

California State University, San Bernardino

CSUSB ScholarWorks

Theses Digitization Project

John M. Pfau Library

2006

Population genetics and phylogeography of the pygmy nuthatch in Southern California

Thomas Alan Benson

Follow this and additional works at: <https://scholarworks.lib.csusb.edu/etd-project>



Part of the [Population Biology Commons](#)

Recommended Citation

Benson, Thomas Alan, "Population genetics and phylogeography of the pygmy nuthatch in Southern California" (2006). *Theses Digitization Project*. 2973.

<https://scholarworks.lib.csusb.edu/etd-project/2973>

This Thesis is brought to you for free and open access by the John M. Pfau Library at CSUSB ScholarWorks. It has been accepted for inclusion in Theses Digitization Project by an authorized administrator of CSUSB ScholarWorks. For more information, please contact scholarworks@csusb.edu.

POPULATION GENETICS AND PHYLOGEOGRAPHY OF THE
PYGMY NUTHATCH IN SOUTHERN CALIFORNIA

A Thesis
Presented to the
Faculty of
California State University,
San Bernardino

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
in
Biology

by
Thomas Alan Benson


June 2006

POPULATION GENETICS AND PHYLOGEOGRAPHY OF THE
PYGMY NUTHATCH IN SOUTHERN CALIFORNIA


A Thesis
Presented to the
Faculty of
California State University,
San Bernardino

by
Thomas Alan Benson
June 2006


Approved by:



Anthony Metcalf, Ph.D., Chair, Biology



Richard Fehn, Ph.D.



James Ferrari, Ph.D.

June 7, 2006

Date

ABSTRACT

I examined the population genetic structure, phylogeography, and subspecies structure of pygmy nuthatch (*Sitta pygmaea*) populations in southern California using mitochondrial DNA sequence data. Population genetic analyses indicate that pygmy nuthatch populations fragmented in the disjunct mountain ranges of southern California exhibit low but significant levels of genetic differentiation. These results correspond to phylogeographic analyses that further suggest that the isolation of pygmy nuthatch populations occurred recently, most likely at the beginning of the current interglacial period (~10,000-20,000 years ago). The mitochondrial DNA data also support a subspecific break between *S. p. melanotis* and *S. p. leuconucha* farther north than proposed subspecific breaks that are based upon morphological characters. These results are similar to patterns found in other bird populations exhibiting similar distributions in California and western North America.

ACKNOWLEDGEMENTS

This thesis would not have been possible without the support of various individuals and organizations. I would like to take this opportunity to thank those people.

I thank my advisor, Tony Metcalf, who provided constant support in all aspects and through all phases of this work, and whom I also consider to be a good friend. I also thank Jim Ferrari and Rich Fehn for their careful critique of this thesis, by which I have produced a better final product.

I thank the staff at the San Bernardino County Museum, Gerald Braden, Mike Rathbun, and Todd Hoggan, for assistance in collecting specimens. I also particularly thank Gene Cardiff, not only for assistance in collecting and preparing specimens, but also for teaching me most of what I know about birds.

I thank the staff at the Louisiana State University Museum of Natural Science, Steve Cardiff, Donna Dittman, Robb Brumfeld, and Chris Witt, for facilitating the tissue grant of both pygmy and brown-headed nuthatch specimens for this research.

I thank my lab mates, Ivan Phillipson and George Walker, for providing technical support and answers to questions they had already figured out on their own.

I thank my fellow graduate students and friends, particularly Ron Spier, Ivan Phillipson, and Heidi Contreras, for their moral support and providing much needed breaks from research, coursework, and teaching.

I thank my family and friends, MaryAnn and Susie Benson, Sean, Jean, Blake, Ryan and Josh Guarino, Bobb Elliott, and Christina Chandler, for their love, support, and understanding and for providing an easy escape from the stress of graduate school.

And finally I thank the following organizations for their financial support without which this research could not have been conducted: Los Angeles Audubon Society Ralph W. Schreiber Student Research Award; San Bernardino Valley Audubon Society Scientific Research Grant; Southern California Academy of Sciences Student Research Award; and multiple California State University San Bernardino (CSUSB) Associated Students Inc., Student Research Grants. This research was also supported by a National Science Foundation MRI grant (DBI#0321209) to A. Metcalf and J. Ferrari and a CSUSB faculty grant to A. Metcalf.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	viii
LIST OF FIGURES	ix
CHAPTER ONE: INTRODUCTION	1
Genetic Differentiation	3
Mitochondrial DNA in Population Studies	7
Sky Island Biogeography	15
Pygmy Nuthatch as a Model Organism	18
Population Genetics and Phylogeography of Pygmy Nuthatches	22
CHAPTER TWO: MATERIALS AND METHODS	27
Data Collection	27
Sequence Analysis	35
Population Genetics	37
Phylogeography	38
Subspecies Limits	42
CHAPTER THREE: RESULTS	45
Sequence Analysis	45
Population Genetics	50
Phylogeography	51
Subspecies Limits	60

CHAPTER FOUR: DISCUSSION	62
Sequence Analysis	63
Population Genetics	63
Phylogeography	65
Subspecies Limits	68
Conservation Implications	73
Conclusions	74
APPENDIX A: HAPLOTYPE LISTS	76
APPENDIX B: DISTANCE MATRICES	98
APPENDIX C: INPUT FILES	100
REFERENCES	113

LIST OF TABLES

Table 1.	Sitta pygmaea Specimen Information.....	28
Table 2.	Variable Sites within 3045-Base Pair Sitta pygmaea Haplotypes.....	46
Table 3.	Gene Statistics for 3045-Base Pair Haplotypes.....	47
Table 4.	Variable Sites within 1639-Base Pair Sitta pygmaea Haplotypes.....	48
Table 5.	Gene Statistics for 1639-Base Pair Haplotypes.....	49
Table 6.	Estimates of Intrapopulation Variability.....	50
Table 7.	Φ_{ST} Values for Analysis of Molecular Variance Comparisons.....	51
Table 8.	Distribution of Sitta pygmaea Haplotypes and Phylogenetic Clades.....	56

LIST OF FIGURES

Figure 1.	<i>Sitta pygmaea</i> Range Map.....	19
Figure 2.	<i>Sitta pygmaea</i> Southern California Range Map.....	21
Figure 3.	Primer Map.....	32
Figure 4.	<i>Sitta pygmaea</i> Subspecific Break Map.....	43
Figure 5.	Bootstrapped Consensus Parsimony Tree.....	53
Figure 6.	Bootstrapped Consensus Neighbor-Joining Tree.....	54
Figure 7.	Consensus Bayesian Tree.....	55
Figure 8.	Minimum Spanning Haplotype Network.....	58
Figure 9.	Nested Cladogram.....	59

CHAPTER ONE
INTRODUCTION

The distribution of genetic variation in geographically structured populations is of fundamental interest to evolutionary and population biologists. For thousands of years biologists have described the variation present in nature and classified organisms into discrete units. In the 1900s, following the development of the Linnaean system of classification, Darwin's theory of evolution, and the biological species concept, biologists began classifying organisms according to their evolutionary relationships. Many questions remained, however, regarding the definition of a species, and whether distinct geographic units were simply local races or reproductively isolated species. Often these questions could not be answered by observation of phenotypic characteristics. The ability to analyze neutral genetic markers (sequences of DNA not influenced by selection) has allowed biologists to measure the evolutionary relationships among populations directly. Observed levels of genetic differentiation among populations can provide a direct link to their evolutionary history. Low genetic differentiation among populations

suggests high levels of gene flow in the present or recent past, indicating connectivity between seemingly disjunct populations. Alternatively, high genetic differentiation among populations suggests a lack of recent gene flow, which may lead to complete isolation of populations and eventually speciation events.

Genetic differentiation is frequently observed in species exhibiting geographic structure (Avice 2000). Populations of such species may be separated by oceans, continents, mountain ranges, valleys, or numerous other geographic barriers. The extent to which these barriers are obstacles to gene flow, along with factors such as effective population size and time since divergence, influences the distribution of genetic variation and the degree of genetic differentiation among populations.

I propose to study the genetic differentiation among southern California populations of the pygmy nuthatch (*Sitta pygmaea*), a small songbird exhibiting a high level of geographic structure and with limited dispersal capabilities (Norris 1958, Kingery and Ghalambor 2001). Populations of pygmy nuthatches occupying disjunct mountain ranges that have been isolated for tens of thousands of years would be expected to show levels of genetic

differentiation consistent with recent divergence. I will examine the genetic structure within and among pygmy nuthatch populations in seven disjunct mountain ranges in southern California. These results will provide insight into the evolutionary ecology and population biology of this species, and allow inferences into the historical biogeography of pygmy nuthatches in southern California. Additionally, the study area encompasses the range of two subspecies of pygmy nuthatch, *S. p. melanotis* and *S. p. leuconucha*, the boundary of which is disputed when based upon analysis of morphological characters (Norris 1958, Grinnell and Miller 1944). This study may provide genetic evidence to corroborate one of these morphological subspecies designations.

Genetic Differentiation

Assessment of variation is the basis of any genetic study in population biology. Genetic differentiation provides information about species evolution and population history. Such data can be used to infer the degree of connectivity among populations via migration, evolutionary relationships of populations within species, and can shape decisions regarding biological conservation issues.

Beginning in the 1970s, an effort was made to correlate observed genetic differentiation with evolutionary, ecological, and biogeographic factors of the species being studied (Avice 1998). This effort, termed intraspecific phylogeography (Avice et al. 1987), compares phylogenetic information from a species to that species' present and historical geographic distribution. Simply stated, the process begins with the identification of genetic markers (e.g., haplotypes of DNA sequences), which are obtained and analyzed for a number of individuals within the population. Gene trees (phylogenetic trees showing the relationship of haplotypes within the population) are produced from the analysis, and these phylogenetic trees are then compared with biogeographic information (i.e., morphological subspecies, geographic subpopulations) to estimate historical gene flow and evolutionary forces and relationships within the species.

The distribution of genetic variation within a species can range from completely unstructured to highly structured (Avice 2000). A genetically unstructured population generally has little geographic subdivision and relatively high gene flow across the range of the population. Highly vagile organisms often form genetically unstructured

populations. For example, several studies have found limited intraspecific genetic differentiation in marine fish species: the combination of widespread distribution, contiguous marine habitat, and high dispersal capability promotes genetic homogeneity (Graves et al. 1984, Gyllensten 1985, Ward et al. 1994).

In contrast, a genetically structured population is often geographically subdivided such that the population consists of several subpopulations among which migration (i.e., gene flow) is restricted to some extent. Many organisms display some degree of genetic population structure, an effect commonly created by the discontinuous distribution of suitable habitat (Hartl and Clark 1997). Subpopulations are separated by biogeographic barriers that can take the form of topographic features, distance, or other abiotic or biotic factors, but vary depending on the resource requirements and evolutionary ecology of the organism being considered. Populations of freshwater fish species occupying different drainages, for example, tend to exhibit high levels of genetic differentiation as a result of terrestrial and marine habitats presenting barriers to migration (Bermingham and Avise 1986, Gyllensten 1985, Ward et al. 1994). Geographic barriers, however, are not

necessarily isolating mechanisms. Although subpopulations are discontinuous, gene flow may still be maintained by individuals migrating or dispersing across these barriers. It is the level of genetic differentiation among subpopulations that is of primary interest to population biologists conducting phylogeographic studies.

Wright (1943) developed a model to describe the effects of migration and genetic drift on the distribution of genetic variation in discontinuous subpopulations. His "island model" proposes that genetic differentiation will occur among discontinuous subpopulations, or continuous subpopulations with short-range dispersal, when the effect of migration is limited (assuming effective population size and selection are sufficiently small). Migration, the movement of an individual from one subpopulation to another, is a homogenizing force that maintains genetic similarity among subpopulations. Drift, or random change in gene frequency, is an opposing evolutionary force that acts to differentiate subpopulations. Migration and drift interact to produce an equilibrium level of genetic variation among subpopulations. When even a few migrants are exchanged among subpopulations per generation, the influence of gene flow is strong enough that the effect of

drift is negated regardless of population size (Wright 1931). In the absence of significant gene flow, however, genetic drift becomes the primary force of evolution acting on subpopulations. This, coupled with the random accumulation of mutations over time, results in genetic differentiation among subpopulations.

This model can be applied to any species that exhibits an island distribution, including those restricted to sky islands. Sky islands consist of similar habitat types distributed on mountains or mountain ranges at high elevation. Because of the geographic distribution and elevational changes often associated with mountain ranges, these habitat patches are often discontinuous. Sky islands are analogous to oceanic islands, except that instead of mountains rising from the ocean floor above sea level, mountains rise above a specific minimum elevation at which certain habitats exist; intervening lower elevations represent oceans of unsuitable habitat that present potential barriers to migration and gene flow.

Mitochondrial DNA in Population Studies

Phylogenetic studies at the intraspecific level attempt to reveal genetic relationships among populations

of a single species or species complex, usually with regard to subspecific designations, morphological characteristics, biogeographic structuring, or a combination of the above. The primary approach used to reveal these relationships is now the direct analysis of DNA. DNA studies have high resolution in determining genetic differentiation, and in some cases show variation and differentiation not revealed by previous methods such as allozyme electrophoresis (Metcalf et al. 2001, Burton and Lee 1994).

Mitochondrial DNA (mtDNA) has proven particularly useful at the intraspecific level. The advantage of using mitochondrial DNA in intraspecific phylogeographic studies stems from the combination of its haploid chromosome number, maternal inheritance, lack of significant recombination, and relatively fast rate of evolution. MtDNA is a circular genome found in the mitochondria of all animal cells, and it is passed to offspring only by females (but see below). These two qualities yield an effective mitochondrial population size that can be one-fourth the effective population size of the nuclear genome (Birky et al. 1983). Decreased effective population size increases the rate of coalescence (the point at which genealogies converge on a common ancestor), which means that a gene

tree based on mtDNA has a better likelihood of reflecting the true species (or population) tree than does a gene tree based on nuclear DNA (Moore 1995). Additionally, because mtDNA rarely recombines, it tends to generate cleaner gene trees than nuclear DNA in which recombination can obscure branching structure (Moore 1995). Another advantage of mtDNA is its relatively quick rate of evolution, which is 5-10 times faster than in single copy nuclear genes (Brown et al. 1979). This makes mtDNA useful for inferring phylogenies of recently diverged taxa or intraspecific population structure. These qualities give mtDNA a distinct advantage over nuclear DNA in inferring the phylogenies within or among closely related genera, species, or subspecies.

MtDNA does have potential limitations that can lead to incorrect phylogenetic inferences if undetected or not properly addressed. Homoplasy (similarity of sequences in different organisms due to convergent evolution rather than shared common ancestry) due to saturation at substitution sites makes mtDNA less effective for analyzing taxa that diverged more than approximately 10 million years ago (Moore and DeFilippis 1997). Nuclear homologues of mtDNA (Numts) are short sequences of the mitochondrial genome

that have been transposed into the nuclear genome, making it possible to amplify and analyze nuclear rather than mtDNA. This is potentially significant because nuclear homologues evolve more slowly than mtDNA (Sorenson and Quinn 1998). Amplification of Numts can be greatly reduced by using purified mtDNA during PCR reactions, or by extracting whole genomic DNA from tissues rich in mtDNA; careful analysis of resultant sequences for anomalies can also be helpful in the detection of Numts (Sorenson and Quinn 1998). Heteroplasmy (the coexistence of multiple mtDNA genotypes within an individual) in the mitochondrial genome can arise via two pathways: paternal leakage or mutation. Paternal leakage involves the transmission of mtDNA from males to their offspring. While this phenomenon is mostly detected in interspecific crosses, it has been recorded intraspecifically in a few species including fruit flies (Kondo et al. 1990), mussels (Hoeh et al. 1991, Zouros et al. 1992), and birds (Kvist 2003). Heteroplasmy can also arise through mutation via replication error. Recombination has been demonstrated in some animal mtDNA (reviewed by Rokas et al. 2003); studies in birds, however, have failed to detect evidence of mtDNA recombination (Berlin and Ellegren 2001, Berlin et al. 2004). Both

heteroplasmy and recombination are believed to be rare and have little impact on phylogeographic studies (Avise et al. 1987, Berlin et al. 2004). Sex biased dispersal may also introduce error into mtDNA-based studies. Because mtDNA is maternally inherited, mtDNA may fail to detect gene flow at the nuclear level when females exhibit philopatry (limited dispersal from the birth site) and males are the dispersing sex. This would reduce the effectiveness of mtDNA over nuclear DNA in inferring species (or population) trees from gene trees (Hoelzer 1997). Thorough research of the life history of the organism and selection of the appropriate molecular marker in accordance with its evolutionary ecology can avoid this type of error. Overall, mtDNA is the most effective genetic marker for inferring phylogenies among recently diverged groups.

Two mitochondrial markers have been widely used in phylogeographic studies: cytochrome *b* (cyt *b*; a protein-coding gene) and the control region (a non-coding region). Cyt *b* has been the most commonly used mitochondrial marker in avian sequence-based studies (Moore 1997). This gene is particularly attractive for phylogenetic studies because most nucleotide substitutions are synonymous at this level. This allows sequences from individuals in different groups

to be easily aligned and compared using amino acid sequences, and allows for informative nucleotide sequence variation at third-position codon sites (Moore 1997). For example, Kircham et al. (2000) used *cyt b* data to show subdivision of cave swallow (*Petrochelidon fulva*) populations in North America into two distinct clades, and that a population in Ecuador is best considered a separate species (*P. rufocollaris*), consistent with recent classification changes based on morphology. Prychitko and Moore (1999) employed *cyt b* to analyze the relationships among 8 genera of woodpeckers (subfamily Picinae), demonstrating its usefulness at higher taxonomic levels. The value of *cyt b* in phylogenetic studies has been well established over the last decade.

The control region is a rapidly evolving portion of the mitochondrial genome that has been employed in a number of phylogenetic studies. For example, analysis of control region sequences have been used to demonstrate restricted gene flow between sedentary and migratory populations of prairie warblers (*Dendroica discolor*) in the eastern United States (Buerkle 1999), Pleistocene origins of disjunct populations of azure-winged magpies (*Cyanopica cyanus*) in Europe and Asia (Fok et al. 2002), and bottlenecking and

recent expansion of red knots (*Calidris canutus*) worldwide (Baker et al. 1994). The utility of the control region in these studies arises from relatively low selective constraints (Baker and Marshall 1997). The control region is divided into Domains I, II, and III based on degree of variability and base composition. Domain II and a short sequence (CSB-I) in Domain III are highly conserved across a wide range of taxa, while Domain I and the remainder of Domain III show large amounts of variation. Approximately two-thirds of the nucleotide sites within the control region are potentially phylogenetically informative, compared to approximately one-third in cytochrome *b*. Studies of human mtDNA have revealed substitution rates in the control region ranging from 2.8 to 5 times faster than the rest of the mitochondrial genome (Cann et al. 1984, Aquadro and Greenberg 1983), and comparisons of red knot mtDNA have shown control region substitution rates approximately 2.6 times faster than *cyt b* in that species (Baker and Marshall 1997). As a result, the control region may reveal differentiation among populations or recently diverged taxa that is undetected by *cyt b* or other mitochondrial or nuclear markers.

