCCACCACA
CATTAAGCCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCTAACAAGCT
TGGCGGCGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCCCTCCTCCATAC
CTCAAAACAACGATCACTCACCTTTCGCCCAATTACCCAGACAATATACTGAATCCTAGTAGC
GGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTG
GACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGTATAA
TCGAAAACAACATCCTAGACCTAGATTAA
NK5899
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGCAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
NK020245
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCTATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTTTCAGTAGATAAAGCCACTCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCTGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC

CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACTCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGACAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
NK029024
        CTGAGGGGCTACGGTAATTACAAATCTGCTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATTTGAGGAGGCTTTTCAGTAGATAAAGCCACTCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGATTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCTGATAATTACACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAGCCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGCGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCCCTT
CTCCATACCTCAAAACAACGATCACTCACCTTTCGCCCAATTACCCAGACAATATACTGAATC
CTAGTAGCGGACCTCCTCATCCTTACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
MI077866   --
TGAGGAGGCTTTTCAGTAGATAAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCT
GCCTTTCATCATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAA
CAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCA
AAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGA
TATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAA
ACCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCTAACAAACTTGGCG
GTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTCCTCCATACCTCAA
AACAACGAGCACTCACTTTCCGCCCAATCACGCAAACAATATACTGAGTCCTAGTAGCGGAC
CTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAA
ACGGCCTCAATTGCCTACTTTGCCATCATTGTTATCTTCATACCAATCGCAGGTATAATCGAA
AACAACATCCTAGACCTAGATTAA
MIM77610   ---

ATTCTTCGCCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTA
TTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCC
ATTCCACCCCTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCAT
AATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCA
CTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGA
TCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTC
ATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAAT
ATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAGCCAGTCGAAT
ACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATCATCGTAATCTTCA
TACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
MIM76828   ---

TCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACC
CAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCCTACTACACAGTCAAAGAC
TTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCT
TGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAG
AATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTG
CTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAA
CGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCT
TATCCTCACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCATCATTATTGGACAAGCAG
CCTCGATTGCCTACTTCGCCATCATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACA
ACATCCTAGACCTAGATTAA
MIM76802   --
TAAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCT
AGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTAAACTC

192

AGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCA
TCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAA
TTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCTCTT
CGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATTAATCTTAT
CAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTC
CGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTTACATGAATC
GGAGGCCAGCCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCAATTGCCTACTT
CGCCATCATCGTAATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCCTAGACCTAGA
TTAA
PAH00123
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAACAC
TCCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCC
TAACAAACTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCA
TCATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCG
CAGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
Mpenn
        CTGAGGGGCTACAGTAATTACAAATCTTCTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTGGAGTGAATCTGAGGTGGCTTCTCAGTAGATAAAGCTACCCTAACACGATTCTTCG
CCTTCCACTTCATTTTACCTTTCATTATTACTGCCCTAGTATTAGTTCACCTTCTATTCCTACA
CGAAACAGGATCAAATAATCCCACCGGTCTAAACTCAGATGCAGACAAAATTCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTTATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCTGATAATTACACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAGTGGTACTTCCTATTCGCCTATGCCATCTTACGATCTATTCCT
AACAAACTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCCCTTATACCATTT
CTCCATACTTCAAAACAACGAGCACTTACTTTCCGTCCAATTACGCAAACAATATACTGAATT
CTAGTAGCTGACCTCCTTATCCTCACATGAATCGGCGGTCAACCAGTCGAATACCCATTTAT
CATCATCGGACAAGCAGCCTCAATTGCCTACTTCGCCATCATCGTTATCTTTATGCCAATCGC
AGGTATAATCGAAATAACATCCTAGACCTAGATTAA
MIM51005
        CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGTGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAACACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGAATAGTCGAAAACAACATCCTAGACCTAGATTAA
MIM51006
        CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT

AACAAGCTTGGTGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAACACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGAATAGTCGAAAACAACATCCTAGACCTAGATTAA
MIM50940
        CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGTGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAACACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGAATAGTCGAAAACAACATCCTAGACCTAGATTAA
MIM50934
        CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGTGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAACACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGAATAGTCGAAAACAACATCCTAGACCTAGATTAA
Stllwater187
        CTGAGGGGCTACAGTTATTACAAATCTACTATCAGCCATTCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTTCATCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
NK020244
        CTGAGGGGCTACAGTAATTACAAATCTGCTATCAGCTATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTTTCAGTAGATAAAGCCACTCTAACACGATTCTTTG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAATCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCTGATAATTACACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGATTTTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCTCTT
CTCCATACCTCAAAACAACGAGCACTCACCTTTCGCCCAATCACTCAGACAATATACTGAGT
CCTAGTAGCGGACCTCCTCATCCTTACATGAATCGGAGGACAACCAGTCGAATACCCATTCA

TCATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCG
CAGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
MI201876
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGCAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
MI217847
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAACAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
SHLLNBR2
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
STEPTO30
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
KIRCH15
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC

ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
DC000191
        CTGAGGGGCTACAGTTATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTTCATCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATAAA
DC000192
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
KEYP0193
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATCATCGTAATCTTCATACCAATCG
CAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
KeyP0194   --
AAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTA
GTATTAGTTCATCTTTTATTCCTACACGAAACAGGATCAAATTACCCAACTGGTCTAAACTCA
GACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCAT
CCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAAT
TACACTCCTGCAAATCCACTCAATACTCCACCACACATTAAACCAGAATGATATTTCCTCTTC

196

GCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTACTAGCATTAATCTTATCA
ATCCTAATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCCG
CCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATCG
GAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAGGCAGCCTCAATTGCCTACTTT
GCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCCTAGACCTAGAT
TAA
KeyP0198   ---

TTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACCG
GTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTTCTAG
GGGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGA
CCCCGATAATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTA
TTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCAT
TAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCAC
TCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCA
CATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATT
GCCTACTTCGCCATCATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTA
GACCTAGATAAA
KeyP0197   ---

CTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTAAAC
TCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTTCTAGGGGGTTCT
CATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGAT
AATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCTC
TTCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATTAATCTT
ATCAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCT
TCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAA
TCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTAC
TTTGCCATCATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCTA
GATAAA
KeyP0196
         CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC

KeyP0199   ---

TTATTACCGCCCTAGTATTAGTTCACCTTTTATTTCTACACGAAACAGGATCAAATAACCCAAC
CGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCC
TAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGA
GACCCCGATAATTACACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATG
ATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTACTAG
CATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTCCTCCATACCTCAAAACAACGAG
CACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATC
CTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTC
AATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACAT
CCTAGACCTAGATAAA
MVZ59382
         CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA

CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCTTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
MVZ59385   --

TGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAA
TAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACCCTTACTACACAGTCA
AAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGA
TATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAACACTCCACCACACATTAA
ACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGCG
GTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCTCTTCTCCA-----------------
----

MVZ59388
        CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAACAC
TCCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCC
TAACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCTCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCTTCATCCTCACATGAATCGGAGGACAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCG
CAGGTATAATCGAAAACAACATCCTAGACCTAGATAAA
MVZ59390
        CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCTCTT
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAGACAATATACTGAAT
CCTAGTAGCGGACCTCTTCATCCTTACATGAATCGGAGGACAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCG
CAGGAATAATCGAAAACAACATCCTAGACCTAGATAAA
MVZ59379
        CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCC--

ATCCATCCCTAACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTGATCCTAGCTCT
CATGCCACTTCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAGACAA
TATACTGAATCCTAGTAGCGGACCTCTTCATCCTTACATGAATCGGAGGACAACCAGTCGAA
TACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTC
ATACCAATCGCAGGTATAATCGAAAACAACATCCTAGACCTAGATAAA

KeyP0201

CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACAAGATTCATCG
CATTCCATCCTA-----
CCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATA
ACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAA
GACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATA
TTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAAC
CAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGT
GTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAA
CAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCT
CCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAG
CAGCCTCGATTGCCTACTTCGCCATCATCGTAATCTTTCATACCAATCGCAGGCATAATCGAA
ACAACATCCTAGACCTAGATTTA

KeyP195

CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACAAGATTCATCG
CATTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCC------------------------------------
----
ATCAATCTTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTT
CCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAAT
CGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACT
TTGCCATCATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCTAG
ATAAA

PAH00118

CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA

103589   --

GCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAAT
AACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACCCTTACTACACAGTCAA
AGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGAT
ATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAA
CCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGTGGT
GTGCTAGCATTAATCTTATCAATCCTGATCCTAGCTCTCATGCCACTCCTCCATACCTCAAAA
CAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCT
CTTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAG
CAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGAATAATCGAAA
ACAACATCCTAGACCTAGATTAA

103761MLong   --

TACCTTTCATTATTACTGCCCTAGTACTAGTTCACCTGTTATTCCTACACGAAACAGGATCAA
ACAACCCAACTGGTCTTAACTCAGACGCAGACAAAATTCCATTTCATCCTTACTTCACAGTCA

AAGATCTTTTAGGGGTCCTCATCCTATTGATAGTTTTCATAATTTTGACTTTATTTTTCCCAGA
TATTCTCGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACTCCACCACACATTAA
ACCAGAATGATATTTTCTCTTCGCATATGCTATCTTACGATCTATCCCCAACAAACTAGGTGG
TGTCCTAGCATTAATCCTATCTATCCTAATCCTAGCCCTCATACCACTACTTCACACCTCAAA
ACAACGAGCACTCACCTTCCGCCCAATCACACAAACAATATACTGAATTCTAGTAGCTGATCT
CCTCATCCTCACATGAATTGGCGGCCAACCAGTCGAATACCCATTCATCATTATTGGTCAAG
CAGCTTCAATTGCCTACTTCACTATTATTGTAATCCTTATACCAATCGCAGGTATGCTTGAAAA
CAACATCCTACACCTAGATTAA

NK20274   ---

TTCCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACACGAAACAGGATCAAA
TAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACCCTTACTACACAGTCAA
AGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGAT
ATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAACACTCCACCACACATTAAA
CCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGT
GTACTAGCGTTAATCTTATCAATCCTAATCCTAGCTCTCATGCCTCTTCTCCATACCTCAAAA
CAACGAGCACTCACCTTTCGCCCAATTACTCAGACAATATACTGAGTCCTAGTAGCGGACCT
CCTCATCCTTACATGAATCGGAGGACAACCAGTCGAATACCCATTCATCATTATTGGACAAG
CAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAA
ACAACATCCTAGACCTAGATTAA

55568   ---
TAAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCT
AGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAATCCAACCGGTCTAAACTC
AGACGCAGACAAAATCCCATTTCACCCTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCA
TCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAA
TTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGATATTTCCTCTT
CGCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGTGGTGTGCTAGCATTAATCTTATC
AATCCTGATCCTAGCTCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCC
GCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATC
GGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACTT
TGCCATCATCGTTATCTTCATACCAATCGCAGGAATAATCGAAAACAACATCCTAGACCTAGA
TTAA

KIRCH11   --
TAAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCT
AGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTAAACTC
AGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCA
TCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAA
TTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCTCTT
CGCCTATGCCATCCTACGATCTATCCCTAACAAACTTGGCGGTGTGCTAGCATTAATCTTATC
AATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCC
GCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTTACATGAATCG
GAGGCCAGCCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACTTC
GCCATTATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCTAGAT
TAAA

656174F   ---
TTTAAAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTGCCTTTCATTATTACCGC
CCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTAAA
CTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTTCTAGGGGTTC
TCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGA
TAATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCT
CTTCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATTAATCT
TATCAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACC
TTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGA
ATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTA
CTTCGCCATCATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCT
AGATTAAA

656174M   --
CCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTAT
TAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACG
CAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTAT
TGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATAC
TCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTA
TGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCT
GATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAA
TTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGC
CAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCAT
CATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
666194   --

ATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACC
GGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCT
AGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGA
GACCCCGATAATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGA
TATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGC
ATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTTCTCCATACCTCAAAACAACGAGC
ACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTCATCC
TTACATGAATCGGAGGACAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCG
ATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATC
CTAGACCTAGATTAA
667696   --

TATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATTAAATAACCCAACT
GGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCT
AGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGA
GACCCCGATAATTACACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATG
ATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCTAACAAACTTGGCGGTGTACTAGC
ATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTCCTCCATACCTCAAAACAACGAGC
ACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCAGACCTCCTTATCCT
CACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAGGCAGCCTCAA
TTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCC
TAGACCTAGATTAA
65643HQ3
          CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTCACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAACAC
TCCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCC
TAACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCG
CAGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
NK9857   --

TTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACTG
GTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCTA
GGGATTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAG
ACCCTGATAATTACACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGT
ATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGCGTACTAGCA
TTAATCTTATCAATCCTAATCCTAGCTCTCATGCCCCTTCTCCATACCTCAAAACAACGATCA

201

CTCACCTTTCGCCCAATTACCCAGACAATATACTGAATCCTAGTAGCGGACCTCCTCATCCTT
ACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGAT
TGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCCT
AGACCTAGATTAA
1727   --

ATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACC
GGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCT
AGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGA
GACCCCGATAATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGA
TATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGC
ATTAATCTTATCAATCCTGATCCTAGCTCTCATGCCTCTTCTCCATACCTCAAAACAACGAGC
ACTCACCTTTCGCCCAATTACTCAGACAATATACTGAGTCCTAGTAGCGGACCTCTTCATCCT
TACATGAATCGGAGGACAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGA
TTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCC
TAGACCTAGATTAA
STEP002   ---
TAAAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTGCCTTTCATTATTACCGCCC
TAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTAAACT
CAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTTCTAGGGGTTCTC
ATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATA
ATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCTCT
TCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATTAATCTTA
TCAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTT
CCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTTACATGAAT
CGGAGGCCAGCCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACT
TCGCCATCATCGTTATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCTAG
ATTAA
DC125   CTGAGGGGCTACAGT---

GAATCCTAGTAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCA
TTCATCATTATTGGACAGGCAGCCTCAATTGCCTATTTTGCCATCATCGTTATCTTCATACCA
ATCGCAGGAATAATCGAAAACAACATCCTAGACCTAGATTAA
MIMO5   --

CAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCTAGGAGTTCTC
ATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATA
ATTACACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGATATTTCCTCT
TCGCCTACGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTACTAGCATTAATCTTAT
CAATCCTAATCCTAGCTCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCC
GCCCAATTACGCAAACAATATACTGAATCCTAGTAGCAGACCTCTTTATCCTCACATGAATCG
GAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAGGCAGCCTCAATTGCCTACTTT
GCCATCATCGTTATCTTCATACCAATCGCAGGAATAATCGAAAACAACATCCTAGACCTAGAT
TAA
KIRCH3
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC

CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
KIRCH4
　　　　　　CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTTATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
NK20275　　　　　　　--

TGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGC
CATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTACTAGCGTTAATCTTATCAATCCTAAT
CCTAGCTCTCATGCCTCTTCTCCATACCTCAAAACAACGAGCACTCACCTTTCGCCCAATTAC
TCAGACAATATACTGAGTCCTAGTAGCGGACCTCCTCATCCTTACATGAATCGGAGGACAAC
CAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCG
TTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCCTAGACCTAGATAAA
DC194Emb　　--
TAAAGCCAACCTCACACGATTCTTCGCCTTCCACTTCATCCTACCTTTCATTATTACCGCCCT
AGTATTAGTTCATCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACTGGTCTAAACTC
AGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCA
TCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAA
TTACACTCCTGCAAATCCACTCAATACTCCACCACACATTAAACCAGAATGATATTTCCTCTT
CGCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTACTAGCATTAATCTTATC
AATCCTAATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCC
GCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATC
GGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAGGCAGCCTCAATTGCCTACTT
TGCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCCTAGACCTAGA
TTAA
DC194　　　　　　--
AAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTA
GTATTAGTTCATCTTTTATTCCTACACGAAACAGGATCAAATTACCCAACTGGTCTAAACTCA
GACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCAT
CCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAAT
TACACTCCTGCAAATCCACTCAATACTCCACCACACATTAAACCAGAATGATATTTCCTCTTC
GCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTACTAGCATTAATCTTATCA
ATCCTAATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCCG
CCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATCG
GAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAGGCAGCCTCAATTGCCTACTTT
GCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCCTAGACCTAGAT
TAA
NK20260
　　　　　　CTGAGGGGCTACAGTAATTACAAATCTGCTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTTTCAGTAGATAAAGCCACTCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTA--

203

MIMO11

```
         CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATTGAAAACAACATCCTAGACCTAGATTAA
```

NK36857   ---

```
TAAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTACCTTTCATTATTACCGCCCT
AGTATTAGTTCACCTTTTATTTCTACACGAAACAGGATCAAATAACCCAACTGGTCTAAACTC
AGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCA
TCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAA
TTACACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGATATTTCCTCTT
CGCCTACGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTACTAGCATTAATCTTAT
CAATCCTAATCCTAGCTCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCC
GCCCAATTACGCAAACAATATACTGAATCCTAGTAGCAGACCTCCTTATCCTCACATGAATCG
GAGGCCAACCAGTCGAATACCCATTCATTATTATTGGACAGGCAGCCTCAATTGCCTACTTT
GCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCCTAGACCTAGAT
TAA
```

NK55568   ---

```
TAAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCT
AGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAATCCAACCGGTCTAAACTC
AGACGCAGACAAAATCCCATTTCACCCTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCA
TCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAA
TTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGATATTTCCTCTT
CGCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGTGGTGTGCTAGCATTAATCTTATC
AATCCTGATCCTAGCTCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCC
GCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATC
GGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACTT
TGCCATCATCGTTATCTTCATACCAATCGCAGGAATAATCGAAAACAACATCCTAGACCTAGA
TTAA
```

STP20EMB   ---

```
AAAGCCACCCTCACACGATTTTTCGCCTTCCACTTCATCCTACCCTTTATTATTACCGCCCTA
GTACTAGTTCACCTCTTATTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTTAACTCA
GACGCAGACAAAATCCCATTTCACCCTTACTACACAGTCAAAGACTTTTTAGGGGTTCTCATC
CTATTGATAGTTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTA
TACTCCTGCAAATCCACTCAACACTCCTCCACACATTAAACCAGAATGGTATTTCCTCTTCGC
CTATGCCATCTTACGATCCATCCCTAACAAGCTAGGTGGTGTGCTAGCATTAATCTTATCAAT
CCTAATCCTAGCTCTCATGCCACTACTTCATACCTCAAAACAACGAGCACTCACCTTCCGCC
CAATTACACAAACAATATACTGAATCCTAGTAGCCGATCTCTTTATCCTCACATGAATCGGCG
GCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCAATTGCCTACTTTGCC
ATTATCGTAATCTTCATACCAATCGCAGGTATGATCGAAAACAACATTCTAGACCTAGATTAA
```

STP20EMB4   ---

```
AAAGCCACCCTCACACGATTTTTCGCCTTCCACTTCATCCTACCCTTTATTATTACCGCCCTA
GTACTAGTTCACCTCTTATTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTTAACTCA
GACGCAGACAAAATCCCATTTCACCCTTACTACACAGTCAAAGACTTTTTAGGGGTTCTCATC
```

CTATTGATAGTTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTCGGAGACCCCGATAATT
ATACTCCTGCAAATCCACTCAACACTCCTCCACACATTAAACCAGAATGGTATTTCCTCTTCG
CCTATGCCATCTTACGATCCATCCCTAACAAGCTAGGTGGTGTGCTAGCATTAATCTTATCAA
TCCTAATCCTAGCTCTCATGCCACTACTTCATACCTCAAAACAACGAGCACTCACCTTCCGCC
CAATTACACAAACAATATACTGAATCCTAGTAGCGGATCTCCTTATCCTCACATGAATCGGCG
GCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCAATTGCCTACTTCGCC
ATTATCGTTATCTTCATACCAATCGCAGGTATGATCGAAAACAACATTCTAGACCTAGATTAA
DC13DC12
          CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATC-------------------------------------
----

NK34915   --

GGTCTTAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCTA
GGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAG
ACCCCGATAATTACACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGAT
ATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTACTAGCAT
TAATCTTATCAATCCTAATCCTAGCTCTCATGCCTCTTCTCCATACCTCAAAACAACGAGCAC
TCACCTTTCGCCCAATTACTCAGACAATATACTGAGTCCTAGTAGCGGACCTCCTCATCCTTA
CATGAATCGGAGGACAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATT
GCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCCTA
GACCTAGATTAA

Appendix H

The 30 *M. montanus* samples used in the analysis of AP5 variation.

Mcalifo
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAACCGTGCAGA
TTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCACGAT
GCCAATGACAAGAGGTTCCAGGTATATGCTCAACAACAGGGTTAGAGGGTCCAGAGGGGAG
GTCACTCCGACCTGGGGCCTTCCTCTGGGGAGGGCAGAGAGAAACAGAAGGGCTCCTCCA
GGTAGCAACCTTGGACATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGC
AGGGTGGCTGTTGACCCTAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAG
GACGTGTTCTCTGACCGTGCCCTCCGCAATATTCCTTGGTACGTGCTAGCTGGAAACCATG--
---
DC00189
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCCAGGGGAGGTCGCTCTGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGT
GCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCT
CTGCAA
76802
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCGAGGGGAGGTCGCTCTGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGT
GCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCT
CTGCAA
DC00119
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCGAGGGGAGGTCGCTCTGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CAGGGGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCC
TAGAGGTGGAATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACTGTGTTCTCTGACCG
TGCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCT
CTGC
DC0193R
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCTGAGGGGAGGTCGCTCCGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CGGGGGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGAATGCCCTTTCTGTCTGTCACAAGAAACCTTTATGGACGTGTTC--------------------
----
STPUNID ---
CCAGAACAGTGCAGATTCTGGGTGCTGACTTCATCATGTCCCTAGGCGAAACTTCTACTTCA
CTGGAGTGCACGATGCCAATGACAACAGGTTCCAGGTTAGAGGGTCCCGAAGGGAGGTCG
CTCCGACCCGGGGCCTTCCTCTGGGGAGGGCGGAGAAAAACAGAAGGGCTCCTCCAGGTA
GCTACCTTGGACATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGG
GGTGGCTGTTGACCCTAGAGGTGGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGG

GACGTGTTTCTCTGACCGTGCCCTCCGCAATATCCCTGGTACGTGCTAGCTGGGGAACCAT
GAACACCTTTGGCAAAC-----
NK05900
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCGAGGGGAGGTCGCTCTGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CAGTGGTTCCTTTGTCCTTGGTATCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGCCCCT
AGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGATGTGTTCTCTGACCGTG
CCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCTCT
GCAA
DC00192
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCGAGGGGAGGTCGCTCTGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCCCCAGGTAGCTACCTTGGACATC
CAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGT
GCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCT
CTGCAA
DC00191
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCGAGGGGAGGTCGCTCTGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGT
GCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCT
CTGCAC
NK29024
GCCCCATTCCACACAGCCAGGGAAATGGCCAATGCTAAAGAAATCGCCAGAACCGTGCAGA
TTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCACGAT
GCCAATGACAAGAGGTTCCAGGTATGTGCTCAACAACAGGGTTAGAGGGTCCCGAGGGGA
GGTCACTCCGACCTGGGGCCTTCCTCTGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCC
AGGTAGCTACCTTGGACATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGG
CAGGGTGGCTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGA
GGACGTGTTCTCTGATCGTGCCTTGTAC--------------------------------------
PAHRN38
GCCCCATTCCACACAGCGCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAACCGTGCAGA
TTATGGGTGCTGACTTCATCATGTCCCTGGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTATACGCTCAACAACAGGGCTAGAGGTTAAAGTGGGGCG
GTCACTTCAACCTGGGGCCCTCCTCTGGGGAGGGCAGAGAAAGGAGAAGAGCTCCTACAG
TAGGTGCCTGGGACACCAAGTGGTTCCTTTGTCCTTGGTCGCTGACTTGGCGAGGAACAGT
CAAGGTAGCTGCTGACCCCAGAGGTGGGGTGCCCTTTCTGTCTGTCACAGGAAACCTTTGA
GGATGTGTTCTCCGACCGTGCCTTGCGCAATGTTCCCTGGTACGTGCTGGCTGGAAACCAT
GACCAC
B76828
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCTGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTTCTTCACCGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCGAGGGGAGGTCGCTCTGACCCGGGGC
CTTCATCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCCCCAGGTAAATACCTTGGACATC
CAGTGGTTCCTTTGTCCATGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGGATGCCCTTTCTCTCTGTCACAGGAAACCTTTGACAACGTGTTCTCTGACCGTG
CCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCTCT
GCAC