Numerous phylogenetic studies have sequenced hypervariable regions of mtDNA to analyze genetic relationships among populations (Avice 2000). For example, Idaghdour et al. (2004) provided evidence of genetic differentiation in the houbara bustard (*Chlamydotis undulata*). Populations in the Middle East (*C. u. macqueenii*) were found to be significantly differentiated from populations in North Africa (*C. u. fuertaventurae* and *C. u. undulata*). The authors were also able to infer that the population on the Canary Islands (*C. u. fuertaventurae*) had been colonized twice, but has been relatively isolated since the second colonization about 20,000 years ago. Other mtDNA sequence-based studies have shown a significant lack of differentiation among populations. Pearce et al. (2002) discovered little geographic subdivision (attributed to high gene flow) among ancient murrelet (*Synthliboramphus antiquus*) populations sampled across the northern Pacific Ocean from China to British Columbia. Kvist et al. (2001) found no significant genetic differentiation among willow tit subspecies (*Parus montanus*) across northern Eurasia. The same study, however, also revealed significant genetic differentiation between the northern subspecies and the more southerly, conspecific songar tit (*P. m. songarus* and

P. m. affinis). These studies bear out the utility of mtDNA in phylogeographic studies, and attest to the range in degree of structure observed in vertebrate populations.

Sky Island Biogeography

Sky islands provide an opportunity to study the effects of genetic differentiation on recently diverged taxa. The coniferous forests and alpine habitats of western North America were much more widespread and contiguous during the last ice age than they are today (Pielou 1991). These habitats underwent several cycles of expansion (downward shift) and contraction (upward shift) during the cooler glacial and warmer interglacial periods in the Pleistocene Epoch approximately 1,000,000 to 10,000 years ago (Dorf 1976, Pielou 1991). Habitat contractions resulted in the creation of sky islands of coniferous forest. The current level of genetic differentiation among sky island populations therefore depends on the degree of isolation during interglacial periods and the extent of population admixture during glacial periods. Sky island populations in western North America potentially may have been geographically isolated for anywhere between

approximately 1,000,000 and 10,000 years (Dorf 1976, Pielou 1991).

Few studies have been conducted on the intraspecific phylogeography of sky island fauna. Lamb et al. (1997) examined the distribution of genetic variation in tassel-eared squirrels (*Sciurus aberti*) on sky islands in the American southwest, revealing distinct eastern and western assemblages, with further genetic subdivision observed in the eastern group. Based on the pattern of differentiation, the authors concluded that climatological events during the Pleistocene may have been important in the divergence of east-west assemblages, while habitat expansion and dispersal were responsible for observed variation within assemblages.

Masta (2000) found that populations of jumping spiders (*Habronattus pugillis*) inhabiting smaller sky islands in southeastern Arizona were monophyletic, while populations on larger islands were not, a result attributed to incomplete lineage sorting. The author concluded that topography (i.e., elevation) played an important role in the differentiation of populations, and that multiple vicariance events were responsible for the observed patterns of genetic variation.

DeChaine and Martin (2004) examined population structure in an alpine butterfly (*Parnassius smintheus*) in the Rocky Mountains. Analysis suggested distinct northern and southern groups separated by a large dispersal barrier, while population structure within these groups was attributed to cyclic habitat fragmentation and expansion caused by climate change during the last 400,000 years.

Barrowclough et al. (2004) found strong geographic structure corresponding to subspecies groups in a study of blue grouse (*Dendragapus obscurus*) in coniferous forests of western North America. Gene flow was extensive, however, among populations within the *D. o. fuliginosus* subspecies group, and within northern and southern clades of the *D. o. obscurus* subspecies group. The authors suggested that differences between subspecies groups were due to separation during the late Pleistocene, while extensive continuous habitat during glacial epochs was responsible for gene flow within groups.

These studies suggest that western North American sky island populations exhibit varying degrees of differentiation consistent with Pleistocene climate change, and that these populations are recently diverged, having become genetically differentiated only within the last

350,000 years during the late Pleistocene (Lamb et al. 1997, Masta 2000, DeChaine and Martin 2004, Barrowclough et al. 2004).

Pygmy Nuthatch as a Model Organism

The pygmy nuthatch (*Sitta pygmaea*) is an ideal species for examining genetic variation among recently diverged sky island taxa. It exhibits well-documented morphological variation among subspecies, has a widespread range showing geographic structure and habitat specificity, and displays extremely low vagility in light of its powers of flight (Bent 1948, Norris 1958). These qualities present the foundation for potentially high levels of genetic differentiation among discontinuous sky island populations of pygmy nuthatches.

Pygmy nuthatches are small songbirds (family Sittidae) that inhabit coniferous forests throughout western North America. They range from the Pacific coast east to the Rocky Mountains, and from British Columbia south into central Mexico (Figure 1). There are currently six recognized subspecies: *S. p. pygmaea* occurring along the Pacific coast in central California; *S. p. melanotis* inhabiting mountains from British Columbia, Canada south to

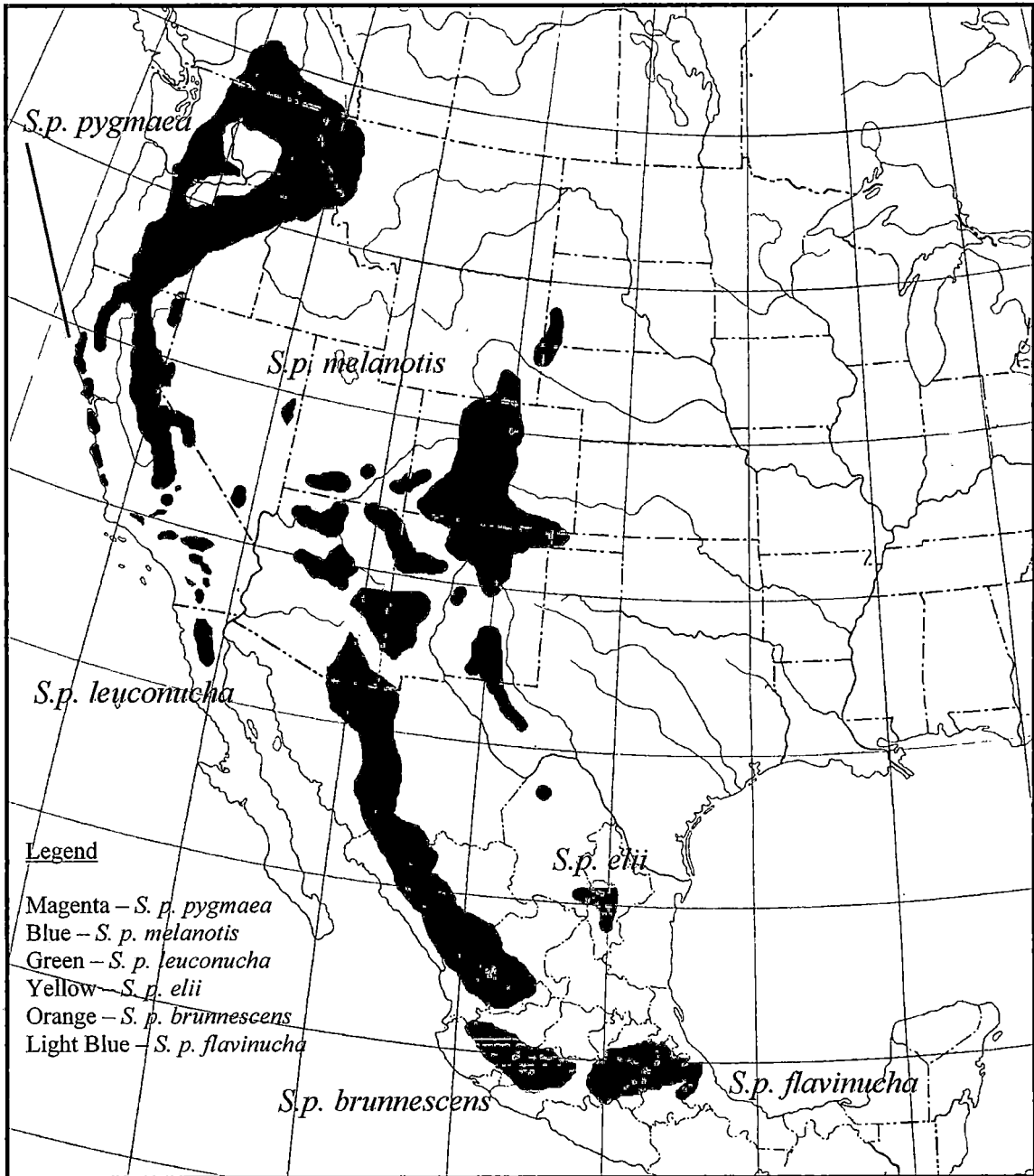


Figure 1. *Sitta pygmaea* Range Map. Approximate range of *Sitta pygmaea* showing distribution of subspecies based on Grinnell and Miller (1944), Norris (1958), National Geographic Society (2002), and Howell and Webb (1995). Map adapted from Little (Little, E.L., Jr. 1971. Atlas of United States trees, volume 1: Conifers and Important Hardwoods. Misc. pub. 1146. U.S. Dept. of Agriculture. Washington, D.C.)

southern California (Pacific interior) and northern Mexico (western North American interior); *S. p. leuconucha* occupying mountains from southern California to Baja California Norte, Mexico; and *S. p. elii*, *S. p. flavinucha*, and *S. p. brunnescens*, which occur in the mountains of central and northern Mexico (Kingery and Ghalambor 2001).

Pygmy nuthatches are restricted to coniferous forests (Small 1994), and as such, their distribution varies from continuous to fragmented throughout their range. In southern California the current patchy distribution of coniferous forests results in large distances between populations of pygmy nuthatches (Figure 2). *S. p. leuconucha* inhabits ponderosa pine forest, generally above 4000 feet elevation, in seven separate populations located in Riverside County in the San Jacinto and Santa Rosa Mountains, and in San Diego County in the Laguna, Cuyamaca, Volcan, and Hot Springs Mountains and on Palomar Mountain. Although most authorities ascribe the Riverside County birds to this subspecies (Grinnell and Miller 1944, Bent 1948, Small 1994), some have described them as belonging to, or intergrading with, *S. p. melanotis* (Norris 1958, Kingery and Ghalambor 2001). *S. p. melanotis* shares similar habitat requirements and ranges from San Bernardino County

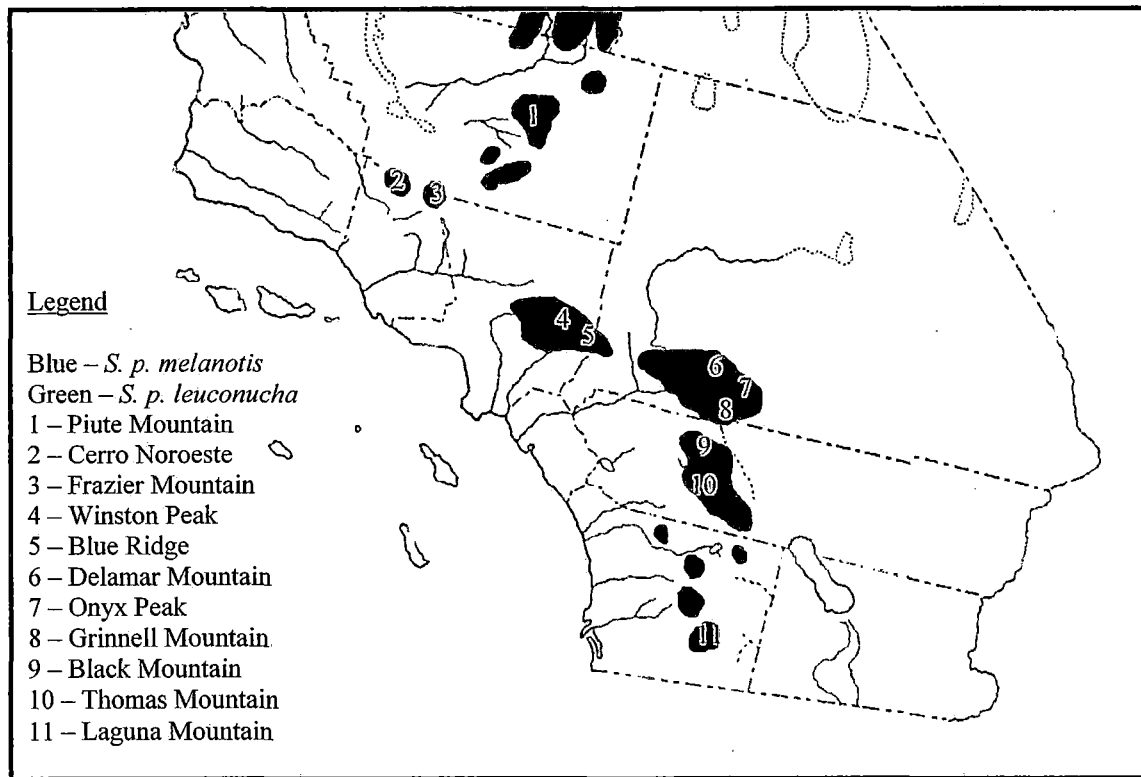


Figure 2. *Sitta pygmaea* Southern California Range Map. Approximate southern California distribution of *Sitta pygmaea* showing subspecies and sampling locations; adapted from Grinnell and Miller (Grinnell, J., and A.H. Miller. 1944. *The Distribution of the Birds of California*. Cooper Ornithological Club. Pacific Coast Avifauna: Number 27. pp. 318-321.)

through northern California. It occurs in disjunct populations in San Bernardino County in the San Bernardino Mountains, Los Angeles County in the San Gabriel Mountains, Ventura County on Frazier Mountain and Mount Pinos, and Kern County on Cerro Noroeste, the Tehachapi Mountains, and the Piute Mountains. The distribution of this subspecies becomes continuous in northern Kern County at the southern

edge of the Sierra Nevada. *S. p. pygmaea* is relatively continuously distributed in San Luis Obispo and Monterey Counties in the Santa Lucia Range and along the coastal fog belt where it prefers long-needled pines (Kingery & Ghalambor 2001).

In addition to being geographically structured, pygmy nuthatches are relatively sedentary. They are a resident species that does not migrate, and shows limited dispersal. Norris (1958) reported mean natal dispersal distances of 286.5 meters, and a maximum recorded dispersal of 533 meters. Therefore, natal dispersal would not promote migration between disjunct mountain ranges. However, pygmy nuthatches are occasionally observed outside their habitat, with a few sightings recorded less than annually in the desert or coastal lowlands (Garrett and Dunn 1981). Thus pygmy nuthatches are capable of rare long distance flights, and could potentially span the distance between disjunct mountain ranges.

Population Genetics and Phylogeography of Pygmy Nuthatches

The primary focus of this study will be to characterize the genetic variation present within and among

populations of two subspecies of pygmy nuthatch occurring in southern California (Figure 2). From these data several other questions relating to the population biology, evolutionary ecology, and biogeography of this species will be addressed. Questions concerning connectivity and gene flow arise about many species with discontinuous distributions. Evolutionary and island models predict that when migration is restricted (and selection is negligible), random drift will act to differentiate populations inhabiting separate islands. This study will examine the degree of genetic differentiation among sky island populations of pygmy nuthatch and assess the hypothesis that sky islands are genetically isolated and that migration (gene flow) is restricted among them.

Up to seven subspecies of pygmy nuthatch have been described based on plumage coloration and skeletal measurements (Norris 1958). Two of these subspecies occur in the study area, with the subspecific designation of one population in dispute. This study will also evaluate whether differences in mitochondrial DNA sequences substantiate these morphological subspecific designations. Analysis of neutral genetic markers could corroborate

morphological subspecies, and shed light on the genetic relationship of disputed populations.

Genetic differentiation can also be used to estimate divergence times of populations and subspecies and infer the historical biogeography of populations. These estimates can be compared with the geological record to determine if paleoclimatological events were important in genetic divergence. Recent studies have shown that Pleistocene climate has played a significant role in both initiating divergence events in North American passerines as well as maintaining previously formed divergence events (Awise and Walker 1998, Johnson and Cicero 2004). The divergence times of pygmy nuthatch populations may be approximated by applying calibrated molecular clocks derived from fossil data from other passerines to observed sequence differentiation (Lovette 2004), potentially allowing the evaluation of the role of Pleistocene climate change on the evolution of pygmy nuthatches. While there is significant variation in molecular evolutionary rates among species, among both genetic markers and species, most rate calibrations of molecular clocks in birds have clustered around 2% per million years (Lovette 2004). A study by Fleischer et al. (1998) on partial *cyt b* sequences

of Hawaiian honeycreepers revealed a substitution rate of 1.6 - 1.9% per million years. For molecular clock estimates to be accurate, however, the populations being compared must be reciprocally monophyletic (Cicero and Johnson 2004).

The results of this project will be an important contribution to the knowledge of the biology of pygmy nuthatches, as well to the understanding of the evolution and ecology of recently diverged taxa, especially sky island species. Little is known about the evolutionary ecology and the population genetics of pygmy nuthatches. This investigation will provide valuable information to fill in the gaps in our knowledge of this species, with particular regard to the evolutionary relationships and history of this species within and among populations and mountain ranges throughout southern California. Significant genetic differentiation, if observed, would indicate that disjunct populations are isolated and gene flow is restricted among populations. Alternatively, lack of genetic differentiation could be attributed to recent or current migration among populations, providing direction for further research into the mechanisms by which a purportedly sedentary species maintains gene flow.

On a broader scale this project will improve the understanding of the population genetics and phylogeography of sky island species. The results of this research will provide insight into the evolution and behavior of sedentary species occurring on sky islands, especially birds.

CHAPTER TWO
MATERIALS AND METHODS

Data Collection

Two of the three subspecies of pygmy nuthatch occurring in southern California, *S. p. melanotis* and *S. p. leuconucha*, are included within the scope of this study. Eleven sampling locations throughout southern California were selected based on representation of subspecies, availability of specimens, and feasibility of collection. Sampling locations were also chosen to create a hierarchical geographic structure, allowing for the analysis of genetic variation within mountain ranges, among mountain ranges, and between subspecies. Seven specimens were sampled from each location, except Thomas Mountain and Piute Mountain where 6 specimens each were sampled. Eight locations represent the widespread *S. p. melanotis* and three locations the southern *S. p. leuconucha*. Three brown-headed nuthatch (*Sitta pusilla*) samples were included for outgroup analysis. Table 1 summarizes subspecies and collection information for all specimens used in this study.

Table 1. *Sitta pygmaea* Specimen Information.