STPCRSN
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCACGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCGAAGGGAGGTCGCTCCGACCCGGGGC
CTTCCTCTGGGGAGGGCGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGT
GCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCT
CTGCAC
CVCREEK
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCACGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCGAAGGGAGGTCGCTCCGACCCGGGGC
CTTCCTCTGGGGAGGGCGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGT
GCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCT
CTGCAC
Stepto6
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCCAGGGGAGGTCGCTCTGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGT
GCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCT
CTGCAC
DC00120
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCACGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCCAGGGGAGGTCGCTCTGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGT
GCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCT
CTGCAC
DC00194
GCCCCATTCCACACAGCCAGGGAAATGGCCAATGCTAAAGAAATCGCCAGAACCGTGCAGA
TTATGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCACGAT
GCCAATGACAAGAGGTTCCAGGTATGTGCTCAACAACAGGGTTAGAGGGTCCCGAGGGGA
GGTCACTCCGACCTGGGGCCTTCCTCTGGGGAGGGCGGAGAGAAACAGAAGGGCTCCTCC
AGGTAGCTACCTTGGACATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGG
CAGGGTGGCTGTTGACCCCAGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGA
GGACGTGTTCTCTGATCGTGCCCTCCGCAATATTCCTTGGTACGTGCTAGCTGGAAACCATG
ACCAC
DC00201
GCCACGGACAGAGAACTCGTCCTCAAAGGTTTCCTGTGACTGACAGAAAGGGCATTCCACC
TCTAGGGTCAACAGCCACCCTGCCGCTGCCACCAAATCAGAGACCAAGGACAAAGGAACCC
CCGGATGTCCAAGGTAGCTACCTGGAGGAGCCCTTCTGTTTTTCTCCACCCTCCCCAGAGG
AAGGCCCCGGGTCGGAGCGACCTCCCCTCAGGACCCTCTAACCTGGAACCTCTTGTCATTG
GCATCATGCACTCCAGTGAAGTAGAAGTTGTCCCCTAGGGACATGATGAAGTCAGCACCCA
GAATCTGCACTGTTCTTGGCGATTTCTTTTAGCATGGGCATTTCCCGGGCTGTTTTGAAT------
----
Kirch04
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT

208

GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCGAGGGGAGGTCGCTCTGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGT
GCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCT
CTGCAA
NK20246 ---------
CACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGATTCTGGGTG
CTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGATGCCAATGAC
AAGAGGTTCCAGGTTAGAGGGTCCCGAGGGGAGGTCGCTCTGACCCGGGGCCTTCCTCTG
GGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAAGTAGCTACCTTGGACATCCAGTGGTTC
CTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCTAGAGGTGGG
ATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGTGCCCTCCGCA
ATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCCTTGGCAACGTCTCTGCA
Step030
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCGAGGGGAGGTCGCTCCGACCCGGGGT
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCGCTCCCAGGTAGCTACCTTGGACATC
CGGGGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGT
GCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCT
CTGCAA
Krch011
GCCCCCTTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCTGAGGGGAGGTCGCTCCGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CGGGGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGAATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGTG
CCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCTCT

Step002
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCCAGGGGAGGTCGCTCCGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CAGTGGTTCCTTTGTCGTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCAGTTGACCCT
AGAGGTGGGATGCCCTTTCTGTCTGTCACTAGAAACCTTTGAGGACGTGTTCTCTGACCGTG
CCCTCCGCAATA--
NK20245
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGAACAGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCGAGGGGAGGTCGCTCTGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAACTACCTTGGACATC
CAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGGATGCCCTTTCTCTCTATCACTAGAAACCTTTGAGGACGTGTTCTCTGACCGTG
CCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTG---------------
N50940 ---
ATCTTGCTGATGTGCGGTATGAATGTCCGTTAATATGGCTTCACTTTGCTAACGAAAAGGCT
GCTCCTCTGGGGGGGTCATCATGAACCCTGACTTTGATGCTATATCCTATACCCTGAACACAC
GAACCCCTGGAGGTCGAGGGTGTTTACTACTGGGTGTTCTTCTGTTTTTCTCCGCCCGCCG
CGAATAAAGGACCCATGTAAAAGCGACCTCCCCTCGGGACCCTCTAACCTGGAACCTCTTGT
CATTGGCATCATGCACTCCAGTGAAGAAGAAGTTGTCCCCTAGGGACATGATGAATTCAGCA
CCCACAATCTGCCGGTTCTGGCGATTTCTTTAGCTTTGGCCATTTCCCGGGCTGTGTGGAAT
GGGGCATT--

Stept8 ---
CCGTGCACATTCTGGGTGCTCACTTCTCATGTCCCTAGGGGACAACTTCTACTTCACTGGAG
TGCACGATGCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCGAAGGGAGGTCGCTCCGA
CCCGGGGCCTTCCTCTGGGGAGGGCGGAGAAAAACAAAAGGGCTCCTCCAGGTAGCTACC
TTGGACATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCT
GTTGACCCTAGAGGTGGGATGCCCTTTCTCTCTATCACAGGAAACCTTTGAGGACGTGTTCT
CTGACCGTGCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGG
CAACGTCTCTGCAC
DC0186 --
TCACCAGGACGTGTGCATTCTGGGTGCTCCCTTCATCATGTCCCTACGTCAACAACTGATAC
TTCACTGGAGTGCATGATGCCAATGACTACAGGTTCCAGGTTCGAGGGTCCCGCTGGGAGG
TCTCCCTGACCCGGGACCTTCCTCTGGGGATGGTGGAGAAAATCCAAAGGGCTCCCTCAGG
TACATACCTTGCACATCCAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAG
GGTGGCTGATGACCCTAGAGGTGGGATGCCCTTTCTGTCTATCACAGGAAACCTTTGAGGA
CGTGTTCTCTGACCGTGCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACC
ACCTTGGCAACGTCTCTGCAC
DC00197
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCTGAGGGGAGGTCTCTCCGACCCGGGGC
CTTCATCTGGCGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAGAGACCTTGGACATC
CGGGGGGTTCCTTTGTCCATGGTATCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTTACCCC
CTAGAGGTGGAACGCCCTTTCCGTCTGTCACAGGAAACCTAGAGGACGTGTTATCATACCGT
GCCCAGAGCAA---
DC0196R -------------
CAGAGGTCTGGAATTATATCTGTTGTCATGGGTTTACTTTCGCTAACGGAGAGGGTGTTCCA
CCTGTAGGGGAAACATCCACCCTGACATTGCCACCGAATCATATACCATGGACACAGGAAC
CCCTGGATGACCAAGGTCTCTACCTGGGGCAGCCCTTCTGGTTTTCTCCGACTCCCCAATTA
AAGGCCCCGGGGCGGAGCGACCTCCCCTCTTGACCCTCTAACCTGGAACCTCTTGTCATTG
GGATCATGCACACCGGTGAATTAGAAGTTGTCCCCTAGGGACATGATGAAGTCAGCACCCA
GAATCTGCACGGTCCTGGCGATTTCTTTAGCATTGGCCATTTCCCGGGCTGTGTGGAATGG
GGCATTA--
Kirch03
GCCCCATTCCACACAGCCCGGGAAATGGCCAATGCTAAAGAAATCGCCAGGACCGTGCAGA
TTCTGGGTGCTGACTTCATCATGTCCCTAGGGGACAACTTCTACTTCACTGGAGTGCATGAT
GCCAATGACAAGAGGTTCCAGGTTAGAGGGTCCCGAGGGGAGGTCGCTCTGACCCGGGGC
CTTCCTCTGGGGAGGGTGGAGAAAAACAGAAGGGCTCCTCCAGGTAGCTACCTTGGACATC
CAGTGGTTCCTTTGTCCTTGGTCTCTGATTTGGTGGCAGCGGCAGGGTGGCTGTTGACCCT
AGAGGTGGGATGCCCTTTCTGTCTGTCACAGGAAACCTTTGAGGACGTGTTCTCTGACCGT
GCCCTCCGCAATATTCCCTGGTACGTGCTAGCTGGAAACCATGACCACCTTGGCAACGTCT
CTGCAC

Appendix I

Proportional likelihoods for ancestral state reconstruction from the cyt*b* gene in *M. montanus*. Nodes are as labeled in Figure 7. The populations corresponding to each node state are characterized as follows: 0, southern Sierra Nevada, 1, Colorado Plateau/Sky Islands, 2, White/Inyo Mountains, 3, eastern Great Basin, 4, southern Rocky Mountains, 5, northern Sierra Nevada, 6, northern Rocky Mountains, and 7, Ash Meadows. The most likely ancestral node for the first bifurcation in the tree (node 9, Fig. 7) and the second bifurcation (node 10, Fig. 5) was the southern Sierra Nevada Mountains. Significant values are shown in bold, and only values ≥ 0.005 are reported.

	0	1	2	3	4	5	6	7
Node 9	**0.99**							
Node 10	**0.78**							**0.11**
Node 11	0.01	**0.36**	0.01	0.02	**0.36**	0.01	0.01	**0.17**
Node 17		0.02	0.01	**0.68**				**0.30**
Node 32	0.01	0.04	0.01	0.02	0.01	0.03	**0.20**	**0.68**
Node 18				**0.35**				**0.63**
Node 6	**0.99**	0.08	0.03	0.03	0.02	0.03	0.02	0.02

APPENDIX J

The 48 *M. montanus* samples analyzed in the microsatellite analyses.

Population ID	Sample Name	Locality
Steptoe	STEPTOE1	Nevada: White Pine Co., Steptoe Wildlife Management Area
	STEPTOEEMB2	Nevada: White Pine Co., Steptoe River
	STEPTOE2	Nevada: White Pine Co., Steptoe River
	STEPTOE3	Nevada: White Pine Co., Steptoe Wildlife Management Area
	STEPTOE6	Nevada: White Pine Co., Steptoe Wildlife Management Area
	STEPTOE8	Nevada: White Pine Co., Steptoe Wildlife Management Area
	STEPTOE9	Nevada: White Pine Co., Steptoe Wildlife Management Area
	SHLLNBRGR2	Nevada: White Pine Co., Steptoe Wildlife Management Area Shallonberger Spring
	SHLLNBRGR3	Nevada: White Pine Co., Steptoe Wildlife Management Area Shallonberger Spring
	STEPTOE10	Nevada: White Pine Co., Steptoe Wildlife Mangement Area, field of wild rye
	STEPTOE15	Nevada: White Pine Co., Steptoe Wildlife Management Area, Fish Pond Springs
	CAVECREEK	Nevada: White Pine Co., Cave Creek
Kirch	KIRCH3	Nevada: Lincoln Co., Kirch Wildlife Management Area
	KIRCH11	Nevada: Lincoln Co., Kirch Wildlife Management Area
	KIRCH15	Nevada: Lincoln Co., Kirch Wildlife Management Area
	KIRCH7	Nevada: Lincoln Co., Kirch Wildlife Management Area
	KIRCH3	Nevada: Lincoln Co., Kirch Wildlife Management Area
	KIRCH4	Nevada: Lincoln Co., Kirch Wildlife Management Area
	KIRCHTL114	Nevada: Lincoln Co., Kirch Wildlife Management Area
	KIRCHTL178	Nevada: Lincoln Co., Kirch Wildlife Management Area
	KIRCHTL118	Nevada: Lincoln Co., Kirch Wildlife Management Area
	KIRCHTL135	Nevada: Lincoln Co., Kirch Wildlife Management Area
	KIRCHTL130	Nevada: Lincoln Co., Kirch Wildlife Management Area
Key Pittman	KP193	Nevada: Lincoln Co., Key Pittman Wildlife Management Area
	KP194	Nevada: Lincoln Co., Key Pittman Wildlife Management Area
	KP196	Nevada: Lincoln Co., Key Pittman Wildlife Management Area
	KP197	Nevada: Lincoln Co., Key Pittman Wildlife Management Area
	KP198	Nevada: Lincoln Co., Key Pittman Wildlife Management Area
	KP199	Nevada: Lincoln Co., Key Pittman Wildlife Management Area
	KP200	Nevada: Lincoln Co., Key Pittman Wildlife Management Area
	KP201	Nevada: Lincoln Co., Key Pittman Wildlife Management Area
Pahranagat	PAH118	Nevada: Lincoln Co., Pahranagat Valley National Wildlife Refuge (NWR)
	PAH119	Nevada: Lincoln Co., Pahranagat Valley NWR
	PAH120	Nevada: Lincoln Co., Pahranagat Valley NWR
	PAH121	Nevada: Lincoln Co., Pahranagat Valley NWR
	PAH122	Nevada: Lincoln Co., Pahranagat Valley NWR
	PAH123	Nevada: Lincoln Co., Pahranagat Valley NWR
	PAH124	Nevada: Lincoln Co., Pahranagat Valley NWR
	PAH125	Nevada: Lincoln Co., Pahranagat Valley NWR