Museum	GenBank					
Accession	Accession	3045-bp	1639-bp			
Number	Number	haplotype	haplotype	Subspecies	Collection Location	Mountain Range
SBCM 57726	DQ153996	A1	A	<i>melanotis</i>	Piute Mountain	Piute
SBCM 57727	DQ194292	A1	A	<i>melanotis</i>	Piute Mountain	Piute
SBCM 57728	DQ153998	B1	B	<i>melanotis</i>	Piute Mountain	Piute
SBCM 57729	DQ153985		C	<i>melanotis</i>	Piute Mountain	Piute
SBCM 57730	DQ194297	B1	B	<i>melanotis</i>	Piute Mountain	Piute
SBCM 57731	DQ194298	B1	B	<i>melanotis</i>	Piute Mountain	Piute
SBCM 57737	DQ155593		D	<i>melanotis</i>	Cerro Noroeste	
SBCM 57738	DQ155594		E	<i>melanotis</i>	Cerro Noroeste	
SBCM 57739	DQ155591		A	<i>melanotis</i>	Cerro Noroeste	
LSU B-41563	DQ194306		A	<i>melanotis</i>	Cerro Noroeste	
LSU B-41564	DQ194307		A	<i>melanotis</i>	Cerro Noroeste	
LSU B-41566	DQ194328		C	<i>melanotis</i>	Cerro Noroeste	
LSU B-41567	DQ194329		C	<i>melanotis</i>	Cerro Noroeste	
SBCM 57734	DQ194308		D	<i>melanotis</i>	Frazier Mountain	
SBCM 57735	DQ194309		D	<i>melanotis</i>	Frazier Mountain	
SBCM 57736	DQ194310		D	<i>melanotis</i>	Frazier Mountain	
LSU B-41557	DQ194330		C	<i>melanotis</i>	Frazier Mountain	
LSU B-41558	DQ194331		C	<i>melanotis</i>	Frazier Mountain	
LSU B-41560	DQ194332		C	<i>melanotis</i>	Frazier Mountain	
LSU B-41561	DQ194314		E	<i>melanotis</i>	Frazier Mountain	
SBCM 57633	DQ154000	E1	E	<i>melanotis</i>	Winston Peak	San Gabriel
SBCM 57634	DQ153986		F	<i>melanotis</i>	Winston Peak	San Gabriel
SBCM 57635	DQ194312		E	<i>melanotis</i>	Winston Peak	San Gabriel
SBCM 57636	DQ194325		C	<i>melanotis</i>	Winston Peak	San Gabriel
SBCM 57637	DQ194311		D	<i>melanotis</i>	Winston Peak	San Gabriel
SBCM 57638	DQ194313		E	<i>melanotis</i>	Winston Peak	San Gabriel
SBCM 57639	DQ194303		A	<i>melanotis</i>	Winston Peak	San Gabriel
SBCM 56877	DQ153987		G	<i>melanotis</i>	Blue Ridge	San Gabriel
SBCM 56878	DQ194334		G	<i>melanotis</i>	Blue Ridge	San Gabriel
SBCM 56898	DQ153997	A2	A	<i>melanotis</i>	Blue Ridge	San Gabriel
SBCM 56899	DQ194293	A1	A	<i>melanotis</i>	Blue Ridge	San Gabriel
SBCM 56900	DQ194294	A1	A	<i>melanotis</i>	Blue Ridge	San Gabriel
SBCM 56901	DQ194295	A1	A	<i>melanotis</i>	Blue Ridge	San Gabriel

Table 1. *Sitta pygmaea* Specimen Information (continued).

Museum	GenBank					
Accession	Accession	3045-bp	1639-bp	Subspecies	Collection Location	Mountain Range
Number	Number	haplotype	haplotype			
SBCM 56902	DQ155595		H	<i>melanotis</i>	Blue Ridge	San Gabriel
SBCM 57596	DQ154001	H1	H	<i>melanotis</i>	Delamar Mountain	San Bernardino
SBCM 57597	DQ194315		H	<i>melanotis</i>	Delamar Mountain	San Bernardino
SBCM 57598	DQ194316		H	<i>melanotis</i>	Delamar Mountain	San Bernardino
SBCM 57599	DQ154004	I1	I	<i>melanotis</i>	Delamar Mountain	San Bernardino
SBCM 57640	DQ194326		C	<i>melanotis</i>	Delamar Mountain	San Bernardino
SBCM 57641	DQ194327		C	<i>melanotis</i>	Delamar Mountain	San Bernardino
SBCM 57642	DQ155592		B	<i>melanotis</i>	Delamar Mountain	San Bernardino
SBCM 57591	DQ154005	J1	J	<i>melanotis</i>	Onyx Peak	San Bernardino
SBCM 57592	DQ153988		K	<i>melanotis</i>	Onyx Peak	San Bernardino
SBCM 57593	DQ154006	L1	L	<i>melanotis</i>	Onyx Peak	San Bernardino
SBCM 57594	DQ155597		L	<i>melanotis</i>	Onyx Peak	San Bernardino
SBCM 57630	DQ194301	J1	J	<i>melanotis</i>	Onyx Peak	San Bernardino
SBCM 57631	DQ153989		M	<i>melanotis</i>	Onyx Peak	San Bernardino
SBCM 57632	DQ194324		L	<i>melanotis</i>	Onyx Peak	San Bernardino
SBCM 56616	DQ155596		J	<i>melanotis</i>	Grinnell Mountain	San Bernardino
SBCM 56617	DQ153990		N	<i>melanotis</i>	Grinnell Mountain	San Bernardino
SBCM 56618	DQ194317		H	<i>melanotis</i>	Grinnell Mountain	San Bernardino
SBCM 56619	DQ153991		O	<i>melanotis</i>	Grinnell Mountain	San Bernardino
SBCM 56678	DQ194336		N	<i>melanotis</i>	Grinnell Mountain	San Bernardino
SBCM 56679	DQ194333		F	<i>melanotis</i>	Grinnell Mountain	San Bernardino
SBCM 56681	DQ194318		H	<i>melanotis</i>	Grinnell Mountain	San Bernardino
SBCM 56813	DQ194335		K	<i>leuconucha</i>	Black Mountain	San Jacinto
SBCM 56814	DQ153999	D1	D	<i>leuconucha</i>	Black Mountain	San Jacinto
SBCM 56816	DQ194296	A1	A	<i>leuconucha</i>	Black Mountain	San Jacinto
SBCM 56817	DQ194304		A	<i>leuconucha</i>	Black Mountain	San Jacinto
SBCM 57765	DQ154002	H2	H	<i>leuconucha</i>	Black Mountain	San Jacinto
SBCM 57766	DQ194319		H	<i>leuconucha</i>	Black Mountain	San Jacinto
SBCM 57767	DQ153992		P	<i>leuconucha</i>	Black Mountain	San Jacinto
SBCM 57583	DQ194320		H	<i>leuconucha</i>	Thomas Mountain	San Jacinto
SBCM 57584	DQ153993		Q	<i>leuconucha</i>	Thomas Mountain	San Jacinto
SBCM 57585	DQ194321		H	<i>leuconucha</i>	Thomas Mountain	San Jacinto
SBCM 57586	DQ194305		A	<i>leuconucha</i>	Thomas Mountain	San Jacinto

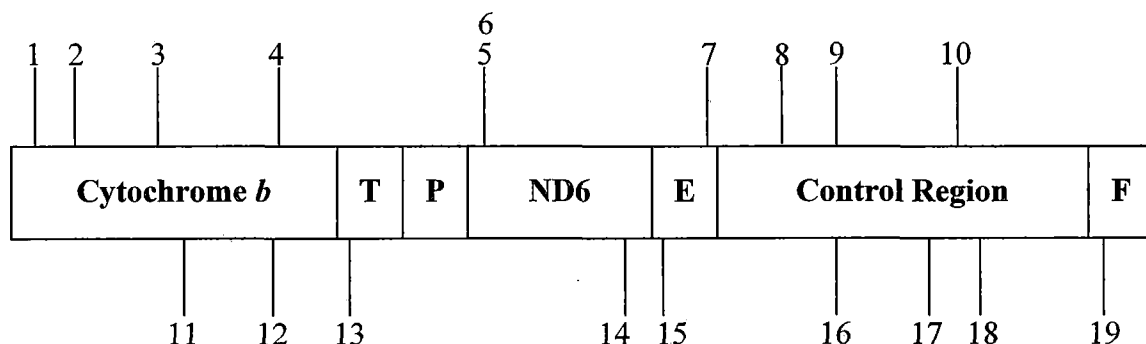
Table 1. *Sitta pygmaea* Specimen Information (continued).

Museum	GenBank					
Accession	Accession	3045-bp	1639-bp			
Number	Number	haplotype	haplotype	Subspecies	Collection Location	Mountain Range
SBCM 57587	DQ194980	D1	D	<i>leuconucha</i>	Thomas Mountain	San Jacinto
SBCM 57588	DQ154003	H3	H	<i>leuconucha</i>	Thomas Mountain	San Jacinto
SBCM 57579	DQ194302	J1	J	<i>leuconucha</i>	Laguna Mountain	Laguna
SBCM 57580	DQ194300	H3	H	<i>leuconucha</i>	Laguna Mountain	Laguna
SBCM 57581	DQ194299	D1	D	<i>leuconucha</i>	Laguna Mountain	Laguna
SBCM 57626	DQ194322		H	<i>leuconucha</i>	Laguna Mountain	Laguna
SBCM 57627	DQ194338		P	<i>leuconucha</i>	Laguna Mountain	Laguna
SBCM 57628	DQ194337		N	<i>leuconucha</i>	Laguna Mountain	Laguna
SBCM 57629	DQ194323		H	<i>leuconucha</i>	Laguna Mountain	Laguna
LSU B-2485	DQ154007	R1	R	<i>S. pusilla</i>	Louisiana	
LSU B-2486	DQ153994		S	<i>S. pusilla</i>	Louisiana	
LSU B-2487	DQ153995		T	<i>S. pusilla</i>	Louisiana	

SBCM-San Bernardino County Museum, LSU-Louisiana State University Museum of Natural Sciences Collection of Genetic Resources; subspecies designations based on Grinnell and Miller (1944).

Tissue samples were acquired from collections at the San Bernardino County Museum and the Louisiana State University Museum of Natural Sciences. Protocols for tissue digestion and DNA extraction follow Hillis et al. (1996) and are summarized as follows. For each specimen, 50 milligrams of frozen liver tissue was macerated on a watch glass and transferred to a 1.5-milliliter centrifuge tube and digested with proteinase-K in a 55°C water bath for a minimum of 8 hours. Total genomic DNA was extracted from the resulting tissue homogenate using 2 phenol-chloroform-isoamyl (25:24:1) and 1 chloroform-isoamyl (24:1) extractions, followed by precipitation with sodium acetate and ethanol. The DNA precipitate was centrifuged and the excess solution decanted. The DNA pellet was air dried in a hood for one hour and then dissolved in distilled water. Concentrations of total genomic DNA were determined by absorbance at 280 nm using a spectrophotometer.

An approximately 3100-base pair fragment of mtDNA was initially amplified and sequenced in 23 samples. This fragment includes the control region, ND6, tRNA^{Glu}, tRNA^{Pro}, tRNA^{Thr}, and a portion of cytochrome *b* (Figure 3). Primers



Name	5'-3' Sequence	Source
1 L14990	CCATCCAACATCTCAGCATGATGAAA	Gill 1999
2 LCYTBc*	AGAGACCTGAAACGTAGGAG	This study
3 LCYTB1	GAGCCTGAGGTGGATTCTCTG	This study
4 LCYTB2	CCGCCTCCGTCCTAGTCTTAT	This study
5 LND6A1	GCCCGAATYGCCCCMCGAGA	This study
6 LND6*	GCCCGAATAGCCCCCGTGA	This study
7 LGLU	ATTCCCACCTGGCCTCTCTCC	This study
8 LCR2*	CATTAACCCAAGTGATCCTAC	This study
9 L437	CTCACGAGAACCGAGCTACT	Tarr 1995
10 LCR1	TAAGTAGACCAACTTCGGCACTGA	This study
11 H15487	GATCCTGTTTCGTGGAGGAAGGT	Pasquet 1998
12 HCYTB*	CAGAAGAGGATTTGTGAGAG	This study
13 H16065	GGAGTCTTCAGTCTCTGGTTTACAAGAC	Gill 1999
14 HND6A	AAYCCYTCTCCYTAYTATGGGGT	This study
15 HGLU*	GGAGAGAGGCCAGGTGGGA	This study
16 HRC437B	ACGTCGAGTAGCTCGTTTCT	This study
17 HCR3*	CAGACGTATGTTCCGCCGTC	This study
18 HCR2	AAAGGTAAATAGCGCGGGTTTCAT	This study
19 H1248	CATCTTCAGTGTCATGCT	Tarr 1995
M13 forward	CACGACGTTGTAAAACGAC	
M13 reverse	GGATAACAATTTACACAGG	

Figure 3. Primer Map.

Approximate locations of primers used in PCR and sequencing. T, P, E, and F are the tRNAs for threonine, proline, glutamic acid, and phenylalanine, respectively. Figure not to scale. *Primers used for amplification of the 1639-base pair fragment.

used for PCR amplification of this fragment were L14990 (Gill 1999) and H1248 (Tarr 1995). Internal sequencing primers were designed from alignments of the mitochondrial genomes of *Corvus frugilegus*, *Vidua chalybeata*, and *Smithornis sharpei* downloaded from GenBank (accession numbers Y18522, AF090341, NC000879, respectively) and pygmy nuthatch sequence data. Figure 3 lists the PCR and sequencing primers used in this study. PCR reactions were carried out using Eppendorf MasterTaq Kit (Brinkmann Instruments, Inc.) in 50 μ L volumes: 5.0 μ L template DNA (approximately 0.2 μ g/mL), 1.0 μ L each primer (20 mM), 0.5 μ L polymerase, 5.0 μ L dNTPs, 5.0 μ L buffer, and 10.0 μ L TaqMaster. The profile for PCR reactions was: initial denaturation at 94°C for 2 minutes; 35 cycles of denaturation at 94°C for 50 seconds, annealing at 50°C for 1 minute, and extension at 72°C for 2 minutes; and final extension at 72°C for 2 minutes. PCR products were visualized with gel electrophoresis using a 0.8% agarose gel and a 1-kilobase ladder, and purified using Montage PCR Centrifugal Filter Devices (Millipore Corporation). Purified PCR products were sequenced by Laragen, Inc. (Los Angeles, CA). Double-stranded sequences were constructed

using Sequitherm EXCEL II DNA Sequencing Kit-LC (Epicentre) in 20 μ L volumes: 1.0 μ L purified PCR product, 1.0 μ L each M13 forward and reverse primers, 7.2 μ L buffer, 1.0 μ L polymerase, and 8.8 μ L distilled water. The profile for cycle sequencing reactions was: initial denaturation at 92°C for 2 minutes; 20 cycles of denaturation at 92°C for 30 seconds, annealing at 54°C for 30 seconds, and extension at 70°C for 1 minute. Cycle sequencing products were sequenced on a LI-COR 4300 automated sequencer. Double-stranded sequences were constructed with AlignIR version 2.0 software (LI-COR, Inc.) and aligned in ClustalX (Thompson et al. 1997).

Sequence Analysis

The goal of this study is to characterize the genetic variation and structure of southern California populations of the pygmy nuthatch, and to determine the evolutionary relationships among these populations using mitochondrial DNA sequence data. This study specifically analyzes the potential effect of isolation on the population genetic structure of pygmy nuthatches in southern California, whether the observed genetic structure corresponds to population geographic structure, and the role of historical

biogeographic processes in shaping population genetic structure. To achieve this goal this research examines genetic structure both within and among mountain ranges. This study also examines the correlation of the genetic data with putative morphological subspecies breaks.

Prior to analysis the reverse complement of the ND6 sequence was taken in both the 3045- and 1639-bp fragments using MacClade 4.01 (Maddison and Maddison 2000). This was done to maintain consistent codon positioning with *cyt b*. A partition homogeneity test was conducted in Paup* 4.10b (Swofford 2003) to determine if the three mitochondrial gene fragments could be combined into a single 1639-base pair sequence for analysis (Farris et al. 1995). Genetic variation is described using basic descriptive statistics including nucleotide frequency, number of haplotypes, number of polymorphic sites, number and type of substitutions, number of transitions (Ti) and transversions (Tv), Ti/Tv ratio, and uncorrected genetic distance among haplotypes. Descriptive statistics were calculated using Paup* 4.10b (Swofford 2003), Mega 2.1 (Kumar et al. 2001) and Modeltest 3.06 (Posada and Crandall 1998), or were done by hand.

Population Genetics

Population genetic analyses use F-statistics or other genetic distance measures to evaluate genetic differentiation within and among populations. Wright's F-statistics are fixation indices that were developed to characterize genetic differentiation within and among subpopulations (Hartl and Clark 1997). The most commonly used of these, F_{ST} , measures genetic differentiation among subpopulations that constitute the total population. F-values range from 0 to 1, with increasing values indicating greater degrees of genetic differentiation. Analogs of F-statistics, such as N_{ST} (Lynch and Crease 1990), K_{ST} (Hudson et al. 1992), and Φ_{ST} (Excoffier et al. 1992) have been developed to analyze DNA sequence data.

Analysis of molecular variance (AMOVA) and Mantel tests use molecular-based F-statistics (or their analogs) to test for significant population structure. AMOVA employs a nested analysis of variance approach to compare levels of genetic differentiation in hierarchically structured populations (Excoffier et al. 1992). Arlequin 2.0 (Schneider et al. 2000) was used to implement the AMOVA. Individuals were grouped into populations by mountain range. Population comparisons were based on a

Kimura 2-parameter distance matrix (calculated by Arlequin) and significance testing of F values and variance components was based on 16000 permutations and evaluated at $p=0.05$.

A Mantel test was used to examine the correlation between population pairwise geographic distances (in km) and population pairwise F_{ST} values (calculated by Arlequin). Significant results from a Mantel test indicate isolation by distance among populations. Significance was determined by running 1000 permutations and evaluating at $p=0.05$.

Fu's F_s , nucleotide diversity (Π), and percentage of variable sites were calculated as measures of intrapopulation variability in Arlequin 2.0 (Schneider et al. 2000). Fu's F_s can be used to detect population expansion, and the relative values of nucleotide diversity (Π) among populations may indicate the direction of that expansion if present (Sgariglia and Burns 2003).

Phylogeography

Evolutionary relationships among haplotypes were estimated using three methods of phylogenetic inference: genetic distance, parsimony, and maximum likelihood. Distance methods use genetic distance data to generate gene

trees. Character data are transformed into pairwise distance values (distances between pairs of haplotypes) and plotted in a matrix. Trees can be constructed from the matrix using various approaches, with two of the more common methods being neighbor-joining and minimum evolution. Neighbor-joining connects nearest neighbors within the matrix until all haplotypes are represented in the tree, while minimum evolution searches multiple trees to produce the tree(s) with the shortest overall distance. Neighbor-joining analysis was conducted in Paup* 4.0b10 (Swofford 2003). Modeltest 3.06 (Posada and Crandall 1998) selected the HKY85 (Hasegawa et al. 1985) model of evolution as the best model for this data set. Trees were constructed using an HKY85 model of evolution, branch-swapping via tree-bisection-reconnection, and by conducting 1000 bootstrap replicates.

Parsimony constructs gene trees based on the fewest number of evolutionary steps required to explain differences among haplotypes. Optimality criteria are defined for the transformation of characters, and the most parsimonious tree (or trees) is selected from all possible trees based on the defined criteria. Maximum parsimony is the most commonly used method of phylogenetic inference

(Swofford et al. 1996). Parsimony analysis was conducted in Paup* 4.0b10 (Swofford 2003). Parsimony trees were constructed using both branch-and-bound (the method guaranteed to find the minimum-length tree) and heuristic searches. Bootstrap analysis was performed using 1000 bootstrap replicates and a full heuristic search, with the initial trees obtained via stepwise addition and branch-swapping via tree-bisection-reconnection.

Maximum likelihood selects the tree that best fits the character data based on evolutionary models chosen by the investigator. The maximum likelihood approach is attractive because it often the estimation method least affected by sampling error, tends to conform to the assumptions of its models, and frequently outperforms other methods when evaluated under multiple models of sequence evolution (Swofford et al. 1996). Maximum likelihood analyses were implemented using MrBayes 3.0b4 (Huelsenbeck and Ronquist 2001). Four chains were run for 30 million generations sampling every 1000th generation with a burnin of 3000 sampled generations. The default settings from MrBayes 3.0b4 were used, except that lset nst = 2 and prset ratepr = variable.

Patterns of historical gene flow were inferred using nested clade analysis (NCA). NCA evaluates the null hypothesis of no geographic association among haplotypes by testing the significance of three values: clade distance (D_c), a measure of the geographic range of a clade, nested clade distance (D_n), a measure of the geographic distribution of a clade relative to its sister clades, and the difference of these values between interior and tip clades (I-T; Templeton 1998). The relationship among significant values within a clade can then be used to infer restricted gene flow, population fragmentation, or population expansion for haplotypes within that clade. A haplotype network was created with TCS 1.18 (Clement et al. 2000) and nested according to the rules established by Templeton et al. (1987) and Templeton and Sing (1993). NCA was implemented in GEODIS 2.2 (Posada et al. 2000) with 10000 random permutations. The inference key of Templeton (1998) was used to infer patterns of historical gene flow from the analysis.

A substitution rate of 1.6-1.9% per million years (Fleischer et al. 1998) was used to estimate the divergence times of haplotypes between pygmy and brown-headed nuthatches and among pygmy nuthatch populations. Because

pygmy nuthatch populations are not reciprocally monophyletic these estimates do not correspond to the divergence times of geographic populations. These estimates do, however, provide an approximation of the maximum divergence time among haplotypes as well as a threshold value for divergence of populations.

Subspecies Limits

The mitochondrial DNA sequence data was used to examine the correlation of genetic structure with morphological subspecies described by both Grinnell and Miller (1944) and Norris (1958). Grinnell and Miller (1944) supported a subspecific break between *S. p. melanotis* and *S. p. leuconucha* at the Banning Pass between the San Bernardino Mountains and the San Jacinto Mountains based on wing length, bill size, and plumage coloration (Figure 4). Norris (1958) supported a break between these two subspecies immediately south of the San Jacinto and Santa Rosa Mountains based primarily on the ratio of bill width to bill depth, but also on wing, tail, bill, and tarsus lengths. Individuals were partitioned by subspecies and an AMOVA was conducted across the subspecific break supported by each author using Arlequin 2.0 (Schneider et

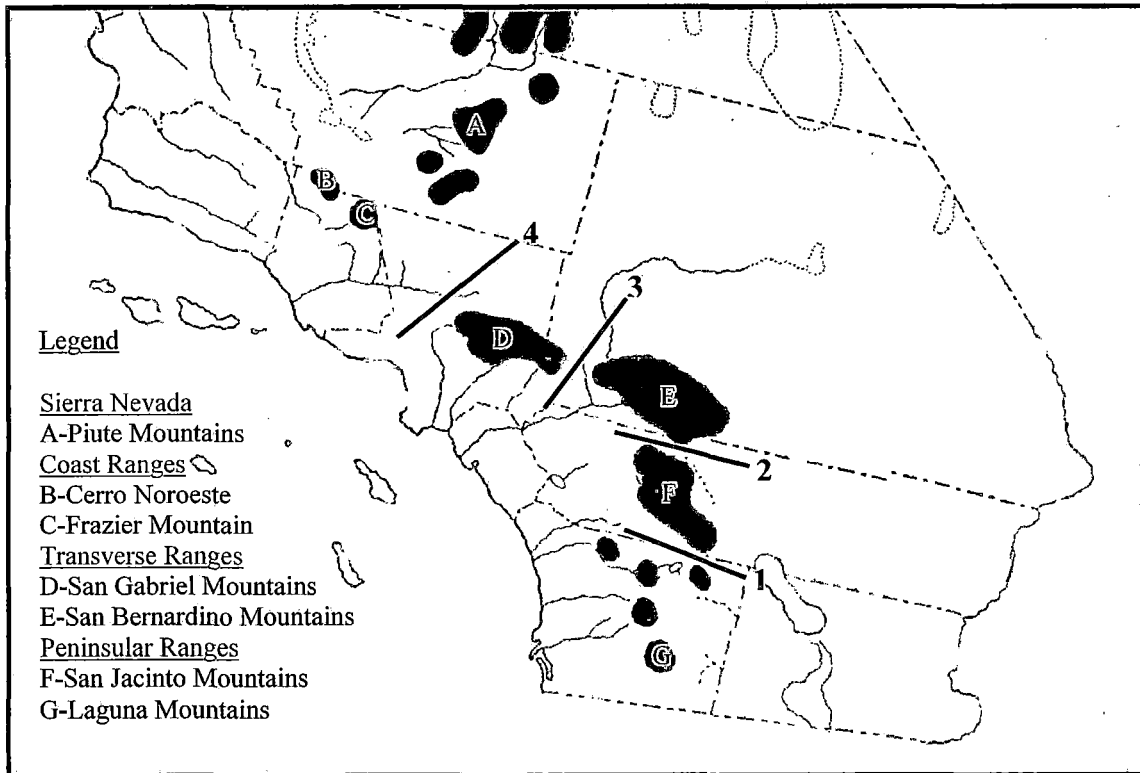


Figure 4. *Sitta pygmaea* Subspecific Break Map. Southern California distribution and hypothesized subspecific breaks between *S. p. melanotis* and *S. p. leuconucha* and their source: 1-immediately south of the San Jacinto and Santa Rosa Mountains (Norris 1958), 2-Banning Pass (Grinnell and Miller 1944), 3-Cajon Pass (this study), 4-immediately north of the San Gabriel Mountains (this study). Map adapted from Grinnell and Miller (Grinnell, J., and A.H. Miller. 1944. The Distribution of the Birds of California. Cooper Ornithological Club. Pacific Coast Avifauna: Number 27. pp. 318-321.).

al. 2000). Observation of the distribution of haplotypes suggested that a stronger genetic subspecific break may occur further north than proposed by Grinnell and Miller (1944) or Norris (1958). Therefore, an AMOVA was also

conducted for each of two additional subspecific breaks: at the Cajon Pass between the San Bernardino and San Gabriel Mountains, and immediately north of the San Gabriel Mountains.

CHAPTER THREE

RESULTS

Sequence Analysis

Analysis of the 3045-base pair genetic marker from 23 pygmy nuthatch specimens in five populations yielded 11 haplotypes (Table 2). Base frequencies for the total sequence were A=0.304, T=0.252, C=0.318, and G=0.126. There were 22 polymorphic sites: 8 in cytochrome *b*, 4 in ND6, and 10 in the control region. All coding region substitutions were synonymous and, with the exception of one 1st-position substitution in *cyt b*, occurred at 3rd positions. The average uncorrected pairwise distance among all haplotypes was approximately 0.0020, substantially lower than values reported for other bird species. The overall transition-transversion ratio was 22.0, with 8, 4, and 9 transitions occurring in *cyt b*, ND6, and the control region respectively, and 1 transversion occurring in the control region. Base frequency and substitution data for the total sequence and for individual gene regions are summarized in Table 3.

Table 2. Variable Sites within 3045-Base Pair *Sitta pygmaea* Haplotypes.

Haplotype	n	Cyt b								ND6				Control Region															
		2	5	7	8	8	8	9	0	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	
A1	6	A	T	A	T	G	C	T	T	C	T	C	C	A	A	A	T	T	G	A	C	C	C	C	C	C	C	C	
A2	1	T
B1	3	A	A	.	.	T	
D1	3	G	.	.	.	A	.	.	.	T	C	.	T	.	.	G	
E1	1	.	C	G	.	A	A	
H1	1	G	.	.	.	A	.	.	.	T	C	.	T	.	.	G	.	.	A	.	T	
H2	1	G	.	.	.	A	.	C	.	T	C	.	T	.	.	G	.	.	A	G	
H3	2	G	.	.	.	A	.	.	.	T	C	.	T	.	.	G	.	.	A	
I1	1	G	.	.	C	A	T	.	C	.	C	T	C	.	A	
J1	3	G	.	.	.	A	.	.	.	T	C	.	T	.	C	G	.	.	A	
L1	1	A	G	.	.	.	C	A	

S. pygmaea haplotype frequency and variable site data for 3045-bp sequences. Numbers above columns indicate nucleotide position relative to 3045-bp sequence; dots indicate identity with haplotype A1.

Table 3. Gene Statistics for 3045-Base Pair Haplotypes.

		Overall (3045 bp)	Cyt <i>b</i> (partial; 1045 bp)	ND6 (519 bp)	Control Region (1248 bp)
Base frequencies	A	0.304	0.273	0.397	0.286
	T	0.252	0.251	0.122	0.304
	C	0.318	0.340	0.394	0.276
	G	0.126	0.136	0.087	0.134
Substitutions	1 st pos	1	1		
	2 nd pos				
	3 rd pos	11	7	4	
	non- coding	10			10
	Total	22	8	4	10
	Ti	21	8	4	9
	Tv	1			1
Ti/Tv	22.0			10.0	

Base frequencies are presented for the light strand of the overall sequence and the individual gene regions. Substitution data are provided for the light strand of *cyt b* and the control region, and the heavy strand of ND6.

The partition homogeneity test found no significant differences among the three gene fragments, therefore they were combined into a single sequence for analysis. Analysis of the 1639-base pair concatenated sequence from 75 pygmy nuthatch specimens in seven populations yielded 17 haplotypes (Table 4), and results were similar overall to the 3045-base pair sequence. Base frequencies for the concatenated sequence were A=0.306, T=0.228, C=0.337, and G=0.129. There were 26 polymorphic sites: 12 in cytochrome *b* and 7 each in ND6 and the control region. In *cyt b* seven 3rd and two 1st-position coding region

Table 4. Variable Sites within 1639-Base Pair *Sitta pygmaea* Haplotypes.

Haplotype	n	Cyt <i>b</i>											ND6							Control Region							
		4	8	1	1	1	3	4	5	5	5	5	5	7	7	8	8	8	9	0	1	1	1	1	1	1	1
A	13	A	G	T	C	T	T	T	A	C	G	T	G	G	G	C	A	A	G	T	A	A	A	G	T	T	G
B	4	A	A	
C	9	A	G	A	
D	8	G	A	A	.	.	G	.	A	.	.	.	G	
E	5	C	.	G	.	.	A	A	
F	2	G	.	.	.	C	C	A	.	A	.	G	C	.	A	
G	2	A	G	.	.	.	G	A	
H	14	G	A	A	.	.	G	.	A	.	.	.	G	.	.	.	A	
I	1	G	C	A	.	A	.	G	C	.	A	
J	4	G	A	A	.	.	G	.	A	.	C	.	G	.	.	.	A	
K	2	G	C	A	A	.	.	G	.	A	.	.	.	G	.	.	.	A	
L	3	A	C	A	
M	1	A	A	.	.	A	
N	3	G	A	T	.	A	.	.	T	G	.	.	C	A	
O	1	G	.	.	T	C	A	.	A	.	G	C	.	A	
P	2	A	A	G	A	
Q	1	G	.	C	A	A	.	.	G	.	A	.	.	.	G	.	.	.	A	

S. pygmaea haplotype frequency and variable site data for 1639-bp sequences. Numbers above columns indicate nucleotide position relative to 1639-bp sequence; dots indicate identity with haplotype A.

substitutions were synonymous, while one 2nd and two 1st-
 position substitutions were non-synonymous. All ND6
 substitutions occurred at the 3rd position and were
 synonymous. The overall transition-transversion ratio was
 26.02, with 12, 7, and 6 transitions occurring in *cyt b*,
 ND6, and the control region respectively, and 1
 transversion occurring in the control region. Base
 frequency and substitution data for the concatenated
 sequence and for individual gene regions are summarized in
 Table 5. The average uncorrected pairwise (p) distance
 among all haplotypes was approximately 0.0035, with a

Table 5. Gene Statistics for 1639-Base Pair Haplotypes.

		Overall (1639 bp)	Cyt <i>b</i> (621 bp)	ND6 (504 bp)	Control Region (514 bp)
Base frequencies	A	0.306	0.276	0.401	0.248
	T	0.228	0.249	0.125	0.301
	C	0.337	0.343	0.393	0.277
	G	0.129	0.131	0.081	0.174
Substitutions	1 st pos	4	4		
	2 nd pos	1	1		
	3 rd pos	14	7	7	
	non- coding	7			7
	Total	26	12	7	7
	Ti	25	12	7	6
	Tv	1			1
Ti/Tv	26.0			7.0	

Base frequencies are presented for the light strand of the
 overall sequence and the individual gene regions.
 Substitution data are provided for the light strand of *cyt b*
 and the control region, and the heavy strand of ND6.

maximum p-distance of 0.0049. The average p-distance between *S. pygmaea* and *S. pusilla* haplotypes was 0.0702.

Population Genetics

Population genetic analyses were used to infer the levels of genetic differentiation within and among pygmy nuthatch populations in southern California. The number of haplotypes, percentage of variable sites, nucleotide diversity (Π), and Fu's F_s were calculated as measures of population variability and are summarized in Table 6. No values for Fu's F_s were significant. Results from the overall AMOVA analysis indicated significant, although relatively low levels of genetic differentiation, reporting $\Phi_{ST}=0.139$ ($p<0.05$) with 13.9% of the total variation

Table 6. Estimates of Intrapopulation Variability.

Mountain range	n	Number of haplotypes	% variable sites	Nucleotide diversity (Π)	Fu's F_s
Piute Mountain	6	3	0.18%	0.0009	0.38
Cerro Noroeste	7	4	0.61%	0.0022	1.04
Frazier Mountain	7	3	0.55%	0.0028	3.27
San Gabriel Mountains	14	7	0.92%	0.0027	0.20
San Bernardino Mountains	21	11	1.16%	0.0031	-1.48
San Jacinto Mountains	13	6	0.67%	0.0023	0.54
Laguna Mountain	7	5	0.79%	0.0026	-0.004

Statistics are based on the 1639-base pair haplotype data set. No values of Fu's F_s were significant.

occurring among mountain ranges (Table 7). Differences within populations accounted for 84.1% of the variance. A Mantel test found a significant correlation (0.462, $p=0.007$) between population pairwise F_{ST} values and geographic distances (Appendix B), suggesting isolation by distance among pygmy nuthatch collecting locations.

Table 7. Φ_{ST} Values for Analysis of Molecular Variance Comparisons.

Comparison	Φ_{ST} value	p-value
Among all mountain ranges	0.139	0.0005
Between Laguna Mountains and all individuals to the north	0.102	0.0517
Between individuals north and south of Banning Pass	0.140	0.0010
Between individuals north and south of Cajon Pass	0.205	0.0001
Between individuals north and south of San Gabriel Mountains	0.093	0.0104

Φ_{ST} values and their significance for overall AMOVA and AMOVAs testing subspecies breaks at various locations. Significant Φ_{ST} s are shown in bold.

Phylogeography

Phylogenetic relationships were evaluated using parsimony, distance, and maximum likelihood-based methods. Consensus trees generated by all three methods were congruent (Figures 5-7), and although each showed well-

supported clades, these clades showed no strong geographic structure. Parsimony analysis using a branch and bound search produced three equally likely minimum length trees differing only in the placement of haplotype N (distributed in the San Bernardino and Laguna Mountains; Table 8). The bootstrap consensus tree (Figure 5), which used a full heuristic search and was congruent with the branch-and-bound analysis, provided strong support for two separate clades containing haplotypes F, I and O (87%; distributed in the San Gabriel and San Bernardino Mountains), and D, H, J, K, and Q (91%; distributed throughout all populations). The consensus tree also weakly supported a clade containing haplotypes C, G, and P (55%; distributed throughout all populations except Piute Mountain). These three clades and the remaining 6 haplotypes formed a basal polytomy with respect to the outgroup.

The bootstrap consensus of the neighbor-joining analysis produced a similar tree to the bootstrapped consensus parsimony tree, supporting clades F-I-O and D-H-J-K-Q at 90% each, and clade C-G-P at 70% (Figure 6). The neighbor-joining analysis also provided weak to moderate

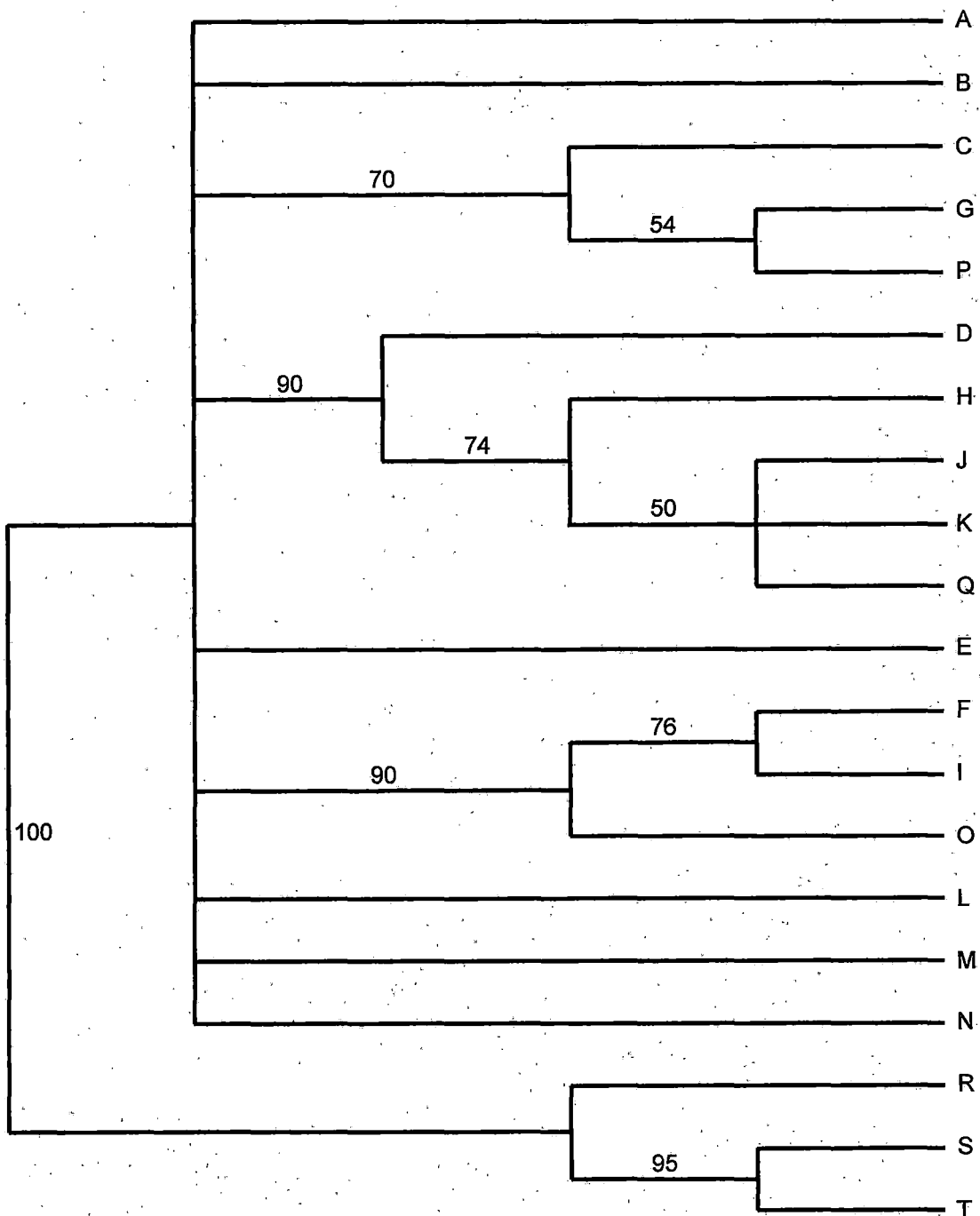


Figure 6. Bootstrapped Consensus Neighbor-Joining Tree. Like the parsimony analysis, the consensus neighbor-joining tree shows strong bootstrap support for clades C-G-P, D-H-J-K-Q, and F-I-O. There is no strong geographic structure among these clades (Table 8).