Appendix J, cont'd

Population ID	Sample Name	Locality
	PAH13	Nevada: Lincoln Co., Pahranagat Valley National Wildlife Refuge
	656174F	Nevada: Lincoln Co., Crystal Springs
	667696	Nevada: Lincoln Co., Crystal Springs
Stillwater	STLLWTR185	Nevada: Churchill Co., Stillwater National Wildlife Refuge
	STLLWTR184	Nevada: Churchill Co., Stillwater National Wildlife Refuge
	STLLWTR183	Nevada: Churchill Co., Stillwater National Wildlife Refuge
	STLLWTR191	Nevada: Churchill Co., Stillwater National Wildlife Refuge
	STLLWTR192	Nevada: Churchill Co., Stillwater National Wildlife Refuge
	STLLWTR186	Nevada: Churchill Co., Stillwater National Wildlife Refuge

APPENDIX K

Samples of *M. montanus* used in analysis of cytochrome *b* variation

Stllwater183
CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAACACTAG
TAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCGCCTTC
CACTTCATCCTACCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACACGAAA
CAGGATCAAACAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACCCTTAC
TACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTAT
TTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACTCCAC
CACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCTAACA
AGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTCCTCC
ATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAG
TAGCGGACCTCTTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATT
ATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGT
ATAATCGAAAACAACATCCTAGACCTAGATTAA
Stllwater184
CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATTCCCTACATCGGCACAACACTAG
TAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCGCCTTC
CACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTTCATCTTTTATTCCTACACGAAA
CAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTAC
TACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTAT
TTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACTCCAC
CACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCTAACA
AGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTAATCCTAGCCCTCATGCCACTCCTC
CATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTA
GTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCAT
TATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGG
TATAATCGAAAACAACATCCTAGACCTAGATTAA
NK005897
CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAACACTAG
TAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCGCCTTC
CACTTCATCCTACCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACACGAAA
CAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTAC
TACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTAT
TTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACTCCAC
CACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCTAACA
AGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTCCTCC
ATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAG
CAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATT
ATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGT
ATAATCGAAAACACCATCCTAGACCTAGATTAA
Stllwater190
CTGAGGGGCTACAGTTATTACAAATCTACTATCAGCCATTCCCTACATCGGCACAACACTAG
TAGAGTGAATCTGAGGGGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCGCCTTC
CACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTTCATCTTTTATTCCTACACGAAA
CAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTAC
TACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTAT
TTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACTCCAC
CACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCTAACA
AGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCCCTCATGCCACTCCTCC
ATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAG
TAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATT
ATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA

Stllwater186

CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATTCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTTCATCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCTTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA

Stllwater185

CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAATGAATCTGAGGGGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA

DC000191

CTGAGGGGCTACAGTTATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTTCATCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATAAA

DC000192

CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA

Stllwater187

CTGAGGGGCTACAGTTATTACAAATCTACTATCAGCCATTCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTTCATCTTTTATTCCTACA

CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
MI201876
　　　　　CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGCAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
MI217847
　　　　　CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAACAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
NK5899
　　　　　CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCTTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAATACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGCAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
PAH00122
　　　　　CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT

AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATCATCGTAATCTTCATACCAATCG
CAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
PAH00124 ---

GCAGACAAAATCCCATTTCACCCTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTA
TTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATA
CTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCT
ATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCC
TGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCA
ATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGG
CCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCA
TCATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
PAH00013
CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATCATCGTAATCTTCATACCAATCG
CAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
PAH00121 ---

TCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTTCTAGG
GGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGAC
CCCGATAATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTAT
TTCCTCTTCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATT
AATCTTATCAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACT
CACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCA
CATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATT
GCCTACTTCGCCATCATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTA
GACCTAGATTAA
PAH10120
CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCG
CAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
PAH00119
CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA

CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCG
CAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
MIM77610 --
ATTCTTCGCCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTA
TTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCC
ATTCCACCCCTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCAT
AATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCA
CTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGA
TCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTC
ATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAAT
ATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAGCCAGTCGAAT
ACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATCATCGTAATCTTCA
TACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
MIM76828 --

TCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACC
CAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCCTACTACACAGTCAAAGAC
TTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCT
TGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAG
AATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTG
CTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAA
CGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCT
TATCCTCACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCATCATTATTGGACAAGCAG
CCTCGATTGCCTACTTCGCCATCATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACA
ACATCCTAGACCTAGATTAA
MIM76802 --
TAAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCT
AGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTAAACTC
AGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCA
TCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAA
TTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCTCTT
CGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATTAATCTTAT
CAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTC
CGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTTACATGAATC
GGAGGCCAGCCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCAATTGCCTACTT
CGCCATCATCGTAATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCCTAGACCTAGA
TTAA
PAH00123
          CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGAGGCTTCTCAGTAGATAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAACAC
TCCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCC
TAACAAACTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCAGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCA

TCATTATTGGACAGGCAGCCTCAATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCG
CAGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
KEYP0193
            CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATCATCGTAATCTTCATACCAATCG
CAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
KeyP0194 --
AAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTA
GTATTAGTTCATCTTTTATTCCTACACGAAACAGGATCAAATTACCCAACTGGTCTAAACTCA
GACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCAT
CCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAAT
TACACTCCTGCAAATCCACTCAATACTCCACCACACATTAAACCAGAATGATATTTCCTCTTC
GCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTACTAGCATTAATCTTATCA
ATCCTAATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCCG
CCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATCG
GAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAGGCAGCCTCAATTGCCTACTTT
GCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCCTAGACCTAGAT
TAA
KeyP0198   1   --

TTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACCG
GTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTTCTAG
GGGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGA
CCCCGATAATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTA
TTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCAT
TAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCAC
TCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCA
CATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATT
GCCTACTTCGCCATCATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTA
GACCTAGATAAA
KeyP0197   --

CTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTAAAC
TCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTTCTAGGGGTTCT
CATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGAT
AATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCTC
TTCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATTAATCTT
ATCAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCT
TCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAA
TCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTAC
TTTGCCATCATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCTA
GATAAA
KeyP0199 --

TTATTACCGCCCTAGTATTAGTTCACCTTTTATTTCTACACGAAACAGGATCAAATAACCCAAC
CGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCC
TAGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGA

219

GACCCCGATAATTACACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATG
ATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTACTAG
CATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTCCTCCATACCTCAAAACAACGAG
CACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATC
CTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTC
AATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACAT
CCTAGACCTAGATAAA
DC194Emb --
TAAAGCCAACCTCACACGATTCTTCGCCTTCCACTTCATCCTACCTTTCATTATTACCGCCCT
AGTATTAGTTCATCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACTGGTCTAAACTC
AGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCA
TCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAA
TTACACTCCTGCAAATCCACTCAATACTCCACCACACATTAAACCAGAATGATATTTCCTCTT
CGCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTACTAGCATTAATCTTATC
AATCCTAATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCC
GCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATC
GGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAGGCAGCCTCAATTGCCTACTT
TGCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCCTAGACCTAGA
TTAA
DC194 ---
AAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTACCTTTCATTATTACCGCCCTA
GTATTAGTTCATCTTTTATTCCTACACGAAACAGGATCAAATTACCCAACTGGTCTAAACTCA
GACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCTAGGAGTTCTCAT
CCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAAT
TACACTCCTGCAAATCCACTCAATACTCCACCACACATTAAACCAGAATGATATTTCCTCTTC
GCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTACTAGCATTAATCTTATCA
ATCCTAATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCCG
CCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATCG
GAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAGGCAGCCTCAATTGCCTACTTT
GCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCCTAGACCTAGAT
TAA
KeyP195
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACAAGATTCATCG
CATTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCC---------------------------------------
----
ATCAATCTTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTT
CCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAAT
CGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACT
TTGCCATCATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCTAG
ATAAA
PAH00118
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT

CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
656174M--
CCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTAT
TAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACG
CAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTAT
TGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATAC
TCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTA
TGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCT
GATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAA
TTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGC
CAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCAT
CATCGTAATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
666194 ---

ATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACC
GGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCT
AGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGA
GACCCCGATAATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGA
TATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGC
ATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTTCTCCATACCTCAAAACAACGAGC
ACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTCATCC
TTACATGAATCGGAGGACAACCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCG
ATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATC
CTAGACCTAGATTAA
667696 ---

TATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATTAAATAACCCAACT
GGTCTAAACTCAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCT
AGGAGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGA
GACCCCGATAATTACACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATG
ATATTTCCTCTTCGCCTACGCCATCTTACGATCCATCCCTAACAAACTTGGCGGTGTACTAGC
ATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTCCTCCATACCTCAAAACAACGAGC
ACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCAGACCTCCTTATCCT
CACATGAATCGGAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAGGCAGCCTCAA
TTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGCAGGTATAATCGAAAACAACATCC
TAGACCTAGATTAA
STEP0006
        CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTCACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATAAA
KIRCH017
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTCACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC

TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATAAA
SHLNBRG3
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
SHLLNBR2
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
STEPTO30
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
KIRCH15
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC

CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
STEP0015
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
STEP002   ---
TAAAAGCCACCCTCACACGATTCTTCGCCTTCCACTTCATCCTGCCTTTCATTATTACCGCCC
TAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAATAACCCAACCGGTCTAAACT
CAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTTCTAGGGGTTCTC
ATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATA
ATTATACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGGTATTTCCTCT
TCGCCTATGCCATCCTACGATCCATCCCTAACAAGCTTGGCGGTGTGCTAGCATTAATCTTA
TCAATCCTGATCCTAGCCCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTT
CCGCCCAATTACGCAAACAATATACTGAATCCTAGTAGCGGACCTCCTTATCCTTACATGAAT
CGGAGGCCAGCCAGTCGAATACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACT
TCGCCATCATCGTTATCTTCATACCAATCGCAGGCATAATCGAAAACAACATCCTAGACCTAG
ATTAA
MIMO5 --

CAGACGCAGACAAAATCCCATTCCACCCTTACTACACAGTCAAAGACTTCCTAGGAGTTCTC
ATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGATATTCTTGGAGACCCCGATA
ATTACACTCCTGCAAATCCACTCAACACTCCACCACACATTAAACCAGAATGATATTTCCTCT
TCGCCTACGCCATCTTACGATCCATCCCTAACAAGCTTGGCGGTGTACTAGCATTAATCTTAT
CAATCCTAATCCTAGCTCTCATGCCACTCCTCCATACCTCAAAACAACGAGCACTCACCTTCC
GCCCAATTACGCAAACAATATACTGAATCCTAGTAGCAGACCTCTTTATCCTCACATGAATCG
GAGGCCAACCAGTCGAATACCCATTCATCATTATTGGACAGGCAGCCTCAATTGCCTACTTT
GCCATCATCGTTATCTTCATACCAATCGCAGGAATAATCGAAAACAACATCCTAGACCTAGAT
TAA
KIRCH3
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
KIRCH4
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA

CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCTATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTTATCTTCATACCAATCGC
AGGCATAATCGAAAACAACATCCTAGACCTAGATTAA
MIMO11
          CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTTCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGGTATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAACTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCCCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTTACATGAATCGGAGGCCAGCCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTCGCCATTATCGTAATCTTCATACCAATCGC
AGGCATAATTGAAAACAACATCCTAGACCTAGATTAA
MVZ59382
          CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATACCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAACAC
TCCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCC
TAACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTACCTCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCTTCATCCTCACATGAATCGGAGGACAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCG
CAGGTATAATCGAAAACAACATCCTAGACCTAGATAAA
MVZ59385 --

TGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACACGAAACAGGATCAAA
TAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACCCTTACTACACAGTCA
AAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAACTTTATTTTTCCCAGA
TATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAACACTCCACCACACATTAA
ACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCTAACAAGCTTGGCG
GTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCTCTTCTCCA------------------
----

MVZ59388
          CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTACACTCCTGCAAATCCACTCAACAC
TCCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCC
TAACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTGATCCTAGCTCTCATGCCACT
CCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAAACAATATACTGAAT
CCTAGTAGCGGACCTCTTCATCCTCACATGAATCGGAGGACAACCAGTCGAATACCCATTCA

```
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCG
CAGGTATAATCGAAAACAACATCCTAGACCTAGATAAA
```
MVZ59390
```
        CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTTCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACCGGTCTAAACTCAGACGCAGACAAAATCCCATTTCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCTCTT
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAGACAATATACTGAAT
CCTAGTAGCGGACCTCTTCATCCTTACATGAATCGGAGGACAACCAGTCGAATACCCATTCA
TCATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCG
CAGGAATAATCGAAAACAACATCCTAGACCTAGATAAA
```
MVZ59379
```
        CTGAGGGGCTACAGTAATTACAAACCTACTATCAGCGATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTCTCAGTAGACAAAGCCACCCTAACACGATTCTTCG
CCTTCCACTTCATCC------------------------------------------------------------------------------------------------
--------------------------------------------------------------------------------------------------------------------------------------
-------------------------------------------------------------
ATCCATCCCTAACAAGCTTGGCGGTGTACTAGCATTAATCTTATCAATCCTGATCCTAGCTCT
CATGCCACTTCTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACGCAGACAA
TATACTGAATCCTAGTAGCGGACCTCTTCATCCTTACATGAATCGGAGGACAACCAGTCGAA
TACCCATTCATCATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTC
ATACCAATCGCAGGTATAATCGAAAACAACATCCTAGACCTAGATAAA
```
NK020114
```
        CTGAGGGGCTACAGTAATTACAAATCTGCTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATTTGAGGAGGCTTTTCAGTAGATAAAGCCACTCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CCTACTACACAATCAAAGACTTCCTAGGGATTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCCGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAGCCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGCGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCCCTC
CTCCATACCTCAAAACAACGATCACTCACCTTTCGCCCAATTACCCAGACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
```
NK029024
```
        CTGAGGGGCTACGGTAATTACAAATCTGCTATCAGCCATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATTTGAGGAGGCTTTTCAGTAGATAAAGCCACTCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACA
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGATTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCTGATAATTACACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAGCCAGAATGGTATTTCCTCTTCGCCTATGCCATCTTACGATCCATCCCT
AACAAGCTTGGCGGCGTACTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCCCTT
CTCCATACCTCAAAACAACGATCACTCACCTTTCGCCCAATTACCCAGACAATATACTGAATC
CTAGTAGCGGACCTCCTCATCCTTACATGAATCGGAGGCCAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
```
NK20245
```
        CTGAGGGGCTACAGTAATTACAAATCTACTATCAGCTATCCCCTACATCGGCACAAC
ACTAGTAGAGTGAATCTGAGGGGGCTTTTCAGTAGACAAAGCCACTCTAACACGATTCTTCG
CCTTCCACTTCATCCTGCCTTTCATTATTACCGCCCTAGTATTAGTCCACCTTTTATTCCTACA
```

```
CGAAACAGGATCAAATAACCCAACTGGTCTAAACTCAGACGCAGACAAAATCCCATTCCACC
CTTACTACACAGTCAAAGACTTCCTAGGGGTTCTCATCCTATTGATAGCTTTCATAATTTTAAC
TTTATTTTTCCCAGATATTCTTGGAGACCCTGATAATTATACTCCTGCAAATCCACTCAACACT
CCACCACACATTAAACCAGAATGATATTTCCTCTTCGCCTATGCCATCCTACGATCCATCCCT
AACAAGCTTGGCGGTGTGCTAGCATTAATCTTATCAATCCTAATCCTAGCTCTCATGCCACTC
CTCCATACCTCAAAACAACGAGCACTCACCTTCCGCCCAATTACTCAAACAATATACTGAATC
CTAGTAGCGGACCTCCTTATCCTCACATGAATCGGAGGACAACCAGTCGAATACCCATTCAT
CATTATTGGACAAGCAGCCTCGATTGCCTACTTTGCCATCATCGTTATCTTCATACCAATCGC
AGGTATAATCGAAAACAACATCCTAGACCTAGATTAA
```

226

Appendix L

Data used in microsatellite analyses

Sample ID	Population	M2	M6	M8	MAR76	MAR80	MAR113
PopOne1	Steptoe	179	147	187	122	204	113
		179	147	189	122	204	117
PopOne2	Steptoe	163	147		114	204	113
		195	149		120	204	113
PopOne3	Steptoe	163	147		114	204	113
		195	149		120	204	117
PopOne4	Steptoe	179	147	189	126	204	113
		181	147	189	126	204	115
PopOne5	Steptoe	179	147	187	122	204	113
		181	149	189	122	204	117
PopOne6	Steptoe	183	147	189	106	204	113
		185	149	189	128	204	113
PopOne7	Steptoe	165	149	185		204	113
		179	149	187		204	119
PopOne8	Steptoe	187	147	189	124	204	111
		189	147	189	124	204	111
PopOne9	Steptoe	179	147			210	113
		181	149			210	115
PopOne10	Steptoe	179	147	187	122	204	113
		181	149	189	122	204	117
PopOne11	Steptoe	179	147	187	118	204	101
		179	149	189	118	204	103
PopOne12	Steptoe	163	147	187	118	204	103
		185	147	187	118	212	119
PopOne13	Steptoe	165					113
		187					117

Appendix L

Data used in microsatellite analyses.

Sample ID	Population	M2	M6	M8	MAR76	MAR80	MAR113
PopTwo1	Kirch	191	147	187	118	204	104
		191	147	187	118	204	104
PopTwo2	Kirch	193	147	185	124	204	101
		193	147	187	124	204	113
PopTwo3	Kirch	191	143		106	204	111
		191	145		106	204	113
PopTwo4	Kirch	177	147	185		212	
		187	147	187		212	
PopTwo5	Kirch	153	147		118	204	103
		232	147		122	204	111
PopTwo6	Kirch	191	141		124	204	109
		191	145		124	204	111
PopTwo7	Kirch		147	185			113
			147	187			113
PopTwo8	Kirch	151	147				101
		179	147				113
PopTwo9	Kirch	151	147	185			
		151	147	187			
PopTwo10	Kirch	179	147		118	204	
		191	147		118	204	
PopTwo11	Kirch	179	147			204	113
		191	147			204	121

Appendix L

Data used in microsatellite analyses.

Sample ID	Population	M2	M6	M8	MAR76	MAR80	MAR113
PopThree1	Key Pittman	191	147	181	118	204	121
		191	147	181	120	204	125
PopThree2	Key Pittman	181	147	185	106	204	109
		187	159	187	108	204	117
PopThree3	Key Pittman	187	147	179	118	200	111
		191	147	179	118	200	121
PopThree4	Key Pittman	187	147	179	116	200	111
		187	147	183	118	200	125
PopThree5	Key Pittman	187	147	179	118	200	109
		187	147	183	118	200	111
PopThree6	Key Pittman	179	147	179	118		109
		185	147	179	118		121
PopThree7	Key Pittman	189	147	179	116	204	117
		193	147	179	120	204	119
PopThree8	Key Pittman	177	147	179	118	200	111
		187	147	181	118	200	113

Appendix L

Data used in microsatellite analyses.

Sample ID	Population	M2	M6	M8	MAR76	MAR80	MAR113
PopFour1	Pahranagat	187	147	178	116		115
		193	147	180	120		117
PopFour2	Pahranagat	189	147	178	120	200	117
		193	161	180	120	200	121
PopFour3	Pahranagat	187		178	116		117
		193		180	120		119
PopFour4	Pahranagat	189	147	177	112	204	117
		193	147	179	120	204	119
PopFour5	Pahranagat	187	147	179	120	200	115
		189	147	179	120	200	117
PopFour6	Pahranagat	179	147	179	122		119
		199	147	179	122		121
PopFour7	Pahranagat	193	147	179	120		117
		199	147	179	120		121
PopFour8	Pahranagat	189	147	177	116	200	115
		193	147	179	120	200	117
PopFour9	Pahranagat	189	147	179	116	204	109
		193	147	179	120	204	111
PopFour10	Pahranagat	179	147	177	114	204	111
		179	147	179	116	204	111
PopFour11	Pahranagat	187	147	177	118	204	109
		189	147	177	120	204	111
PopFour12	Pahranagat	194	147	185	122	212	117
		194	149	185	122	212	119

Appendix L

Data used in microsatellite analyses.

Sample ID	Population	M2	M6	M8	MAR76	MAR80	MAR113
PopFive2	Stillwater	189	147	185	106	204	109
		191	159	189	118	204	121
PopFive3	Stillwater	187	147	189	108	204	111
		191	147	189	110	204	113
PopFive4	Stillwater	193	147	187	108	204	111
		197	147	187	116	204	117
PopFive5	Stillwater	193	147	189	124	204	115
		197	159	189	126	204	117
PopFive1	Stillwater	189	147	185		204	110
		189	151	189		204	110

LITERATURE CITED

Ad hoc Committee for Acceptable Field Methods in Mammalogy (1987) Acceptable field methods in mammalogy: preliminary guidelines approved by the American Society of Mammalogists. Journal of Mammalogy 68: 1-18.

Allen, B.D. and Anderson, R.Y. (2000) A continuous high-resolution record of late Pleistocene climate variability from the Estancia Basin, New Mexico. GSA Bulletin 112: 1444-1458.

Alvarez-Castenada, T. (2005) *Peromyscus melanotis*. Mammalian Species 764: 1-4.

Anderson, S. (1959) Distribution, variation, and relationships of the montane vole, *Microtus montanus*. University of Kansas Publications, Museum of Natural History 9: 415-511.

Arbogast, B.S., Browne, R.A. and Weigl, P.D. (2001) Evolutionary genetics and Pleistocene biogeography of North American tree squirrels (*Tamiasciurus*). Journal of Mammalogy 82: 302-319.

Armstrong, D.M. (1972) Distribution of mammals in Colorado. University of Kansas Museum of Natural History Monograph 3: 1-415.

Avise, J.C. (2000) Phylogeography: the history and formation of species. Cambridge MA. Cambridge University Press. 464 pp.

Baars, D.L. (2000) The Colorado Plateau: a geologic history. University of New Mexico Press, Albuquerque, NM. 254pp.

Balloux, F. (2001) EASYPOP (version 1.7). A computer program for the simulation of population genetics. Heredity 92: 301-302.

Beerli, P. and Felsenstein, J. (2001) Maximum likelihood estimation of a migration

    matrix and effective population sizes in n subpopulations by using a coalescent

    approach. Proceedings of the National Academy of Sciences 98: 4563-4568.

Beever, E.A., Brussard, P.F. and Berger, J. (2003) Patterns of apparent extirpation among

    isolated populations of pikas (*Ochotona princeps*) in the Great Basin. Journal of

    Mammalogy 84: 37-54.

Benson, L.V., Burdett, J.W., Kashgarian, M., Lund, S.P., Phillips, F.M. and Rye, R.O.