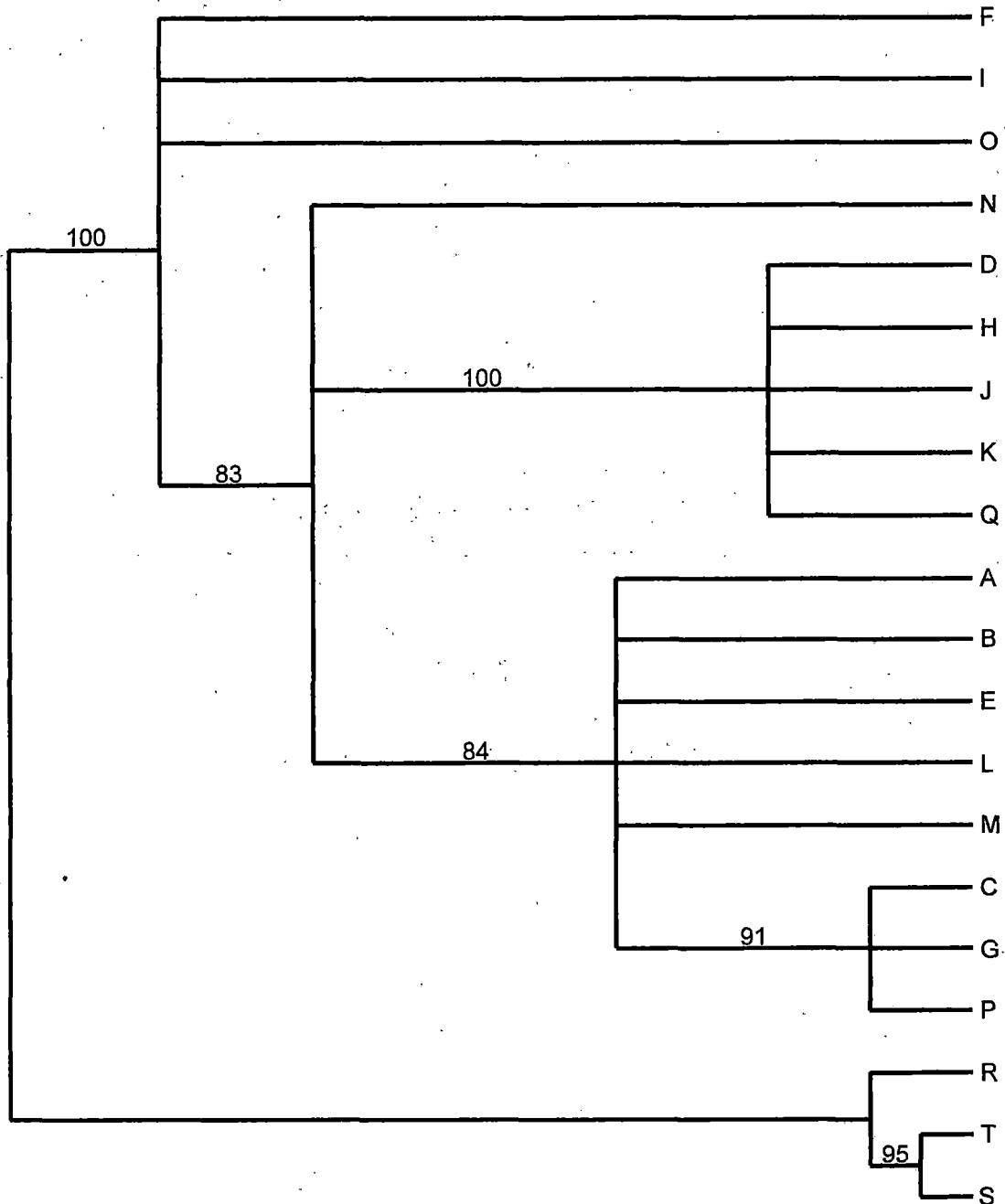


Figure 7. Consensus Bayesian Tree.

The Bayesian tree shows strong support for clades D-H-J-K-Q, C-G-P, A-B-E-L-M-CGP, and N-DHJKQ-ABELMCGP. The Bayesian analysis supports more structure than the parsimony and neighbor-joining analyses, however, lacks the support for clade F-I-O shown by these two analyses. Again there is no strong geographic structure among these clades (Table 8).

Table 8. Distribution of *Sitta pygmaea* Haplotypes and Phylogenetic Clades.

Subspecies	Mountain Range	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
<i>S. p. melanotis</i>	Piute	2	3	1														
	Cerro Noroeste	3		2	1	1												
	Frazier			3	3	1												
	San Gabriel	5		1	1	3	1	2	1									
	San Bernardino		1	2			1		5	1	3	1	3	1	2	1		
<i>S. p. leuconucha</i>	San Jacinto	3			2				5			1					1	1
	Laguna				1				3		1				1		1	
	<i>n</i>	13	4	9	8	5	2	2	14	1	4	2	3	1	3	1	2	1
Represented subspecies		ml	m	m	ml	m	m	m	ml	m	ml	ml	m	m	ml	m	l	l

Subspecies	Mountain Range	A-B-E-L-M	C-G-P	D-H-J-K-Q	F-I-O	N
<i>S. p. melanotis</i>	Piute	5	1			
	Cerro Noroeste	4	2	1		
	Frazier	1	3	3		
	San Gabriel	8	3	2	1	
	San Bernardino	5	2	9	3	2
<i>S. p. leuconucha</i>	San Jacinto	3	1	9		
	Laguna		1	5		1
	<i>n</i>	26	13	29	4	3

Upper portion of table shows the distribution of *Sitta pygmaea* haplotypes with number of individuals of each haplotype occurring in each population. Represented subspecies indicates the subspecies represented by each haplotype: m=*melanotis*, l=*leuconucha*, and ml=*melanotis* and *leuconucha*. Example: haplotype A contains 2 individuals sampled from Piute Mountain and is represented by both *S. p. melanotis* and *S. p. leuconucha*. Lower portion of table shows distribution of *Sitta pygmaea* phylogenetic clades with number of individuals of each clade occurring in each population. Example: 5 individuals within clade A-B-E-L-M were sampled from Piute Mountain. Subspecies designations are based on Grinnell and Miller (1944).

support for groupings within these three clades.

Haplotypes F and I (distributed in the San Gabriel and San Bernardino Mountains) were grouped together with 76% bootstrap support, and H, J, K and Q (distributed in the San Gabriel, San Bernardino, San Jacinto and Laguna Mountains) were grouped with 74% bootstrap support. As in the parsimony consensus tree, the three main clades were arranged in a basal polytomy with the remaining 6 haplotypes with respect to the outgroup.

Bayesian analysis produced a consensus tree similar overall to parsimony and distance-based methods, but with more resolution of the interior nodes (Figure 7).

Haplotypes F, I, and O and a clade containing the remaining 14 haplotypes (83% posterior probability) formed a basal polytomy with respect to the outgroup. Within the clade containing the 14 haplotypes, haplotype N stood alone, the D-H-J-K-Q clade was supported at 100%, and a second clade including haplotypes A, B, E, L, M, C, G, and P (distributed throughout all populations) was supported at 84%, with haplotypes C, G, and P comprising an additional group (91% posterior probability) within this clade.

Phylogeographic relationships were also analyzed using nested clade analysis. TCS produced a single haplotype

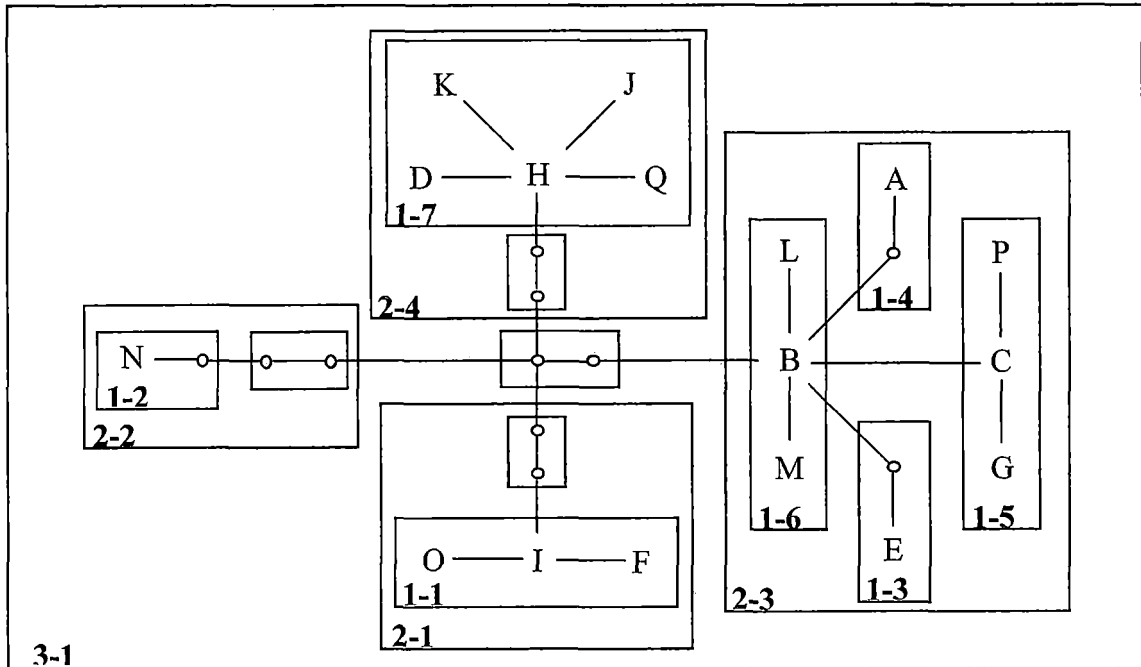


Figure 9. Nested Cladogram.

Nested cladogram showing the successive grouping of sister clades (haplotypes). Nested clade analysis uses these clades in conjunction with geographic information to infer evolutionary history of the clades.

significantly small, and the D_n value for I-T was significantly small, leading to an inference of contiguous range expansion (inference key 1-2-11-12no). Within Clade 2-3, D_n values for Clade 1-6 and I-T were significantly large, leading to an inference of restricted gene flow with isolation by distance (inference key 1-2-11-17-4no). Clade 3-1 also suggested restricted gene flow with isolation by distance (inference key 1-2-3-4no) based on a significantly small D_c value for Clade 2-4.

Based on a substitution rate of 1.6–1.9% (Fleischer et al. 1998), and a mean uncorrected pairwise genetic distance of 7.02% between *S. pygmaea* and *S. pusilla* haplotypes, pygmy nuthatches diverged from brown-headed nuthatches approximately 4.4–3.7 million years ago. The maximum uncorrected pairwise genetic distance among pygmy nuthatch haplotypes (several pairings) is 0.49%. Using the same substitution rate these haplotypes diverged approximately 250,000 to 300,000 years ago.

Subspecies Limits

The results of the AMOVA for the hypothesized subspecific breaks are summarized in Table 7. The AMOVA analysis provided no significant genetic support for Norris' hypothesis of a subspecific break immediately south of the Santa Rosa Mountains (Figure 4). The AMOVA analysis did provide significant genetic support for Grinnell and Miller's hypothesis of a subspecific break at the Banning Pass. The $\Phi_{ST}=0.140$ indicates that 14% of the genetic differentiation among pygmy nuthatches can be accounted by differences between individuals north and south of the Banning Pass. This is similar to the value among mountain ranges irrespective of subspecies. I tested two additional

hypotheses based on my observation of the haplotype distribution data (Table 8). Both AMOVA analyses yielded significant results. An $\Phi_{ST}=0.205$ was reported for the AMOVA across the Cajon Pass, indicating that over 20% of the genetic differentiation among pygmy nuthatches can be accounted by differences between individuals north and south of Cajon Pass. This was the highest Φ_{ST} reported among the four AMOVAs conducted at hypothesized subspecific breaks. A significant $\Phi_{ST}=0.093$ was reported for the subspecific break immediately north of the San Gabriel Mountains.

CHAPTER FOUR

DISCUSSION

Overall the analyses reveal that pygmy nuthatch populations in southern California are currently stable with restricted gene flow among populations. There is significant genetic differentiation of haplotypes among populations and a significant association between linear distance and population subdivision. There is, however, no strong geographic structure of haplotypes among populations. Considered together, this suggests that: 1) historical gene flow was extensive, 2) populations were isolated relatively recently, and 3) the current distribution of pygmy nuthatches is the result of vicariance rather than dispersal. Analyses of the subspecific break between *S. p. melanotis* and *S. p. leuconucha* indicate that the strongest genetic difference occurs at the Cajon Pass, which is further north than the hypothesized subspecific breaks supported by morphological characters.

Sequence Analysis

Analysis of both the 3045- and 1639-base pair sequences revealed low levels of genetic variation within southern California populations of the pygmy nuthatch. Overall genetic diversity was low, as was the degree of genetic differentiation among populations and subspecies (Table 7). Nucleotide composition and substitution data of both the 3045 and 1639-base pair sequences (Tables 3 and 5) are generally consistent with values reported for other birds at the intraspecific level (Baker and Marshall 1997, Moore and DeFilippis 1997). Specifically the low G content is characteristic of D-loop in birds (Baker and Marshall 1997), and the high proportion of 3rd position, synonymous substitutions within coding regions is consistent with recently diverged birds (Moore and DeFilippis 1997). The average pairwise genetic distance among haplotypes (0.0035) is also similar to values found in other passerines (Buerkle 1999, Sgariglia and Burns 2003, Burns and Barhoum 2006).

Population Genetics

Results of the population genetic analyses revealed a low but significant level of genetic differentiation among

populations (Table 7). Specifically, the AMOVA indicated that 13.9% of the genetic variance is due to differences among rather than within mountain ranges. There is also a significant isolation by distance effect among collecting locations. These results suggest that there has been extensive historical gene flow, but that this gene flow has been restricted in the recent past. The lack of lineage sorting from the phylogenetic analyses also suggest that gene flow was only recently restricted.

These results are consistent with the historical distribution of the pygmy nuthatch's coniferous forest habitat. Studies in the sky islands of Arizona (Masta 2000) and the Rocky Mountains (DeChaine and Martin 2004) have shown that habitats moved southward and to lower elevations during glacial periods. These results suggest the coniferous forests in California would have been more contiguous and widely distributed at lower elevations during the last glacial period approximately 100,000 to 10,000 years ago, providing greater opportunity for migration among pygmy nuthatch populations at that time. During the present interglacial period, coniferous forests have contracted and elevated, becoming isolated. This study suggests that this, combined with the sedentary

nature of pygmy nuthatches, has resulted in the recent restriction of gene flow among mountain ranges.

Estimates of nucleotide diversity (Π) were similar to values reported in other intraspecific studies of passerines, while the number of haplotypes [per number of individuals] per population was generally lower (Table 6; Buerkle 1999, Sgariglia and Burns 2003). Values of F_u 's F_s for all populations were not significant (Table 6), indicating that population sizes are currently stable and that no populations have undergone a recent range expansion. These results further support the recent isolation of pygmy nuthatch populations and the role of vicariance in the separation of populations. Pygmy nuthatches have remained relatively sedentary during periods of interglacial isolation while gene flow during glacial periods was facilitated by proximity.

Phylogeography

Parsimony, neighbor-joining, and maximum likelihood analyses revealed strong to moderate support for four distinct haplotype clades: F-I-O, D-H-J-K-Q, A-B-E-L-M-C-G-P, and C-G-P by itself (Figures 5-7). Of these, only the F-I-O clade shows any association with geography.

Haplotypes F, I, and O are restricted to the San Gabriel and San Bernardino Mountains within the Transverse Ranges of southern California (Table 8). The remaining clades in the phylogenetic analysis are widespread in their distribution and occur in all or all but one of the sampled mountain ranges (Table 8).

Nested clade analysis revealed the overall pattern of historical gene flow in southern California pygmy nuthatch populations to be influenced by both past contiguous range expansion and restriction of gene flow with isolation by distance. Clade 1-7 (D-H-J-K-Q), which is represented in all but the Piute Mountains, was inferred to have arisen through contiguous range expansion. The majority of individuals represented in this clade are distributed in the San Bernardino and San Jacinto Mountains (Table 8), with fewer individuals represented to the north and south. This suggests that these may have been concentrated in a glacial refugium in the vicinity of the San Bernardino and San Jacinto Mountains and then expanded northward as glaciers retreated during interglacial episodes. This pattern of haplotype distribution, however, may be an artifact of the sampling design with the San Bernardino and San Jacinto Mountains being two of the most heavily sampled

ranges. Clades 2-3 (A-B-E-L-M-C-G-P; represented in all populations except Laguna Mountain) and 3-1 (entire cladogram) were inferred to be the result of restricted gene flow with isolation by distance. With the exception of a few widespread haplotypes (A, D, and H), all of the haplotypes are geographically restricted either from the San Bernardino Mountains to the south (haplotypes I-Q) or from the San Bernardino Mountains to the north (haplotypes B-C and E-G). This suggests that pygmy nuthatches have experienced significant but local gene flow during past glacial episodes as habitats became connected. The Mantel test also showed a significant isolation by distance effect among pygmy nuthatch collecting locations.

The maximum divergence time between southern California pygmy nuthatch haplotypes based on a 1.6-1.9% substitution rate (Fleischer et al. 1998) is approximately 250,000 to 300,000 years ago. While geographic populations are not reciprocally monophyletic, this estimate provides a maximum threshold value for the restriction of gene flow among populations, providing further evidence for the recent isolation of pygmy nuthatch populations. Thus southern California populations have become isolated on

disjunct mountain ranges well within the last 300,000 years.

Due to the sedentary nature of pygmy nuthatches it is likely that even though their habitat was widespread during glacial epochs, dispersal was still limited and migration occurred only among adjacent populations. Current populations on isolated mountain ranges exchanged genes with populations on adjacent isolated mountain ranges during glacial periods when their habitat was connected at lower elevations.

Subspecies Limits

Pygmy nuthatch populations in southern California exhibit clinal variation in morphology: mean size increases and plumage coloration becomes paler from north to south (Kingery and Ghalambor 2001). This clinal variation encompasses two described subspecies, but exactly where this subspecific break occurs is uncertain (Figure 4; Grinnell and Miller 1944, Norris 1958). Analysis of the distribution of sequence variation using AMOVA supports a significant break within the Transverse Ranges at the Cajon Pass between the San Gabriel and San Bernardino Mountains (Table 7). Twenty percent of the genetic differentiation

can be accounted by individuals north and south of the Cajon Pass. By comparison, the mtDNA data accounted for only 14% of the genetic variance across Banning Pass between the San Bernardino Mountains (Transverse Ranges) and the San Jacinto Mountains (Peninsular Ranges), the subspecific break supported by Grinnell and Miller (1944) based on morphological characters. This is similar to the level of variation (13.9%) explained by the overall AMOVA analyzing differences among all mountains ranges. Our mtDNA data revealed no statistically significant genetic differentiation between individuals in the Laguna Mountains and those to the north, providing no genetic support for Norris' (1958) proposed subspecific break.

Recent studies on comparative phylogeography of California fauna provide support for a subspecific break at the Cajon Pass, at least in an evolutionary perspective. A comprehensive review of phylogeographic studies of California fauna found that most species exhibit a distinct genetic break near (north of) the Transverse Ranges (Figure 4), although this pattern was not as apparent in birds (Calsbeek et al. 2003). More recently studies by Sgariglia and Burns (2003), Burns and Barhoum (2006), and Alexander and Burns (unpublished) have provided additional

phylogeographic information on California's avifauna supporting this genetic break near the Transverse Ranges.

In a study of the phylogeography of the wrentit (*Chamaea fasciata*), Burns and Barhoum (2006) found a strong genetic break among populations at the Cajon Pass, an area with no strong morphological variation and which did not conform to a described morphological subspecies break. This is nearly identical to the pattern found in this pygmy nuthatch study, except wrentit populations south of the Cajon Pass form a monophyletic grouping whereas pygmy nuthatch populations do not. The reason for the lack of monophyly observed in pygmy nuthatches may be differences in habitat during Pleistocene climate changes. The pygmy nuthatch is a resident of coniferous forest while the wrentit inhabits coastal sage scrub and chaparral. Both habitats probably experienced similar down slope and southward shifts during glacial periods, however, because coastal sage scrub and chaparral occur at lower elevation, they probably were more isolated at these times restricting gene flow among wrentit populations (Burns and Barhoum 2006). Alternatively, coniferous forests were probably more widespread and contiguous during glacial periods

(Barrowclough et al. 2004) facilitating opportunities for gene flow among adjacent pygmy nuthatch populations.

A study on the white-headed woodpecker (*Melanerpes albolarvatus*), with which the pygmy nuthatch is co-distributed in Pacific North America, also found a nearly monophyletic grouping of populations from Mount Pinos south through the Peninsular Ranges (Alexander and Burns unpublished). In this case, however, the authors postulated that the southern populations were colonized by a southward expansion from the north, and that monophyly was the result of a founder event or bottleneck as opposed to vicariance caused by the occurrence of a southern glacial refugium for wrentits (Alexander and Burns unpublished, Barhoum and Burns 2006). Thus while similar patterns of genetic differentiation are observed among many co-distributed taxa, the proposed evolutionary processes by which these patterns occur are different for different species.

This research supports three significant genetic breaks of varying degree at the northern edge of the San Gabriel Mountains, at the Cajon Pass between the San Gabriel and San Bernardino Mountains, and at the Banning Pass at the northern edge of the San Jacinto Mountains.

The strongest genetic break occurs at the Cajon pass, in contrast to the strongest morphological break observed by Grinnell and Miller (1944) occurring at the Banning Pass. That the subspecific breaks supported by mitochondrial DNA and morphological data do not correspond is not unusual or unprecedented. Genetically-based morphological traits may represent adaptations to local conditions, and would be subject to local environmental selective pressures that would not affect neutral mtDNA markers (Burns and Barhoum 2006). In a review of avian mitochondrial DNA studies, Zink (2004) found that 97% of continental North American bird subspecies lack a corresponding significant population genetic structure.

While a subspecific break at the Cajon Pass is supported by our mtDNA data and evidence from other species, caution is warranted. Much debate has occurred over the concept of subspecies and the criteria for diagnosing them (Patten and Unitt 2002). Recently Rising (2005) has argued that subspecies should be distinctive, geographically isolated populations. While southern California populations meet the criteria of being geographically isolated, these populations are more clinal and less distinctive. This study encompasses only a small

portion of the range of *S. p. melanotis*, which is widely distributed throughout western North America, and does not include samples from all of the populations representing *S. p. leuconucha*, particularly those in northern Baja California Norte, Mexico, and several populations located between the Laguna and San Jacinto Mountains in southern California. Therefore one must also consider that all of these populations may comprise a single subspecies. Even Grinnell and Miller (1944) recognized the cline in morphological variation and considered the placement of the subspecific break between *S. p. melanotis* and *S. p. leuconucha* "arbitrary".

Conservation Implications

The population and phylogenetic data generated from this research also has important application to the management and conservation of land in southern California. Long term habitat loss and fragmentation have made habitat conservation and connectivity a priority for land management in southern California. The South Coast Missing Linkages Project has recently identified the pygmy nuthatch as an important indicator for identifying and preserving dispersal corridors among large habitat preserves on public

and private lands (South Coast Wildlands 2005). The results of this research, however, suggest that pygmy nuthatches are a poor choice as indicators of corridor usage. The current restriction of gene flow among mountain ranges suggests that migration and dispersal among mountain ranges is a rare event. Due to their sedentary nature and habitat affinity, pygmy nuthatches would be better suited as indicators of habitat quality for coniferous forests. This research provides baseline genetic data which could be used to monitor changes in genetic composition of pygmy nuthatch populations over time, from which inferences about the health and quality of their coniferous forest habitat could be made. For example, if at some time in the future a population of pygmy nuthatches exhibited reduced genetic diversity relative to this baseline data, this could indicate a reduction in population size corresponding to a decline in the quality or availability of suitable habitat.

Conclusions

Pygmy nuthatch populations in southern California exhibit relatively low levels of genetic variation and differentiation. During glacial periods greater than about 10,000 years ago, coniferous forests were more widespread

and contiguous (Pielou 1991, Dorf 1976), allowing for extensive gene flow among populations. The current wide distribution of haplotypes and lack of lineage sorting in the phylogenetic analysis is the result of this historical gene flow. Within the last 10,000 years during the current interglacial period, coniferous forests have become isolated in disjunct mountain ranges. As a result pygmy nuthatch populations have undergone a recent restriction of gene flow, and show significant (although low) levels of genetic differentiation among mountain ranges. This genetic differentiation is greatest at the Cajon Pass, which may also represent a subspecific break between *S. p. melanotis* and *S. p. leuconucha*. If so, this subspecific break would occur further north than previously thought, and would not conform to proposed morphologically-based subspecific breaks.

APPENDIX A
HAPLOTYPE LISTS

3045-BASE PAIR HAPLOTYPE LIST

A1

CTTTGGATCACTCCTAGGTATCTGCTTAGTACACCCAAATCGTAACAGGCCTGCTACTAGCAACACACTACA
CCGCAGACACCTCCCTAGCCTTTGCCTCAGTCGCCCATACCTGCCGAAACGTCCAGTTCGGCTGACTTATC
CGCAATCTCCACGCAAACGGAGCCTCATTCTTCTTTATCTGCATCTACCTTCACATCGGCCGGGGAATCTA
CTACGGCTCATACCTAAATAAAGAGACCTGAAACGTAGGAGTTATTTCTCCTACTAACCCTCATAGCAACTG
CTTTTCGTAGGATACGTTCTACCATGAGGACAAATATCCTTCTGAGGGGCTACAGTCATCACCAACCTATTC
TCCGCAATCCCCTACATCGGCCAAACACTAGTAGAATGAGCCTGAGGTGGATTCTCTGTGACAAACCTAC
ATTAACACGATTCTTCGCCCTGCACTTCTTACTTCCATTTCGTCATCGCAGGCATTACATTAGTTCACCTCA
CCTTCTACACGAAACAGGCTCAAACAACCCCTTAGGAATCCCTCAGACTGTGACAAAATTCATTCCAC
CCGTACTACTCCACAAAAGACATCCTAGGGTTCGCTATCATGCTTATTTCTCCTAGTCACTCTAGCACTATT
CTCCCCAAACCTACTAGGAGACCCAGAAAACCTCACACCAGCTAACCCCTAGTACACCCCCACACATCA
AACCCGAATGATATTTCCCTATTTCGCATACGCTATTCTCCGATCCATCCCAAACAACTAGGAGGAGTACTA
GCCCTTGCCGCTCCGCTCCTAGTCTTATTCCTTATGCCCTTACTACACACATCCAAACAAAGCTCAATAAC
CTTCCGCCACTCTCACAAATCCTCTTCTGAACCTCTAGTAGCTGACCTCTTCATTCTTTACCTGAGTAGGCA
GCCAACAGTCGAACACCCATTTCATCATCATCGGACAACCTAGCTTCACTCGCCTACTTCCAAATCATCCTA
GTTCTGTTCCTACTCGCATCCATCCTAGAAAATAAACTCCTAAAACTTTAAACCACTCTAATAGTTTATAA
AAACATTTGGTCTTGTAAAGCCAAAGATTGAAGACTACACCCCTCTTATAGAGTTTCAACACCTCAGAAAAGAA
GGAATCAAACCCCTTATCGCCAACCTCCCAAAGCTGGAATTTTTATTAACCTACTTTCTGACCCCTCCCTCCC
CCCTAAACAGCCCGAATAGCCCCCGTGATAACCCCGCACCAACTCCAGCACTACAAATAAAGTTAACAA
CAATCCCCAACCACCAATCAAAAACAGCCCCGCCCCACGAATAAAAACTGCTACCCACCAAAAATCCG
TAGGAGCCACAGACAACCCCTTACCATTCACTGTACCCTATTACCACACACTCAACCCCTCCTCCAATT
ACCACCCCCACAGCAACAACCAACCCCATCCCTAACCCATACCCAACAACCCCTAAATCTCCCCACGACTC
CGGATAAGGATCCGCCGCCAAAGAGACAGAATAAACAAACACCACCAACATCCCACCTAAATAAACATAAA
CCAACACCAAGGACACGAAAGAAACCCCTAAACTTACCAATCAACCACACCCAGCAACAGATGCCAAAAC
AACCCAACCACCCCATAAATAAGGAGAAGGATTAGACGCCACTCCCAAAGCCCCCAAACAAAGCACAGCCC
TAAAAATAATACAAATTTCTATCATAAATCCCACCTGGCCTCTCTCCAGGATCAATGGCCTGAAAAACCATC
GTTATAAAAAATTTAACTACGGGAATCCTACCCCCCTTCCCCCCCCAGCATGTTTTCTCATGCTTTTTTCG
GGTATGTGGTCTGCATCGCTCTCTTTGCCACATCAGACAACCTACTATAATGTAGGAACTCCACGCCATA
CGTATTGCCTTGCCTCAAAAACCTCAAACATATCGCCATAAGATGTTGTTCAAGCTATTACACCTCTAG
GTACTTTCTGTTTTCAGGTACCATAAAACCAAGTGATCCTTACCTCCAGCCGAAGCCGAGCTTACTTAA
GATCGACTATATCTCCCTGCGCATACAAAATCCTCTCCTCGGATACGGGTGAATGTCCCAGCACACCTTTGC
ATTCCAAAGTCATAAACCTTAGCCACCTCCTAGGATTTATTACTCTCCAACAGCCTTCAAGAACTCCCAA
GCCAGAGAACCTGGTTATCTATTAATCGTGTCTTCTCACGAGAACCAGGCTACTCAACGTGAGTTTATACTTA
CGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAACGTGCCACCGGTTGTAATTTGAGGA
ACGCTACTTGATTATTCCTATTTCAATTTGTTCTTACAGATACATTTGCTTGGGGATGGACGCTCCTCATAT
ATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTTCTTTTTGGGGTCTCTTCAATAAGCC
CTTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCACATGACCGGCGAACATACGACTGCC
CGAAACCACCTAGAATGTCATGGTTATTGGATTAAGTAGACCAACTTCGGCACTGATGCACTTTGACCCCA
TTCATGAAACCCGCGCTATTTACCTTTTTAAGTAGCCGATAATGTAATGGTTGGCAGACATAGTTAATTTT
TTCTCTTCTTCTAGGAATTTATCACCCAAACCCCAATTTTTTCGTCAATTTTTTCGTTCTGTTTTTTTTTTT
CTTTTTTGTGTTGTCATTTTTCGTTTCATTTTAGACAAACAAACAGCTATATAACCCTACATTTTATCAAAA
CTTAACCATCAATTCATTCACAACAAACTTTCTCTAAAATTCACCCACTCACAAATCAATCACCATTTCAT
CATTTTTAACAACGTTTAAACAAACAAACTACCATCATAAACATCACAAACCCAGCAACTTACACCCATCCC
TAAAAACCAACAAAAATACAAAACAACAATTCAGAATAACCAATAACATCCTACCACCAAC

A2

CTTTGGATCACTCCTAGGTATCTGCTTAGTACACCCAAATCGTAACAGGCCTGCTACTAGCAACACACTACA
CCGCAGACACCTCCCTAGCCTTTGCCTCAGTCGCCCATACCTGCCGAAACGTCCAGTTCGGCTGACTTATC
CGCAATCTCCACGCAAACGGAGCCTCATTCTTCTTTATCTGCATCTACCTTCACATCGGCCGGGGAATCTA
CTACGGCTCATACCTAAATAAAGAGACCTGAAACGTAGGAGTTATTTCTCCTACTAACCCTCATAGCAACTG
CTTTTCGTAGGATACGTTCTACCATGAGGACAAATATCCTTCTGAGGGGCTACAGTCATCACCAACCTATTC

TCCGCAATCCCCACATCGGCCAAACACTAGTAGAATGAGCCTGAGGTGGATTCTCTGTGACAAACCCCTAC
ATTAACACGATTCTTCGCCCTGCACCTCTTACTTCCATTCGTTCATCGCAGGCATTACATTAGTTACCTCA
CCTTCCCTACACGAAACAGGCTCAAACAACCCCTTAGGAATCCCCCAGACTGTGACAAAATTCCATTCAC
CCGTACTACTCCACAAAAGACATCCTAGGGTTCGCATATCATGCTTATTTCTCCTAGTCACTCTAGCACTATT
CTCCCCAAACCTACTAGGAGACCCAGAAAACCTCACACCAGCTAACCCCTAGTACACCCCCACACATCA
AACCCGAATGATATTTCCATTCGCATACGCTATTCCTCCGATCCATCCCAAACAAACTAGGAGGAGTACTA
GCCCTTGCCGCCCTCCGTCCTAGTCTTATTCCTTATGCCCTTACTACACACATCCAAACAACGCTCAATAAC
CTTCGGCCCACTCTCACAATCCTCTTCTGAACCTAGTAGCTGACCTCTTCATTCTTACCTGAGTAGGCA
GCCAACAGTGAACACCCATTTCATCATCATCGGACAACCTAGCTTCACCTCGCCTACTTCACAATCATCCTA
GTTCTGTTCCCCTCGCATCCATCCTAGAAAATAAACTCCTAAAACCTTAAACCAACTCTAATAGTTTATAA
AAACATTTGGTCTTGTAAGCCAAAGATTGAAGACTACACCCCTTCTTAGAGTTTCAACACCTCAGAAAGAAA
GGAATCAAACCCCTTATCGCCAACCTCCCAAAGCTGGAATTTTTTATTAACTACTTTCTGACCCTCCCTTCCC
CCCTAAACAGCCCGAATAGCCCCCGTGATAACCCCGCACCAACTCCAGCACTACAAAATAAAGTTAACAA
CAATCCCCAACCCAAATCAAAAACAGCCCGCCCCCACGAATAAAAACTGCTACCCCAACAAATCCG
TACGAGCCACAGACAACCCCTTACCATTCACTGTACCCTATTCACCACACACTCAACCCCTCCTCCAATT
ACCACCCCCACAGCAACAACCAACCCCATCCCTAACCCATACCCAAACAACCCCTAAATCTCCCCACGACTC
CGGATAAGGATCCGCCGCCAAAGAGACAGAATAAAACAACACCACCAACATCCACCTAAATAAACTATAA
CCAACACCAAGGACACGAAAGAAACCCCTAAACTTACCAATCAACCACACCCAGCAACAGATGCCAAAACC
AACCCAACCCCCATAATAAGGAGAAGGATTAGACGCCACTCCCAAAGCCCCAAAACAAGCACAGCCC
TAAAAATAATACAAATTCATATAAATCCCACCTGGCCTCTCTCCAGGATCAATGGCCTGAAAAACCATC
GTTATAAAAAATTTAACTACGGGAATCCTACCCCCCTTCCCCCCCAGCATGTTTTCTCATGCTTTTTTCG
GGTATGTGGTTCTGCATCGCTCTCTTGGCCACATCAGACAACCTACTATAATGTAGGAAACTCCACGCCATA
CGTATTGCCCTTGCCCTCCAAAACCTCAAACATTCATCGCCCATAGATGTTGTTCAAGCTATTACACCTTAG
GTACTTTCTTGTTCAGGTACCATAAAACCAAGTGATCCTACCTCCAGCCGAAGCCGCAAGCGTTACTTAA
GATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACGGGTGAATGTCCAGCACACCTTTGC
ATTCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTCTCCAACAGCCTTCAAGAACTCCCAA
GCCAGAGAACCTGGTTATCTATTAATCGTGTTCCTCACGAGAACCGAGCTACTCAACGTCACTTATACTTA
CGGTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAACGTGCCACCGGTTGTAATTTACAGGA
ACGTACTTTGATTATTCCTATTCATTTGTTCTTACAGATACATTTGCTTGGGGATGGACCTCCTCATAT
ATATCCGTAATCGCGCTAGCCACAGGGCCGTTTTTACTCTTCTTTTTCTTTTTGGGGTCTCTTCAATAAGCC
CTTCAAGTGCGTAGCAGGAGTTATCTTCTCTTGACATGTCCATCACATGACCGGCGAACATACGACTGCC
CGAAACCACCTAGAATGTCATGGTTATTGGATTAAGTAGACCAACTTCGGCACTGATGCACCTTTGACCCCA
TTCATGAAACCCGCGCTATTTACCTTTTTAAGTAGCCGATAATGTAATGGTTGGCAGACATAGTTAATTTT
TTCTCTTCTTCTAGGAATTATCACCCAAACCCCAATTTTTCGTCAATTTTTCTGTTCTGTTTTTTTTTTTT
CTTTTTGTTTGTCAATTTTTCGTTCAATTTTAGACAAACAACAGCTATATAACCCCTACATTTATCAAACA
CTTAACCATCAATTCATTCACAACAACCTTCTCTAAAATTCACCCACTCACAAATCAATCACCATTTCAT
CATTTTAAACAACGTTTAAACAACAACAACTACCATCATAAACATCACAAACCCAGTAACTTACACCCATCCC
TAAAAACCAAACAATAACAAACAACAATTCAGAATAACCAATAACATCCTACCACCAAC