    (1996) Climatic and hydrologic oscillations in the Owens Lake Basin and

    adjacent Sierra Nevada, California. Science 274: 746-748.

Betancourt, J.L., Rylander, K.A., Penalba, C. and McVickar, J.L. (2001) Late Quaternary

    vegetation history of Rough Canyon, south-central New Mexico, USA.

    Palaeogeography, Palaeoclimatology, Palaeecology 165: 71-95.

Beyer, H.L. (2004) Hawth's Tools for ArcGIS. Available at

    http://www.spatialecology.com/htools.

Boyer, A. (2008) Extinction patterns in the avifauna of the Hawaiian Islands. Diversity

    and Distributions 14: 509-517.

Bromberg, C, Cash, H., Curtis, P., Goebel III, C.J., Irwin, L., Singer, J.,

    Van Hoewyk, D. and Winkelplek, J. (1995) Sequencher. Gene Codes

    Corporation. Ann Arbor, Michigan.

Brown, J.H. (1984) On the relationship between abundance and distribution of species.

    American Naturalist 124: 255-279.

Brown, J.H. (1978) The theory of insular biogeography and the distribution of boreal birds and mammals. K. Intermountain Biogeography: a Symposium (ed. by T. Harper and J.L. Reveal). Great Basin Naturalist Memoirs 2: 209-227.

Brown, J.H. (1971) Mammals on mountaintops: non-equilibrium insular biogeography. American Naturalist 105: 467-478.

Brown, J.H. and Lomolino, M.V. (1998) Biogeography. Second Edition. Sinauer Associates, Inc. Sunderland, Massachusetts.

Brown, W.M., George, M. and Wilson, A. (1979) Rapid evolution of animal mitochondrial DNA. Proceedings of the National Academy of Sciences 76: 1967-1971.

Bull, J.C., Pickup, N.J., Pickett, B., Hassell, M.P. and Bonsall, M.B. (2006) metapopulation risk is increased by environmental stochasticity and assemblage complexity. Proceedings of the Royal Society B 274: 87-96.

Burbridge, A.A. and Manly, B.F.J. (2002) Mammal extinction on Australian islands: causes and conservation implications. Journal of Biogeography 29: 465-473.

Carstens, B.C., Brunsfeld, S.J., Demboski, J.R., Good, J.M. and Sullivan, J. (2005) Investigating the evolutionary history of the Pacific Northwest Mesic Forest Ecosystem: hypothesis testing within a comparative phylogeographic framework. Evolution 59: 1639-1652.

Castoe, T.A., Spencer, C.L. and Parkinson, C.L. (2007) Phylogeographic structure and historical demography of the western diamondback rattlesnake (*Crotalus atrox*): A perspective on North American desert biogeography. Molecular Phylogenetics and Evolution 42: 193-212.

Catzeflis, F. M., Sheldon, F.H., Ahlquist, J.E. and Sibley, C.G. (1987) DNA-DNA hybridization evidence of the rapid rate of muroid rodent evolution. Molecular Biology and Evolution 4: 242-253.

Ceballos, G. and Oliva, G. (2006) Diversidad y conservacion de los mamiferos de Mexico. Mamiferos Silvestres De Mexico (ed. by G. Ceballos and G. Oliva), pp.21-49. Fondo De Cultura Economica USA. 986pp.

Chaline, J. and Graf, J.D. (1988) Phylogeny of the Arvicolidae (Rodentia): Biochemical and paleontological evidence. Journal of Mammalogy 69: 22-33.

Channel, R. and Lomolino, M.V. (2000) Trajectories to extinction: spatial dynamics of the contraction of geographical ranges. Journal of Biogeography 27: 169-179.

Clement, M., Posada, D. and Crandall, K.A. (2000) TCS: a computer program to estimate gene genealogies. Molecular Ecology 9: 1657-1660

Coates, A.G. and Obando, J.A. (1996) The geologic evolution of the Central American Isthmus. Evolution and Environment in Tropical America (ed. by J.B.C. Jackson, A.F. Budd and A.G. Coates), pp.21–56. University of Chicago Press, Chicago.

Conroy, C.J. and Cook, J.A. (2000) Phylogeography of a post-glacial colonizer: *Microtus longicaudus* (Muridae: Rodentia). Molecular Ecology 9: 165-175.

Conroy, C.J. and Cook, J.A. (1999) MtDNA evidence for repeated pulses of speciation within arvicoline and murid rodents. Journal of Mammalian Evolution 6: 221-245.

Conroy, C.J., Hortelano, Y., Cervantes, F. and Cook, J.A. (2001) The phylogenetic position of southern relictual species of *Microtus*. Mammalian Biology 66: 332-344.

Conroy, C.J. and Neuwald, J.L. (2008) Phylogeographic study of the California vole, *Microtus californicus*. Journal of Mammalogy 89: 755-767.

Cracraft, J. and Plum, R.O. (1988) Patterns and processes of diversification: speciation and historical congruence in some Neotropical birds. Evolution 42: 603-620.

Crawford, D.L. and Smith, F. A. (*in prep*.). Examining refugial dynamics for the montane vole (*Microtus montanus*) in the Great Basin.

Crawford, M.H. (2007) Anthropological Genetics: Theory, Methods and Applications. Cambridge University Press, Cambridge, UK.

Creer, S., Malhotra, A. and Thorpe, R.S. (2003) Assessing the phylogenetic utility of four mitochondrial genes and a nuclear intron in the Asian pit viper Genus, *Trimeresurus*: separate, simultaneous, and conditional data combination analyses. Molecular Biology and Evolution 20: 1240-1251.

Crespi, E.J., Rissler, L.J. and Browne, R.A. (2003) Testing Pleistocene refugia theory: phylogeographical analysis of *Desmognathus wrighti*, a high-elevation salamander in the southern Appalachians. Molecular Ecology 12: 969-984.

Davis, R. and Callahan, J.R. (1992) Post-Pleistocene dispersal in the Mexican vole (*Microtus mexicanus*): an example of an apparent trend in the distribution of southwestern mammals. Great Basin Naturalist 52: 262-258.

Davis, R., Dunford, C. and Lomolino, M.V. (1988) Montane mammals of the American southwest: the possible influence of post-Pleistocene colonization. Journal of Biogeography 15: 841-848.

Debry, R.W. and Sheshadri, S. (2001) Nuclear intron sequences for phylogenetics of closely related mammals: an example using the phylogeny of *Mus*. Journal of Mammalogy 82: 280-288.

DeChaine, E.G. and Martin, A.P. (2004) Historic cycles of fragmentation and expansion in *Parnassius smintheus* (Papilionidae) inferred using mitochondrial DNA. Evolution 58: 113-127.

Drummond, A.J., Ho, S.Y.W., Phillips, M.J. and Rambaut, A. (2006) Relaxed phylogenetics and dating with confidence. PLOS Biology 4: 699-710.

Dudik, M., Phillips, S.J. and Schapire, R.E. (2004) Performance guarantees for regularized maximum entropy density estimation. Proceedings of the 17th Annual Conference on Computational Learning Theory. ACM Press, New York, pp. 655-662.

Duncan, R.P. and Forsyth, D.M. (2006) Modeling population persistence on islands: mammal introductions in the New Zealand archipelago. Proceedings of the Royal Society 273: 2969-2975.

Dunham, J., Peacock, M., Tracy, C.R., Nielsen, J. and Vinyard, G. (1999) Assessing extinction risk: integrating genetic information. Conservation Biology (online) 3: 2.

Durish, N.D., Halcomb, K.E., Kilpatrick, C.W. and Bradley, R.D. (2004) Molecular systematics of the *Peromyscus truei* species group. Journal of Mammalogy 85: 1160-1169.

Dyurgerov, M.B. and Meier, M.F. (2000) Twentieth-century climate change: evidence from small glaciers. Proceedings of the National Academy of Sciences 97: 1406-1411.

Eckstein, R.L., O'Neill, R.A., Danihelka, J., Otte, A. and Kohler, W. (2006) Genetic structuring among and within peripheral and central populations of three endangered floodplain violets. Molecular Ecology 15: 2367-2379.

Edwards, C.W. and Bradley, R.D. (2002) Molecular systematics and historical phylobiogeography of the *Neotoma mexicana* species group. Journal of Mammalogy 83: 20-30.

Eisenberg, J.F. (1981) The mammalian radiation. University of Chicago Press, Chicago, Illinois, USA.

Elith, J. (2002) Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. (ed. by S. Ferson and M.

Burgman), pp. 39-58. Quantitative methods for conservation biology. Springer, NY.

Epps, C.W., McCullough, D.R., Wehausen, J.D., Blech, V.C. and Rechel, J.L. (2004) Effects of climate change in population persistence of desert-dwelling mountain sheep in California. Conservation Biology 18: 102-113.

Evanno, G., Regnault, S. and Goudet, J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14: 2611-2620.

Excoffier, L., Laval, G. and Schneider, S. (2006) ARLEQUIN, (version 3.11): An integrated software package for population genetic analysis. Evolutionary Bioinformatics Online 1: 47-50.

Fa, J.E. and Morales, L.M. (1993) Patrones de diversidad e mamíferos de México. Diversidad Biológica de México. Orígenes y Distribución. (ed. by T.P. Ramamoorthy, R. Bye, A. Lot, and J. Fa). pp. 315-354. Instituto de Biología, UNAM.

Fagan, W.F., Aumann, C., Kennedy, C.M. and Unmack, P.J. (2005) Rarity, fragmentation, and the scale dependence of extinction risk in fishes. Ecology 86: 34-41.

Falush, D., Stephens, M. and Pritchard, J.K. (2003) Inference of population structure using multilocis genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567-1587.

Faunmap Working Group (R.W. Graham, E.L.Lundelius, Jr., M.A.Graham, E.K. Schroeder, R.S.Toomey, III, E.Anderson, A.D.Barnosky, J.A.Burns, C.S. Churcher, D.K.Grayson, R.D.Guthrie, C.R.Harington, G.T.Jefferson, L.D.Martin, H.G.McDonald, R.E.Morlan, H.A.Semken, Jr., S.D.Webb, L.Werdelin, and M.C. Wilson). (1996) Spatial response of mammals to late quaternary environmental fluctuations. Science 272: 1601–1606.

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

Felsenstein, J. (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. Journal of Molecular Evolution 17: 368-376.

Findley, J.S. (1969) Biogeography of Southwestern boreal and desert mammals (ed. by J. K. Jones, Jr.), pp. 113-128. Contributions in Mammalogy, Miscellaneous Publications of the University of Kansas Museum of Natural History 51: 1-428.

Findley, J.S., Harris, A.H., Wilson, D.E. and Jones, C. (1975) Mammals of New Mexico. University of New Mexico Press, Albuquerque. 360 pp.

Findley, J.S. and Jones, C. (1962) Distribution and variation of the voles of the genus *Microtus* in New Mexico and adjacent areas. Journal of Mammalogy 43: 154-166.

Fisher, R.A. (1930) Gene frequencies in a cline determined by selection and diffusion. Biometrics 6: 353-361.

Floyd, C.H., Van Vuren, D.H. and May, B. (2005) Marmots on Great Basin mountaintops: using genetics to test a biogeographic paradigm. Ecology 86: 2145-2153.

Frankham, R. (2000) Genetics and conservation: commentary on Elgar and Clode. Australian Biologist 13:45-54.

Frankham, R. (1998) Inbreeding and extinction: island populations. Conservation Biology 12: 665-675.

Frankham, R. (1997) Do island populations have less genetic variation than mainland populations? Heredity 78: 311-327.

Frey, J.K., Bogan, M.A. and Yates, T. (2007) Mountaintop island age determines species richness of boreal mammals in the American Southwest. Ecography 30: 231-240.

Frey, J.K. (1989) Morphologic variation in the Mexican vole (*Microtus mexicanus*). Masters thesis, Emporia State University, Emporia, KS.

Frey, J.K., Root, J.J., Jones, C.A., Calisher, C.H. and Beatty, B.J. (2002) New records of the Mogollon vole, *Microtus mogollensis* (Mearns, 1890), in southwestern Colorado. Western North American Naturalist 62: 120-123.

Friesen, V. (2000). Introns. Molecular Methods in Ecology.(ed. by A.J. Baker), pp. 274-294. Blackwell.