B1

CTTTGGATCACTCCTAGGTATCTGCTTAGTACCCCAAATCGTAACAGGCCTGCTACTAGCAACACACTACA
CCGACAGACACCTCCCTAGCCTTTGCCTCAGTCGCCCATACCTGCCGAAACGTCCAGTTCCGGCTGACTTATC
CGCAATCTCCACGCAAACGGAGCCTCATCTCTTTATCTGCATCTACCTTCACATCGGCCGGGGAATCTA
CTACGGCTCATACCTAAATAAAGAGACCTGAAACGTAGGAGTTATTCCTACTAACCCCTCATAGCAACTG
CTTTCGTAGGATACGTTCTACCATGAGGACAAATATCCTTCTGAGGGGCTACAGTCATCACCACCTATTTC
TCCGCAATCCCCACATCGGCCAAACACTAGTAGAATGAGCCTGAGGTGGATTCTCTGTGACAAACCCCTAC
ATTAACACGATTCTTCGCCCTGCACCTCTTACTTCCATTCGTTCATCGCAGGCATTACATTAGTTTACCTCA
CCTTCTTACAGTAAACAGGCTCAAACAACCCCTAGGAATCCCCCAGACTGTGACAAAATTTCCATTTCCAC
CCGTACTACTCCACAACAAGACATCCTAGGGTTCGCATATCATGCTTATTTCTCCTAGTCACTCTAGCACTATT
CTCCCCAAACCTACTAGGAGACCCAGAAAACCTCACACCAGCTAACCCCTAGTACACCCCCACACATCA
AACCCGAATGATATTTCCATTCGCATACGCTATTCCTCCGATCCATCCCAAACAACTAGGAGGAGTACTA
GCCCTTGCCGCCCTCCGTCCTAGTCTTATTCCTTATACCCCTTACTACACACATCCAAACAACGCTCAATAAC
CTTCGGCCCACTCTCACAATCCTCTTCTGAACCTAGTAGCTGACCTCTTCATTCTTACCTGAGTAGGCA
GCCAACAGTGAACACCCATTTCATCATCATCGGACAACCTAGCTTCACCTCGCCTACTTCACAATCATCCTA

GTTCTGTTCCCACTCGCATCCATCCCTAGAAAATAAACTCCTAAAACTTTAAACCAACTCTAATAGTTTATAA
AAACATTGGTCTTGTAGCCAAAGATTGAAGACTACACCCCTTCTTAGAGTTTCAACACCTCAGAAAAGAAA
GGAATCAAACCCCTTATCGCCAACCTCCCAAAGCTGGAATTTTTATTAACTACTTTCTGACCCCTCCCTTCCC
CCCTAAACAGCCCGAATAGCCCCCGGTGATAACCCCGCACCAACTCCAGCACTACAAATAAAGTTAACAA
CAATCCCCAACCACCAATCAAAAACAGCCCCGCCCCCACGAATAAAAACTGCTACCCACCAAAAATCCG
TACGAGCCACAGACAACCCCTCACCATTCACTGTACCCTATTCACCACACACTCAACCCCTCCTCCAATT
ACCACCCCCACAGCAACAACCAACCCCATCCCTAACCCATACCCAACAACCCCTAAAATCTCCCACGACTC
CGGATAAGGATCCGCCGCCAAAGAGACAGAATAAACAAACACCACCAACATCCCACCTAAAATAAACTATAA
CCAACACCAAGGACACGAAAGAAACCCCTAAACTTACCAATCAACCACACCCAGCAACAGATGCCAAAACC
AACCCAACCACCCATAATAAGGAGAAGGATTAGACGCCACTCCCAAAGCCCCAAAACAAAGCACAGCCC
TAAAAATAATACAAATTCATATAAATCCCACCTGGCCTCTCTCCAGGATCAATGGCCTGAAAAACCATC
GTTATAAAAAATTTAACTACGGGAATCCTACCCCCCTTCCCCCCCAGCATGTTTTCTCATGCTTTTTTCG
GGTATGTGGTCTGTCATCGCTCTCTTGGCCACATCAGACAACACTATAATGTAGGAACTCCACGCCATA
CGTATTGCCCTTGCCTCCAAAACCTCAAACATTATCGCCATAAGATGTTGTCAAGCTATTACACCTTAG
GTACTTTCTTGTTCAGGTACCATAAAACCAAGTGATCCTACCTCCAGCCGAAGCCGCAAGCGTTACTTAA
GATCGACTATAATCTCCCTGCGCATACAAATCCTCTCCTCGGATACGGGTGAATGTCCAGCACACCTTTGC
ATTCCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTCTCCAACAGCCTTCAAGAACTCCCAA
GCCAGAGAACCTGGTTATCTATTAATCGTGTCTCAGGAGAACCGAGCTACTCAAGCTCAGTTATACCTTA
CGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAACGTGCCACCGGTTGTAATTTAGGA
ACGCTACTTGATTATTCCTATTCAATTTGTTCTTACAGATACATTTGCTTGGGGATGGACGCTCCTCATAT
ATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTTTCTTTTTGGGGTCTCTTCAATAAAC
CTTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCACATGACCGGCGAACATACGACTGCC
CGAAACCACCTAGAATGTATGGTTATTGGATTAAGTAGACCAACTTCGGCACTGATGCACTTTGACCCCA
TTCATGAAACCCGCGTATTTACCTTTTTAAGTAGCCGATAATGTAATGGTTGGCAGACATAGTTAATTTT
TTCTCTTCTCTAGGAATTATCACCCAAACCCCAATTTTTCGTCAATTTTTCGTTCTGTTTTTTTTTTTT
CTTTTTGTTTGTCTATTTTTGTTTATTTTAGACAAACAAACAGCTATATAACCCCTACATTTATCAAAA
CATTAAACCATCAATTCATTCACAACAAACTTCCCTCTAAAATTCACCCACTCACAAATCAATCACCATT
TATTTTAAACAACGTTTAAACAACAAACACTACCATCATAAACATCACAAACCCAGCAACTTACACCCATCCC
TAAAAACCAACAAAAATACAAAACAACAATTCAGAATAACCAATAACATCCTACCACCAAC

D1

CTTTGGATCACTCCTAGGTATCTGCTTAGTCACCCAAATCGTAACAGGCCTGCTACTAGCAACACACTACA
CCGACAGACACCTCCCTAGCCTTTGCCTCAGTCGCCCATACCTGCCGAAACGTCCAGTTCCGGCTGACTTATC
CGCAATCTCCACGCAAACGGAGCCTCATCTCTTTATCTGCATCTACCTTCACATCGGCCGGGAATCTA
CTACGGCTCATAACCTAAATAAAGAGACCTGAAACGTAGGAGTTATCTCCTACTAACCCCTCATAGCAACTG
CTTTGCTAGGGTACGTTCTACCATGAGGACAAATATCCTTCTGAGGGGCTACAGTCATACCAACCTATTC
TCCGCAATCCCTACATCGGCCAAACACTAGTAGAATGAGCCTGAGGTGGATTCTCTGTCGACAACCCCTAC
ATTAACACGATTTCTCGCCCTGCATCTTACTTCCATTCGTATCGCAGGCATTACATTTAGTTACCTCA
CCTTCCCTACAGAAACAGGCTCAAACAACCCCTTAGGAATCCCTCAGACTGTGACAAAATTCATTTCCAC
CCGTACTACTCCACAAAAGACATCCTAGGGTTGCTATCATGCTTATTTCTCCTAGTCACTCTAGCACTATT
CTCCCCAAACCTACTAGGAGACCCAGAAAACCTTACACCAGCTAACCCCTAGTCACACCCCCACACATCA
AACCCGAATGATATTTCTTATTCGCATACGCTATTTCCGATCCATCCCAAACAAACTAGGAGGAGTACTA
GCCCTTGGCCCTCCGTCCTTAGTCTTATTCCTTATAACCTTACTACACACATCCAAACAACGCTCAATAAC
CTTCCGCCACTCTCACAAATCCTCTTCTGAACCTTAGTAGCTGACCTCTTCACTTCTTACCTGAGTAGGCA
GCCAACAGTCGAACACCCATTCATCATCATCGGACAACCTAGCTTCACTCGCCTACTTCAAAATCATCCTA
GTTCTGTTCCCACTCGCATCCATCCTAGAAAATAAACTCCTAAAACCTTTAAACCAACTCTAATAGTTTATAA
AAACATTGGTCTTGTAGCCAAAGATTGAAGACTACACCCCTTCTTAGAGTTTCAACACCTCAGAAAAGAAA
GGAATCAAACCCCTTATCGCCAACCTCCCAAAGCTGGAATTTTTATTAACTACTTTCTGACCCCTCCCTCCC
CCCTAAACAGCCCGAATAGCCCCCGGTGATAACCCCGCACCAACTCCAGCACTACAAATAAAGTTAACAA
CAATCCCCAACCACCAATCAAAAACAGCCCCGCCCCCACGAATAAAAACTGCTACCCACCAAAAATCCG
TACGAGCCACAGACAACCCCTCACCATTCACTGTACCCTATTCACCACACACTCAACCCCTCCTCCAATT
ACCACCCCTACAGCAACAACCAACCCCATCCCTAACCCATACCCAACAACCCCTAAAATCCCCCAGGACTC
CGGATAAGGATCCGCCGCCAAAGAGACAGAATAAACAAACACCACCAACATCCCACCTAAAATAAACTATAA
CCAACACCAAGGACACGAAAGAAACCCCTAAACTTACCAATCAACCACACCCAGCAACAGATGCCAAAACC

AACCCAACCACCTCCATAATAAGGAGAAGGATTAGACGCCACTCCCAAAGCCCCAAAACAAAGCACAGCCC
TAAAAATAATACAAATTCCTATCATAAATCCCACCTGGCCTCTCTCCAGGATCAATGGCCTGAAAAACCATC
GTTATAAAAAATTTAACTACGGGAATCCTACCCCCCCCCTCCCCCCCCAGCATGTTTTCTCATGCTTTTTTCG
GGTATGTGGTCTGCATCGCTCTCTTTGCCACATCAGACAACACTATAAATGTAGGAACTCCACGCCATA
CGTATTGCCTTGCCCTCCAAAACCTCAAACATTTATCGCCCATAAAGATGTTGTTCAAGCTATTACACCTCTAG
GTACTTTCTTGTTCAGGTACCATAAAACCAAGTGATCCTACCTCCAGCCGAAGCCGCAAGCGTTACTTAA
GATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACGGGTGAATGTCCCAGCACACCTTTGC
ATTCCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTCTCCAACAGCCTTCAAGAACTCCCAA
GCCAGAGAACCTGGTTATCTATTAATCGTGTCTCTCACGAGAACCAGCTACTCAACGTCAGTTATACTTA
CGTTATTTGGTGTGAGGGGCATGCATCTAATAAACTTGATCTCCAACGTGCCACCGGTTGTAATTTTCAGGA
ACGCTACTTGATTATTCCTATTCAATTGTTCTTACAGATACATTTGCTTGGGGATGGACGCTCCTCATAT
ATATCCGTAATCGCGGCTAGCCACAGGCGTTTTTACTCTTCTTTTTCTTTTTGGGGTCTCTTCAATAAGCC
CTTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCACATGACCGGCGAACATACGACTGCC
CGAAACCACCTAGAATGTCATGGTTATTGGATTAAGTAGACCAACTTCGGCCTGATGCACCTTTGACCCCA
TTCATGAAACCCGCGCTATTTACCTTTTTAAGTAGCCGATAATGTAATGGTTGGCAGACATAGTTAATTTT
TTCTCTTCTCTAGGAATTTACCCAAACCCCAATTTTTCGTCAATTTTTCGTTCGTTTTTTTTTTTTTTT
CTTTTTGTGTGTCATTTTTCGTTCATTTTAGACAAACAAACAGCTATATAACCCTACATTTATCAAACA
CTTAACCATCAATTCATTACAAACAAACTTTCTCTAAAATTCACCCACTCACAAATCAATCACCATTTCAT
CATTTTAAACAACGTTTAAACAACAAACTACCATCATAAACATCACAAACCCAGCAACTTACACCCATCCC
TAAAACCAAACAAAATACAAAACAACAATTCAGAAATAACCAATAACATCCTACCACCAAC

E1

CTTTGGATCACTCCTAGGTATCTGCTTAGTCACCCAAATCGTAACAGGCCTGCTACTAGCAACACACTACA
CCGACAGACACCTCCCTAGCCTTTGCCTCAGTCGCCCATACCTGCCGAAACGTCCAGTTCGGCTGACTTATC
CGCAATCTCCACGCAAACGGAGCCTCATCTTCTTTATCTGCATCTACCTTCACATCGGCCGGGGAATCTA
CTACGGCTCATACCTAAATAAAGAGACCTGAAACGTAGGAGTTAATCTCCTACTAACCCCTCATAGCAACTG
CTTTCTGTTAGGATACGTTCTACCATGAGGACAAATATCCTCTGAGGGGCTACAGTCATCACCAACCTATT
TCCGCAATCCCCACATCGGCCAAACACTAGTAGAATGAGCCTGAGGTGGATTCCTGTGACAAACCTTAC
ATTAACACGATTTCTCGCCCTGCACCTCTTACTTCCATTCGTTCATCGCAGGCATTACATTAGTTACCTCA
CCTTCCCTACACGAAACAGGCTCAAACAACCCCTTAGGAATCCCCTCAGACTGCGACAAAATTCCTATCCAC
CCGTACTACTCCACAAAAGACATCCTAGGGTTCGCTATCATGCTTATTTCTCCTAGTCACTCTAGCACTATT
CTCCCCAAACCTACTAGGAGACCCAGAAAACCTTACACCAGCTAACCCCTAGTTCACACCCCCACACATCA
AACCCGAATGATATTTCTTATTCGCATACGCTATTTCTCGGATCCATCCCAAACAAACTGGGAGGAGTACTA
GCCCTTGCCGCTCCGCTCTTAGTCTTATTCCTTATACCTTACTACACACATCCAAACAACGCTCAATAAC
CTTCCGCCCACCTCTCACAAATCCTCTTCTGAACCTTAGTAGCTGACCTCTTCAATCTTACCTGAGTAGGCA
GCCAACAGTCGAACACCCATTCATCATCATCGGACAACCTAGCTTCACTCGCCTACTTCACAATCATCCTA
GTTCTGTTCCCACTCGCATCCATCCTAGAAAATAAACTCCTAAAACCTTAAACCAACTTAATAGTTTATAA
AAACATTTGGTCTTTGTAAGCCAAAGATTGAAGACTACACCCCTTCTTAGAGTTTCAACACCTCAGAAAGAAA
GGAATCAAACCCCTTATCGCCAACCTCCCAAAGCTGGAATTTTTTATTAAACTACTTTCTGACCCCTCCCTTCCC
CCCTAAACAGCCCGAATAGCCCCCGGTGATAACCCCGCACCAACTCCAGCACTACAAAATAAAGTTAAACAA
CAATCCCCAACCACCAATCAAAAACAGCCCCGCCCCCAGCAATAAAAAACTGCTACCCCAACCAAAATCCG
TACGAGCCACAGACAACCCCTTACCATTCACTGTACCCTATTCACCACACACTCAACCCCTCCTCCAAT
ACCACCCCCACAGCAACAACCAACCCCATCCCTAACCCATACCCAACAACCCCTAAATCTCCCCACGACTC
CGGATAAGGATCCGCCGCAAAGAGACAGAATAAACAAACACCACCAACATCCCACCTAAATAAACTATAA
CCAACACCAAGGACACGAAAGAAACCCCTAAACTTACCAATCAACCACACCCAGCAACAGATGCCAAAAC
AACCCAAACCCCAATAATAAGGAGAAGGATTAGACGCCACTCCCAAAGCCCCAAAACAAAGCACAGCC
TAAAAATAATACAAATTCCTATCATAAATCCCACCTGGCCTCTCTCCAGGATCAATGGCCTGAAAAACCATC
GTTATAAAAAATTTAACTACGGGAATCCTACCCCCCCCCTCCCCCCCCAGCATGTTTTCTCATGCTTTTTTCG
GGTATGTGGTCTGCATCGCTCTCTTTGCCACATCAGACAACACTATAAATGTAGGAACTCCACGCCATA
CGTATTGCCTTGCCCTCCAAAACCTCAAACATTTATCGCCCATAAAGATGTTGTTCAAGCTATTACACCTCTAG
GTACTTTCTTGTTCAGGTACCATAAAACCAAGTGATCCTACCTCCAGCCGAAGCCGCAAGCGTTACTTAA
GATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACGGGTGAATGTCCCAGCACACCTTTGC
ATTCCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTCTCCAACAGCCTTCAAGAACTCCCAA
GCCAGAGAACCTGGTTATCTATTAATCGTGTCTCTCACGAGAACCAGCTACTCAACGTCAGTTATACTTA

CGGTTATTGGTGT CAGGGGCATACATCTAATAAACTTGATCTCCAACGTGCCACCGGTTGTAATTT CAGGA
ACGCTACTTGATTATTCCTATTC AATTGTTCTT CACAGATACATTTGCTTGGGGATGGACGCTCCTCATAT
ATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTTCTTTTGGGGTCTCTTCAATAAACC
CTTCAAGTGC GTAGCAGGAGTTATCTTCCCTTTGACATGTCCATCACATGACCGGCGAACATACGACTGCC
CGAAACCACCTAGAAATGTCATGGTTATTGGATTAAGTAGACCAACTTCGGCACTGATGCACCTTTGACCCCA
TTCATGAAACCCGCGCTATTTACCTTTTTAAGTAGCCGATAATGTAATGGTTGGCAGACATAGTTAATTTT
TTCTCTTCTTCTAGGAATTATCACCCAAACCCCAATTTTTTCGTCAATTTTTTCGTTCTGTTTTTTTTTTTTT
CTTTTTTGTGTCATTTTCGTTCAATTTTAGACAAACAAACAGCTATATAACCCTACATTTATCAAACA
CTTAACCATCAATTCATTCACAACAAACTTTCCCTCTAAAATTCACCCACTCACAAATCAATCACCATTTCAT
CATTTTAAACAACGTTTAAACAAACAAAACCTACCATCATAAACATCACAAACCCAGCAACTTACACCCATCCC
TAAAAACCAAAACAAAAATACAAAACAACAATTCAGAATAACCAATAACATCCTACCACCAAC