Fu, Y.X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147**:** 915-925.

Fu, Y.X. and Li, W.H. (1993). Statistical tests of neutrality of mutations. Genetics 133**:** 693-709.

Galbreath, K.E. and Cook, J.A. (2004) Genetic consequences of Pleistocene glaciations for the tundra vole (*Microtus oeconomus*) in Beringia.
Molecular Ecology 13: 135-148.

Galewski, T., Tilak, M., Sanchez, S., Chevret, P., Paradis, E. and Douzery, E.J.P. (2006) The evolutionary radiation of Arvicolinae rodents (voles and lemmings): relative contribution of nuclear and mitochondrial DNA phylogenies. BMC Evolutionary Biology 6: 80.

Garza, J.C., and Williamson, E.G. (2001) Detection of reduction in population size using data from microsatellite loci. Molecular Ecology 10: 305-318.

Gebremedhin, B., Ficetola, G.F., Naderi, S., Rezaei, H.R., Maudet, C., Rioux, D., Luikart, G., Flagstad, O., Thuiller, W. and Taberlet, P. (2009) Combining genetic and ecological data to assess the conservation status of the endangered Ethiopian walia ibex. Animal Conservation 12: 89-100.

Genovart, M., Oro, D., Juste, J. and Bertorelle, G. (2007) What genetics tell us about the conservation of the critically endangered Balearic shearwater? Biological Conservation 137: 283-293.

Goldstein, D.B. and Pollock, D.D. (1997) Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. Journal of Heredity 88: 335-342.

Goudet, J. (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://www.unil.ch/izea/softwares/fstat.html.

Grayson, D.K. (1993) The Desert's Past: a natural prehistory of the Great Basin. Smithsonian Institution Press, Washington, D.C. 356 pp.

Grayson, D.K. and Livingston, S.D. (1993) Missing mammals on Great Basin mountains: Holocene extinctions and inadequate knowledge. Conservation Biology 7: 527-532.

Hadly, E.A., Ramakrishnan, U., Chan, Y.L., van Tuinen, M., O'Keefe, K., Spaeth, P.A. and Conroy, C. J. (2004) Genetic Response to Climatic Change: Insights from Ancient DNA and Phylochronology. PLoS Biology 2: 1-10.

Haffer, J. (1969) Speciation in Amazonian forest birds. Science 165: 131-137.

Hafner, M.S. and Sullivan, R.M. (1995) Historical and ecological biogeography of

    Nearctic pikas (Lagomorpha: Ochotonidae). Journal of Mammmalogy

    76: 302-321.

Hall, E.R. (1981) Mammals of North America. Second edition. John Wiley and Sons,

    New York. 1181 pp.

Hanski, I. (1997) Metapopulation dynamics: from concepts and observations to predictive

    models. Metapopulation biology: ecology, genetics and evolution. (ed. by I.

    Hanski, I. and M. Gilpin), pp. 69-92. Academic Press, Ca.

Harpending, H.C. (1994) Signature of ancient population growth in a low-resolution

    mitochondrial DNA mismatch distribution. Human Biology 66: 591-600.

Harrill, J.R. and Prudic, D.E. (1998) Aquifer systems in the Great Basin region of

    Nevada, Utah, and adjacent states-summary report. U. S. Geological Survey

    Professional Paper 1409-A: A1-A66. Washington, D.C.

Harris, A.H. (1985) Late Pleistocene vertebrate paleoecology of the West. University of

    Texas Press, Austin.

Hawkins, B.A. and Porter, E.E. (2003) Does herbivore diversity depend on plant

    diversity? The case of California butterflies. American Naturalist 161: 40-49.

Hayes, J. P. and Harrison, R.G. (1992) Variation in mitochondrial DNA and the

    biogeographic history of woodrats (*Neotoma*) in the eastern United States.

    Systematic Biology 41: 331–344.

Heaney, L.R. (2000) Dynamic disequilibrium: a long-term, large-scale perspective on the

    equilibrium theory of island biogeography. Global Ecology and

    Biogeography 9: 59-74.

Henry, C.D. (2008) Ash-flow tufts and paleovalleys in northeastern Nevada: implications for Eocene paleogeography and extension in the Sevier hinterland, northern Great Basin. Geosphere 4: 1-35.

Hewiit, G. (2000) The genetic legacy of the Quaternary Ice ages. Nature 405: 907-913.

Hewitt, G. (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. Biological Journal of the Linnaean Society 58: 247-276.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A. (2005a) Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25: 1965-1978.

Hijimans, R.J., Guarino, L., Jarvis, A., O'Brien, R., Mathur, P., Bussink, C., Cruz, M., Barrantes, I. and Rojas, E. (2005b) DIVA-GIS Version 5.2 User Manual.

Ho, S.Y.W. (2007) Calibrating molecular estimates of substitution rates and divergence times in birds. Journal of Avian Biology 38: 409-414.

Hoffman, E.A. and Blouin, M.S. (2004) Historical data refute recent range contractions cause of low genetic diversity in isolated frog populations. Molecular Ecology 13: 271-276.

Holmgren, C.A., Penalba, M.C., Rylander, K.A. and Betancourt, J.L. (2003) A 16,000 14C yr B.P. packrat midden series from the USA-Mexico borderlands. Quaternary Research 60: 319-329.

Huang, Z., Liu, N. and Zhou, T. (2005) A comparative study of genetic diversity of peripheral and central populations of Chukar partridge from northwestern China. Biochemical Genetics 43: 613-621.

Huelsenbeck, J.P. and Ronquist, F. (2001) MRBAYES: Bayesian inference of phylogeny. Bioinformatics 17: 754-755.

Hugall, A., Moritz, C., Moussalli, A. and Stanisic, J. (2002) Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail *Gnarosophia bellendenkerensis* (Brazier 1875). Proceedings of the National Academy of Sciences 99: 6112-6117.

Hutchinson, D.W. (2002) Testing the central/peripheral model: analyses of microsatellite variability in the eastern collared lizard (*Crotaphytus collaris collaris*). American Midland Naturalist 149: 148-162.

Hutchinson, D.W. and Templeton, A.R. (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influence of gene flow and genetic drift on the distribution of genetic variability. Evolution 53: 1898-1914.

Hutchinson, D.W. (2003) Testing the central/peripheral model: analyses of microsatellite variation in the eastern collared lizard (*Crotaphytus collaris collaris*). American Midland Naturalist 149: 148-162.

Irwin, D.M., Kocher, T.D. and Wilson, A.C. (1991) Evolution of the cytochrome *b* gene of mammals. Journal of Molecular Evolution 32: 128-144.

Ishabashi, Y., Yoshinaga, Y. and Saitoh, T. (1999) Poloymorphic microsatellite DNA markers in the field vole *Microtus montebelli*. Molecular Ecology 8: 163-164.

Isaac, J.B., Jones, K.E., Gittleman, J.L. and Purvis, A. (2005) Correlates of species richness in mammals: body size, life history, and ecology. The American Naturalist 165: 600-607.

Jaarola, M., Martinkova, N., Gunduz, I., Brunhoff, C., Zima, J., Nadachowski, A., Amori, G., Bulatova, N.S., Chondropoulos, B., Fraguedakis-Tsolis, S., Gonzalez-Esteban, J., Jose Lopez-Fuster, M., Kandaurov, A.S., Kefelioglu, H., da Luz Mathias, M., Villate, I. and Searle, J.B. (2004) Molecular phylogeny of the speciose vole genus *Microtus* (Arvicolnae, Rodentia) inferred from mitochondrial DNA sequences. Molecular Phylogenetics and Evolution 33: 647-663.

Jass, C.N. (2004) Mid to Late Holocene mammals from Bighorn Cave, Arizona. The Southwestern Naturalist 49: 75-82.

Judd, S.R. (1980) Observations of the chromosome variation in *Microtus mexicanus* (Rodentia: Microtinae). Mamm. Chromosome Newsletter 21: 110-113.

Kalinowski, S.T. (2004) Counting alleles with rarefaction: private alleles and hierarchical sampling designs. Conservation Genetics 5: 539-543.

Karron, J.D. (1987) A comparison of levels of genetic polymorphism and self incompatibility in geographically restricted and widespread plant congeners. Evolutionary Ecology 1: 47-58.

Keller, B.L. (1985) Reproductive Patterns. I*n* Biology of New World *Microtus* (ed. by R.H. Tamarin). American Society of Mammalogists Special Publication 8: 725-778.

Kimura, M. (1983) The neutral theory f molecular evolution. Cambridge University Press, Cambridge. 388 pp.

Knowles, L.L., Carstens, B.C. and Keat, M.L. (2007) Coupling genetic and eological models to examine how past populations distributions to divergence. Current Biology 17: 940-946.

Lacy, R.C. (2000) Structure of the VORTEX simulation model for population viability analysis. Ecological Bulletin 48: 191-203.

Lamb, T., Jones, T.R. and Wettstein, P.J. (1997) Evolutionary dynamics and phylogeography of tassel-eared squirrels (*Sciurus aberti*). Journal of Mammalogy 78: 117-133.

Lambin, X. Aars, J., Piertney, S.B. and Telfer, S. (2005) Inferring pattern and process in small mammal metapopulations: insights from ecological and genetic data. Ecology, genetics and evolution of metapopulations. (ed. by I. Hanski and O.E. Gaggiotti), pp. 515-540. Elsevier Academic Press, Burlington, Mass.

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

Lawlor, T.E. (1998) Biogeography of Great Basin mammals: paradigm lost? Journal of Mammalogy 79: 1111-1130.

Leberg, P.L. (2002) Estimating allelic richness: effects of sample size and bottlenecks. Molecular Ecology 11: 2445-2449.

Lee, M.R. and Endler, F.P.B. (1977) Karyotypes of eight species of Mexican rodents (Muridae). Journal of Mammalogy 54: 479-487.

Leon-Paniagua, L., Navarro-Siguenza, A.G., Hernandez-Banos, B.E. and Morales, J.C. (2007) Diversification of the arboreal mice of the genus *Habromys* (Rodentia: Cricetidae: Neotominae) in the Mesoamerican highlands. Molecular Phylogenetics and Evolution 42: 653-664.

Lesica, P. and Allendorf, F.W. (1995) When are peripheral populations valuable for conservation? Conservation Biology 9:753-760.

Lessa, E.P., Cook, J.A. and Patton, J.L. (2003) Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. Proceedings of the National Academy of Sciences 100: 10331-10334.

Liu, C.R., Berry, P.M., Dawson, T.P. and Pearson, R.P. (2005) Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28: 385-393.

Lomolino, M.V., Brown, J.H and Davis, R. (1989) Island biogeography of montane forest mammals in the American southwest. Ecology 70: 180-194.

Lomolino, M.V. and Channel, R. (1995) Splendid isolation: patterns of range collapse in endangered mammals. Journal of Mammalogy 76: 335-347.

MacArthur, R.H. and Wilson, E.O. (1967) The theory of island biogeography. Princeton University Press, Princeton, New Jersey, USA.

Mace, G.M. and Lande, R. (1991) Assessing extinction threats: toward a reevaluation of IUCN threatened species categories. Conservation Biology 5: 148-157.

Maddison, W.P. and Maddison, D.R. (2009) Mesquite: a modular system for evolutionary analysis. Version 2.6. (http://mesquiteproject.org).

Mantel, N. (1987) The detection of disease clustering and a generalized regression approach. Cancer Research 27: 209-220.

Matthee, C.A., Eick, G., Willows-Munro, S., Montgelard, C., Pardini, A.T. and Robinson, T.J. (2006) Indel evolution of mammalian introns and the utility of non-coding nuclear markers in eutherian phylogenetics. Molecular Phylogenetics and Evolution 42: 827-837.

Maxson, J. and Tikoff, B. (1996) Hit and run collision model for the Laramide orogeny, western United States. Geology 24: 968-972.

McCormack, J.E., Bowen, B.S. and Smith, T.B. (2008) Integrating paleoecology and genetics of bird populations in two sky island archipelagos. BMC Biology 6: 23.

McDonald, K.A. and Brown, J.H. (1992) Using montane mammals to model extinctions due to global change. Conservation Biology 6: 409-415.

Medrano, J.F., Aasen, E. and Sharrow, L. (1990) DNA extraction from nucleated red blood cells. Biotechniques 8: 43.

Messier, W. and Stewart, C.B. (1995) Episodic adaptive evolution of primate lyzozymes. Nature 9: 151-154.

Metzger, D. G., et al. (1973) Geohydrology of the Parker-Blythe-Cibola area, Arizona and California. U. S. Geological Survey Professional Papers 486-G. 130pp.

Mittermeier, R.A., Robles, G.P. and Hoffman, M. (2005) Hotspots revisited: Earth's Biologically Richest and Most Endangered Terrestrial Ecoregions. Conservation International, Washington, D.C. 392 pp.

Modi, W.S. (1987) Phylogenetic analyses of chromosomal banding patterns among the Nearctic Arvicolidae (Mammalia: Rodentia). Systematic Zoology 36: 109-136.

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

Morafka, D.J. (1977) A biogeographic analysis of the Chihuanhuan Desert through its herptofauna. Dr. W. Junk, B.V.M. Publishers, The Hague, The Netherlands.

Nei, M. (1987) Molecular Evolutionary Genetics. Columbia University Press, New York, USA.

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

Osborn, G. and Bevis, K. (2001) Glaciation in the Great Basin of the western United States. Quaternary Science Reviews 20: 1377-1410.

Pacheco, N.M., Congdon, B.C. and Friesen, V.L. (2002) The utility of nuclear introns for investigating hybridization and genetic introgression: a case study involving *Brachyramphus* murrelets. Conservation Genetics 3: 175-182.

Patterson, B.D. (1982) Pleistocene vicariance, montane islands, and the evolutionary divergence of some chipmnuks. Journal of Mammalogy 63: 387-398.

Peat, H.J., Clarke, A. and Convey, P. (2007) Diversity and biogeography of the Antarctic flora. Journal of Biogeography 34: 132-146.

Peppers, L.L. and Bradley, R.D. (2000) Molecular systematics of the genus *Sigmodon*. Journal of Mammalogy 81: 332-343.

Peterson, A.T., Ortega-Huerta, M.A., Bartley, J., Sanchez-Cordero, V., Soberon, J., Buddemeier, R.H. and Stockwell, D.R.B. (2002) Future projections for Mexican faunas under global climate change. Nature 416: 626-628.

Petit, R., Mousadik, A. and Pons, O. (1998) Identifying populations for conservation on the basis of molecular markers. Conservation Biology 12: 844-855.

Phillips, S.J. and Dudik, M. (2008) Modeling of species distributions with Maxent: new extensions and comprehensive evaluation. Ecography 31: 161-175.

Philips, S.J., Anderson, R.P. and Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. Ecological Modeling 190: 231-259.

Pimm, S.L., Russell, G.J. and Brooks, T.M. (1995) The future of biodiversity. Science 269: 347-350.

Plume, R.W. (1995) Hydrogeologic framework of the Great Basin region of Nevada, Utah, and adjacent states. U.S. Geological Survey Professional Paper 1409-B, Washington, D.C., 64 pp.

Polhemus, D.A. and Polhemus, J.T. (2002) Basins and Ranges: the biogeography of aquatic true bugs (Insecta: Heteroptera) in the Great Basin. Great Basin Aquatic Systems History. Smithsonian Contribution to the Earth Sciences Vol. 33. (ed.by R. Hershler, D.B. Madsen and D.R. Currey), pp. 235-254. Smithsonian Institution Press, Washington, D.C.

Posada, D. and Crandall, K.A. (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14: 817-818.

Pruett, C.L. and Winkler, K. (2008) The effects of sample size on population genetic diversity in song sparrows *Melospiza melodia*. Journal of Avian Biology 39: 252-256.

Raxworthy, C.J., Pearson, R.G., Rabibisoa, N., Rakotondrazafy, A.M., Ramanamanjato, J.B., Raselimanana, A.P., Wu, S., Nussbaum, R.A. and Stone, D.A. (2008) Extinction vulnerability of tropical montane endemism from warming and upslope displacement: a preliminary appraisal for the highest massif in Madagascar. Global Change Biology 14: 1-18.

Reheis, M.C., Sarna-Wojicki, A.M., Reynolds, R.L., Repenning, C.A. and Miflin, M.D. (2002) Pliocene to Middle Pleistocene lakes in the western Great Basin: ages and connections. Great Basin Aquatic Systems History. Smithsonian Contribution to the Earth Sciences Vol. 33. (ed.by R. Hershler, D.B. Madsen and D.R. Currey), pp. 53-120. Smithsonian Institution Press, Washington, D.C.

Reiman, B.E. and McIntyre, J.D. (1993) Demographic and Habitat Requirements for Conservation of Bull Trout. General Technical Report INT-302. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station, Boise, ID. 38pp.

Repenning, C.A. (1990) Of mice and ice in the Late Pliocene of North America. Arctic 43: 314-323.

Riddle, B.R., Dawson, M.N., Hadly, E.A., Hickerson, M.J., Mantooth, S.J. and Yoder, A.D. (2008) The role of molecular genetics in sculpting the future of integrative biogeography. Progress in Physical Geography 32: 173-202.

Rogers, A.R. (1995) Genetic evidence for a Pleistocene population expansion. Evolution 49: 608-615.

Rowe, K.C., Heske, E.J., Brown, P.W. and Paige, K.N. (2004) Surviving the ice: Northern refugia and post-glacial colonization. Proceedings of the National Academy of Sciences 101: 10355-10359.

Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X. and Rozas, R. (2003) DnaSP: DNA polymorphisms analyses by the coalescent and other methods. Bioinformatics 19: 2496-2497.

Sagarin, R.D., Gaines, S.D. and Gaylord, B. (2006) Moving beyond assumptions to understand abundance distributions across the ranges of species. Trends in Ecology and Evolution 21: 524-530.

Sagarin, R.D. and Gaines, S.D. (2002) The 'abundant centre' distribution: to what extent is it a biogeographic rule? Ecology Letters 5: 137-147.

Schwartz, M.K., Mills, L.S., Ortega, Y., Ruggiero, L.F. and Allendorf, F.W. (2003) Landscape location affects genetic variation of Canada Lynx (*Lynx canadensis*). Molecular Ecology 12: 1807-1816.

Sefc, K.M., Payne, R.B. and Sorenson, M.D. (2003) Microsatellite amplification from museum feather samples: effects of fragment size and template concentration on genotyping errors. The Auk 4: 982-989.

Selkoe, K.A. and Toonen, R.J. (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecology Letters 9: 615-629.

Sheperd, L.D., Millar, C.D., Ballard, G., Ainley, C. and Lambert, D.M. (2005) Microevolution and mega-icebergs in the Antarctic. Proceedings of the National Academy of Sciences 102: 16717-16922.

Slatkin, M. and Maddison, W.P. (1989) A cladistic measure of gene flow inferred from the phylogeny of alleles. Genetics 123: 603-613.

Slatkin, M. (1985) Gene flow in natural populations. Annual Review of Ecology and Systematics 16: 393-430.

Smith, M.F. and Patton, J.L. (1993) The diversification of South American rodents: evidence from the mitochondrial sequence data for the Akodontine tribe. Biological Journal of the Linnean Society 50: 149-177.

Spielman, D., Brook, B.W. and Frankham, R. (2004) Most species are not driven to extinction before genetic factors impact them. Proceedings of the National Academy of Sciences 101: 15261-15264.

Stacey, C.A. and Post, D.M. (2009) Effects of disturbance by humans on small mammals in a Chihuahuan Desert ecosystem. The Southwestern Naturalist 54: 272-278.

Steele, C. A. and Storfer, A. (2006) Coalescent-based hypothesis testing supports
multiple Pleistocene refugia in the Pacific Northwest for the Pacific giant
salamander (*Dicamptodon tenebrosus*). Molecular Ecology 15: 2477-2487.

Steppan, S.J., Adkins, R.M. and Anderson, J. (2004) Phylogeny and divergence-date
estimates of rapid radiations in muroid rodents based on multiple nuclear genes.
Systematic Biology 53: 533-553.

Sunnucks, P. (2000) Efficient genetic markers for population biology. Trends in Ecology
and Evolution 15: 199-203.

Sullivan, J., Arellano, E. and Rogers, D.S. (2000) Comparative phylogeography of
Mesoamerican highland rodents: concerted versus independent response to past
climatic fluctuations. The American Naturalist 155: 755-768.

Swenson, N. (2008) The past and future influence of geographic information systems on
hybrid zone, phylogeographic and speciation research. Journal of Evolutionary
Biology 21: 421-434.

Swofford, D.L. (2002) PAUP*: phylogenetic analysis using parsimony (*and other
methods). Version 4.0b10. Sinauer Associates, Inc. Publishers. Sunderland,
Massachusetts.

Szpiech, Z.A., Jakobsson, M. and Rosenberg, N.A. (2008) ADZE: a rarefaction approach
for counting alleles private to combinations of populations. Bioinformatics 24:
2498-2504.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA
polymorphism. Genetics 123: 585-595.

Tamura K, Dudley, J., Nei, M. and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596-1599. (http://www.kumarlab.net/publications)

Toledo, V.M. (1982) Pleistocene changes of vegetation in tropical Mexico. Biological Diversification in the Tropics (ed. by G.T. Prance), pp. 93-111. Columbia University Press, New York.

Trammell, E.J., Berry, K., Bassett, S. and Sada, D.W. (2008) Distribution and recovery of vegetational assemblages in Ash Meadows National Wildlife Refuge. Nevada. The Southwestern Naturalist 53: 326-324.

Van Devender, T.R. (1987) Holocene vegetation and climate in the Puerto Blanco Mountains, southwestern Arizona. Quaternary Research 27: 51-72.

Vawter, L. and Brown, W.M. (1986) Nuclear and mitochondrial DNA comparisons reveal extreme rate variation in the molecular clock. Science 234: 194-196.

Vignieri, S.N. (2005) Streams over mountains: influence of riparian connectivity on gene flow in the Pacific jumping mouse (*Zapus trinotatus*). Molecular Ecology 14: 1925-1937.

Vrba, E.S. (1993) Turnover-pulses, the Red Queen, and related topics. American Science 293: 418-452.

Wagner, R.S., Miller M.P., Crisafulli, C.M. and Haig, S.M. (2005) Geographic variation, genetic structure, and conservation unit designation in the Larch Mountain salamander (*Plethodon larselli*). Canadian Journal of Zoology 83: 396–406.

Walker, D. and Avise, J.C. 1(998) Principles of phylogeography as illustrated by
 freshwater terrestrial turtles in the southeastern United States. Annual Review of
 Ecology and Systematics 29: 23-58.

Walser, B. and Heckel, G. (2008) Microsatellite markers for the common vole (*Microtus
 arvalis*) and their cross-species utility. Conservation Genetics 9: 479-481.

Wegener, A. (1966). The origins of continents and oceans. Dover Publications, NY.

Weinrich, D.M. (2001) The rates of molecular evolution in rodent and primate
 mitochondrial DNA. Journal of Molecular Evolution 52: 40-50.

Weir, B.S. and Cockerham, C.C. (1984) Estimating F-statistics for the analysis of
 population structure. Evolution 38: 1358-1370.

Wilhelm, D.E. (1982) Zoogeographic and evolutionary relationships of selected
 populations of *Microtus mexicanus*. Occasional Papers of the Museum of Texas
 Tech University 75: 1-29.

Wright, S. (1943) Isolation by genetics. Genetics 28: 114-138.

Yang, Z. and Nielsen, R. (1998) Synonymous and nonsynonymous rate variation in
 nuclear genes of mammals. Journal of Molecular Evolution 46: 409-418.

Young, J.A. and B.A. Sparks, B.A. (1985) Cattle in the cold desert. Utah State University
 Press, Logan, UT.

Zakrzewski, R.J. (1985) The fossil record. Biology of New World *Microtus* (ed. by
 R.H. Tamarin), pp. 1–51. Special Publication No. 8, American Society of
 Mammalogists. Shippensburg, Pa.

Zwickl, D.J. (2006) Genetic algorithm approaches for the phylogenetic analysis of large

biological sequence datasets under the maximum likelihood criterion. Ph.D.

dissertation. The University of Texas at Austin.