H1

CTTTGGATCACTCCTAGGTATCTGCTTAGT CACCCAAATCGTAACAGGCCTGCTACTAGCAACACACTACA
CCG CAGACACCTCCCTAGCCTTTGCCTCAGTCGCCCATACCTGCCGAAACGTCCAGTTCGGCTGACTTATC
CGCAATCTCCACGCAAACGGAGCCTCATTCTTCTTTATCTGCATCTACCTTACATCGGCCGGGGAATCTA
CTACGGCTCATACTAAATAAAGAGACCTGAAACGTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTG
CTTTCGTAGGGTACGTTCTACCATGAGGACAAATATCCTTCTGAGGGGCTACAGTCATCACCACCTATTC
TCCGCAATCCCTTACATCGGCCAAACACTAGTAGAATGAGCCTGAGGTGGATTCTCTGTGCACAACCCCTAC
ATTAACACGATTCTTCGCCCTGCACCTTCTTACTTCCATTCGTTCATCGCAGGCATTACATTAGTTACACTCA
CCTTCCCTACACGAAAACAGGCTCAAACAACCCCTAGGAATCCCTCAGACTGTGACAAAATTCATTCCAC
CCGTACTACTCCACAAAAGACATCCTAGGGTTCGCTATCATGCTTATTCTCCTAGTCACTCTAGCACTATT
CTCCCCAAACCTACTAGGAGACCCAGAAAACCTCACACCAGCTAACCCTTAGTACACCCCCACACATCA
AACCCGAATGATATTTCTATTCGCATACGCTATTCTCCGATCCATCCCAAACAAACTAGGAGGAGTACTA
GCCCTTGCCGCTCCGTCCTAGTCTTATTCCTTATAACCTTACTACACACATCCAAACAACGCTCAATAAC
CTTCCGCCCACTCTACAAATCCTCTTCTGAACCTAGTAGCTGACCTCTTCTTACTTACCTGAGTAGGCA
GCCAACAGTCGAACACCCATTCATCATCTCGGACAACCTAGCTTCACTCGCTACTTCAACAATCATCCTTA
GTTCTGTTCCCACTCGCATCCCTAGAAAATAAACTCCTAAAACCTTAAACCAACTCTAATAGTTTATAA
AAACATTTGGTCTTGTAAGCCAAAGATTGAAGACTACACCCCTTCTTAGAGTTTCAACACCTCAGAAAGAAA
GGAATCAAACCCCTTATCGCCAACCTCCCAAAGCTGGAATTTTTATTAAACTACTTTCTGACCCTCCCTTCCC
CCCTAAAACAGCCCGAATAGCCCCCGTGATAACCCCGCACCAACTCCAGCACTACAAATAAAGTTAAACA
CAATCCCAACCAACCAATCAAACAACAGCCCCGCCCCCAGCAATAAAAAACTGCTACCCCAACCAAAATCCG
TAGGAGCCACAGACAACCCCTTACCATTCACTGTACCCTATTACCACACACTCAACCCCTCCTCCAATT
ACCACCCCTACAGCAACAACCAACCCCATCCCTAACCCATACCCAACAACCCCTAAATCCCCCAGGACTC
CGGATAAGGATCCGCGCCAAAGAGACAGAATAAACAAACACCACCAACATCCCACCTAAATAAACTATAA
CCAACACCAAGGACACGAAAGAAACCCCTAAACTTACCAATCAACCACACCCAGCAACAGATGCCAAAACC
AACCCAAACCACTCCATAATAAGGAGAAGGATTAGACGCCACTCCCAAAGCCCCCAAACAAAGCACAGCCC
TAAAAATAATACAAATTCATCATAATTCACCACCTGGCCTCTCTCCAGGATCAATGGCCTGAAAAACCATC
GTTATAAAAAATTTAACTACGGGAATCCTACCCCCCTTCCCCCCCCAGCATGTTTTCTCATGCTTTTTTCG
GGTATGTGGTCTGTCATCGCTCTCTTTGCCACATCAGACAACCTACTATAATGTAGGAACTCCACGCCATA
CGTATTGCCTTGCCCTCAAAAACCTCAAACATTATCGCCATAAGATGTTGTTCAAGCTATTACACCTCTAG
GTACTTTCTTGTTCAGGTACCATAAACC AAGTGATCCTACCTCCAGCCGAAGCCGCAAGCGTTACTTAA
GATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACGGGTGAATGTCCCAGCACACCTTTGC
ATTCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTCTCCAACAGCCTTCAAGAACTCCCAA
GCCAGAGAACCTGGTTATCTATTAATCGTGTCTTCTCACGAGAACCAGCTACTCAACGTCAAGTTATACTTA
CGGTTATTGGTGT CAGGGGCATGCATCTAATAAACTTGATCTCCAACGTGCCACCGGTTGTAATTT CAGGA
ACGCTACTTGATTATTCCTATTC AATTGTTCTT CACAGATACATTTGCTTGGGGATGGACGCTCCTCATAT
ATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTTCTTTTTGGGGTCTCTTCAATAAACC
CTTCAAGTGC GTAGCAGGAGTTATCTTCCCTTTGACATGTCCATCACATGACCGGCGAACATACGACTGCC
CGAAACCACCTAGAAATGTCATGGTTATTGGATTAAGTAGACCAACTTCGGCACTGATGCACCTTTGACCCCA
TTCATGAAACCCGCGCTATTTACCTTTTTAAGTAGCCGATAATGTAATGGTTGGCAGACATAGTTAATTTT
TTCTCTTCTTCTAGGAATTATCACCCAAACCCCAATTTTTTCGTCAATTTTTTCGTTCTGTTTTTTTTTTTT
CTTTTTTGTGTCATTTTCGTTCAATTTTAGACAAACAAACAGCTATATAACCCTACATTTATCAAACA
CTTAACCATCAATTCATTCACAACAAACTTTCCCTCTAAAATTCACCCATTACAAATCAATCACCATTTCAT

CATTTTAAACAACGTTTAAACAAACAAAACCTACCATCATAAACATCACAAACCCAGCAACTTACACCCATCCC
TAAAAACCAAACAAAAATACAAAACAACAATTCAAGAATAACCAATAACATCCTACCACCAAC

H2

CTTTGGATCACTCCTAGGTATCTGCTTAGTACCCCAAATCGTAACAGGCCTGCTACTAGCAACACACTACA
CCGAGACACCTCCCTAGCCTTTGCCTCAGTCGCCCATACCTGCCGAAACGTCCAGTTCGGCTGACTTATC
CGCAATCTCCACGCAAACGGAGCCTCATTCTTCTTTATCTGCATCTACCTTACATCGGCCGGGGAATCTA
CTACGGCTCATAACCTAAATAAAGAGACCTGAAACGTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTG
CTTTCGTAGGGTACGTTCTACCATGAGGACAAATATCCTTCTGAGGGGCTACAGTCATACCAACCTATTC
TCCGCAATCCCCTACATCGGCCAAACACTAGTAGAATGAGCCTGAGGTGGATTCTCTGTGACAAACCTTAC
ATTAACACGATTCTTCGCCCTGCACCTTCTTACTTCCATTTCGTTCATCGCAGGCATTACATTAGTTCACCTCA
CCTTCCCTACACGAAAACAGGCTCAAACAACCCCTAGGAATCCCCTCAGACTGTGACAAAATTTCCATTCCAC
CCGTACTACTCCACAAAAGACATCCTAGGGTTCGCTATCATGCTTATTCTCCTAGTCACTCTAGCACTATT
CTCCCCAAACCTACTAGGAGACCCAGAAAACCTCACACCAGCTAACCCCTAGTACACCCCCACACATCA
AACCCGAATGATATTTCTTATTCGCATACGCTATTCTCCGATCCATCCCAAACAAACTAGGAGGAGTACTA
GCCCTTGCCGCTCCGTCTAGTCTTATTCTTATAACCTTACTACACACATCCAAACAACGCTCAATAAC
CTTCCGCCACTCTCACAAATCCTCTTCTGAACCTCTAGTAGCTGACCTCTTCATCCTTACCTGAGTAGGCA
GCCAACAGTCGAACACCCATTCATCATCATCGGACAACCTAGCTTCACTCGCCTACTTACAATCATCCTA
GTTCTGTTCCCACTCGCATCCATCCTAGAAAATAAACCTCTAAAACCTTTAACCAACTCTAATAGTTTATAA
AAACATTGGTCTTGTAAGCCAAAGATTGAAGACTACACCCCTTCTTAGAGTTTCAACACCTCAGAAAGAAA
GGAATCAAACCTTATCGCCAACTCCCAAAGCTGGAATTTTTATTAAACTACTTTCTGACCCTCCCTTCCC
CCCTAAACAGCCCCGAATAGCCCCCGTGATAACCCCGCACCAACTCCAGCACTACAAATAAAGTTAAACA
CAATCCCCAACCAATCAAAAACAGCCCCGCCCCACGAATAAAAACTGCTACCCCACCAAATCCG
TAGGAGCCACAGACAACCCCTTACCATTCACTGTACCCTATTACCACACACTCAACCCCTCCTCCAATT
ACCACCCCTACAGCAACAACCAACCCCATCCCTAACCCATACCCAACAACCCCTAAATCCCCCACGACTC
CGGATAAGGATCCGCCCAAAGAGACAGAAATAACAACACCACCAACATCCCACCTAAATAAACTATAA
CCAACACCAAGGACAGAAAGAAACCCCTAAACTTACCAATCAACCACACCAGCAACAGATGCCAAAAC
AACCAACCACTCCATAATAAGGAGAAGGATTAGACGCCACTCCCAAAGCCCCAAACAAGCACAGCC
TAAAAATAATCAAAATCTATCATAAATCCCACCTGGCCTCTCTCCAGGATCAATGGCCTGAAAACCATC
GTTATAAAAAATTTAACTACGGGAATCCTACCCCCCTTCCCCCCCCAGCATGTTTTCTCATGCTTTTTTCG
GGTATGTGGTTCTGCATCGCTCTCTTTGCCACATCAGACAACCTACTATAATGTAGGAACTCCACGCCATA
CGTATTGCCTTGCCCTCCAAAACCTCAAACATTATCGCCATAAGATGTTGTTCAAGCTATTACACCTCTAG
GTACTTTCTTGTTTCAGGTACCATAAACCCTAAGTATCCTACCTCCAGCCGAAGCCGAAGCGTTACTTAA
GATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACGGGTGAATGTCCCAGCACACCTTTGC
ATTCCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTCTCCAACAGCCTTCAAGAACTCCCAA
GCCAGAGAACCTGGTTATCTATTAATCGTGTCTCAGGAGAACCGAGCTACTCAACGTGAGTTATACTTA
CGGTTATTGGTGTGAGGGCATGCATCTAATAAACTTGATCTCCAACGTGCCACCGGTTGTAATTTAGGA
ACGCTACTTGATTATTCTTATTCATTTGTTCTTACAGATACATTTGCTTGGGGATGGACGCTCCTCATAT
ATATCCGTAATCGCGGCTAGCCACAGGCGTTTTTTACTCTTCTTTTTCTTTTTGGGGTCTCTTCAATAAACC
CTTCAAGTGCCTAGCAGGAGTTATCTTCCCTTTGACATGTCCATCACATGACCGGCGAACATACGACTGCC
CGAAACCACCTAGAATGTCATGGTTATTGGATTAAGTAGACCAACTTCGGCACTGATGCACTTTGACCCCA
TTCATGAAACCCGCGCTATTTACCTTTTTAAGTAGCCGATAATGTAATGGTTGGCAGACATAGTTAATTTT
TTCTCTTCTTCTAGGAATGTCACCCAAACCCCAATTTTTTCGTCAATTTTTCGTTCGTTTTTTTTTTTTT
CTTTTTTGTGTTGTCATTTTTCGTTCATTTTAGACAAAACAACAGCTATATAACCCTACATTTATCAAACA
CTTAACCATCAATTCATTCACAACAAACTTCCCTAAAATTCACCCACTCACAAATCAATCACCATTTCAT
CATTTTAAACAACGTTTAAACAAACAAAACCTACCATCATAAACATCACAAACCCAGCAACTTACACCCATCCC
TAAAAACCAAACAAAAATACAAAACAACAATTCAAGAATAACCAATAACATCCTACCACCAAC

H3

CTTTGGATCACTCCTAGGTATCTGCTTAGTACCCCAAATCGTAACAGGCCTGCTACTAGCAACACACTACA
CCGAGACACCTCCCTAGCCTTTGCCTCAGTCGCCCATACCTGCCGAAACGTCCAGTTCGGCTGACTTATC
CGCAATCTCCACGCAAACGGAGCCTCATTCTTCTTTATCTGCATCTACCTTACATCGGCCGGGGAATCTA
CTACGGCTCATAACCTAAATAAAGAGACCTGAAACGTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTG
CTTTCGTAGGGTACGTTCTACCATGAGGACAAATATCCTTCTGAGGGGCTACAGTCATACCAACCTATTC

TCCGCAATCCCCACATCGGCCAAACACTAGTAGAATGAGCCTGAGGTGGATTCTCTGTGACAACCCCTAC
ATTAACACGATTCTTCGCCCTGCACCTCTTACTTCCATTTCGTTCATCGCAGGCATTACATTAGTTACACCTCA
CCTTCCCTACACGAAACAGGCTCAAACAACCCCTAGGAATCCCCCTCAGACTGTGACAAAATTCATTCCAC
CCGTACTACTCCACAAAAGACATCCTAGGGTTCGCTATCATGCTTATTCTCCTAGTCACTCTAGCACTATT
CTCCCCAAACCTACTAGGAGACCCAGAAAACCTCACACCAGCTAACCCCTAGTACACCCCCACACATCA
AACCCGAATGATATTTCCATTCGCATACGCTATTTCCGATCCATCCCAAACAAACTAGGAGGAGTACTA
GCCCTTGCCGCCCTCCGTCCTAGTCTTATTCCTTATACCCCTTACTACACACATCCAAACAACGCTCAATAAC
CTTCCGCCCACTCTCACAATCCTCTTCTGAACTCTAGTAGCTGACCTCTTCATTCTTACCTGAGTAGGCA
GCCAACAGTCGAACACCCATTCATCATCATCGGACAACCTAGCTTCACTCGCCTACTTACAATCATCCTA
GTTCTGTTCCCACTCGCATCCATCCTAGAAAAATAAACTCCTAAAACTTAAACCAACTTAATAGTTTATAA
AAACATTGGTCTTGTAAGCCAAAGATTGAAGACTACACCCCTCTTAGAGTTTCAACACCTCAGAAAGAAA
GGAATCAAACCCCTTATCGCCAACCTCCCAAAGCTGGAATTTTTTATTAACTACTTTCTGACCCCTCCCTTCCC
CCCTAAACAGCCCGAATAGCCCCCGTGATAACCCCGCACCAACTCCAGCACTACAAAATAAAGTTAACAA
CAATCCCCAACCACCAATCAAAAACAGCCCGCCCCCAGCAATAAAAACTGCTACCCCAACAAAATCCG
TAGGAGCCACAGACAACCCCTTACCATTCACTGTACCCTATTCACCACACACTCAACCCCTCCTCCAATT
ACCACCCCTACAGCAACAACCAACCCCATCCCTAACCCATACCCAACAACCCCTAAATCCCCCAGACTC
CGGATAAGGATCCGCCGCCAAAGAGACAGAAATAAACAAACACCACCAACATCCCACCTAAATAAACTATAA
CCAACACCAAGGACACGAAAGAAACCCCTAACTTACCAATCAACCACACCAGCAACAGATGCCAAAACC
AACCAACCCTCCATAATAAGGAGAAGGATTAGACGCCACTCCCAAAGCCCCAAAACAAAGCACAGCCC
TAAAAATAATACAAATTCATATAATCCCACCTGGCTCTCTCCAGGATCAATGGCTGAAAAACCATC
GTTATAAAAAATTTAACTACGGGAATCCTACCCCCCTTCCCCCCCAGCATGTTTTCTCATGCTTTTTTCG
GGTATGTGGTCTGCATCGCTCTTTTGCACATCAGACAACCTACTATAATGTAGGAACTCCACGCCATA
CGTATTGCCCTTGCCCTCCAAAACCTCAAACATTATCGCCCATAAAGATGTTGTTCAAGCTATTACACCTCTAG
GTACTTTCTTTGTTTCAGGTACCATAAAACCAAGTGATCCTACCTCCAGCCGAAGCCGCAAGCGTTACTTAA
GATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACGGGTGAATGTCCCAGCACACCTTTGC
ATTCCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTCTCCAACAGCCTTCAAGAACTCCCAA
GCCAGAGAACCTGGTTATCTATTAATCGTGTCTTCTCACGAGAACCAGCTACTCAACGTCAGTTATACTTA
CGGTATTTGGTGTACAGGGGCATGCATCTAATAAACTTGATCTCCAACGTGCCACCGGTTGTAATTTAGGA
ACGCTACTTGATTTATCCTATTCATTTGTTCTTACAGATACATTTGCTTGGGGATGGACGCTCCTCATAT
ATATCCGTAATCGCGCTAGCCACAGCCGTTTTTACTCTTCTTTTCTTTTGGGGTCTCTTCAATAAAC
CTTCAAGTGCCTAGCAGGATTATCTTCTCTTGTGACATGTCCATCACATGACCCGGAACATACGACTGCC
CGAAACCACCTAGAATGTATGGTTATTGGATTAAAGTAGACCAACTTCGGCACTGATGCACTTTGACCCCA
TTATGAAACCCGCGCTATTTACCTTTTTTAAGTAGCCGATAATGTAATGGTTGGCAGACATAGTTAATTTT
TTCTCTTCTTCTAGGAATTTATCACCCAAACCCCAATTTTTTCGTCAATTTTTTCGTTCTGTTTTTTTTT
CTTTTTTGTGTTGTCATTTTTCGTTTCTTTTAGACAAAACAAACAGCTATATAACCCTACATTTATCAAAA
CTTAACCATCAATTCATTCACAACAAACTTTCTCTAAAAATCACCCACTCACAAAATCAATCACCATTCA
CATTTTAAACAACGTTTAAACAAACAAACTACCATCATAAACATCACAAACCCAGCAACTTACACCCATCCC
TAAAAACCAACAAAATACAAAACAACAATCAAGAAATAACCAATAACATCCTACCACCAAC

I1

CTTTGGATCACTCCTAGGTATCTGCTTAGTTCACCCAAATCGTAACAGGCCTGCTACTAGCAACACACTACA
CCGCAGACACCTCCCTAGCCTTTGCCCTCAGTCGCCCATACCTGCCGAAACGTCCAGTTTCGGCTGACTTATC
CGCAATCTCCACGCAAACGGAGCCTCATCTTCTTTATCTGCATCTACCTTACATCGGCCGGGGAATCTA
CTACGGCTCATACTAAATAAAGAGACCTGAAACGTAGGAGTTATTCTCCTACTAACCCCTCATAGCAACTG
CTTTCTGTTAGGTTACGTTCTACCATGAGGACAAATATCCTTCTGAGGGGCTACAGTCACTACCAACCTATT
TCCGCAATCCCCACATCGGCCAAACACTAGTAGAATGAGCCTGAGGTGGATTCTCTGTGACAACCCCTAC
ATTAACACGATTCTTCGCCCTGCACCTCTTACTTCCATTTCGTTCATCGCAGGCATTACATTAGTTACACCTCA
CCTTCCCTACACGAAACAGGCTCAAACAACCCCTAGGAATCCCCCTCAGACTGTGACAAAATTCATTCCAC
CCGTACTACCTCCACAAAAGACATCCTAGGGTTCGCTATCATGCTTATTCTCCTAGTCACTCTAGCACATAT
CTCCCCAAACCTACTAGGAGACCCAGAAAACCTCACACCAGCTAACCCCTAGTACACACCCCAACACATCA
AACCCGAATGATATTTCCATTCGCATACGCTATTTCTCCGATCCATCCCAAACAAACTAGGAGGAGTACTA
GCCCTTGCCGCCCTCCGTCCTAGTCTTATTCCTTATACCCCTTACTACACACATCCAAACAACGCTCAATAAC
CTTCCGCCCACTCTCACAATTTCTTCTGAACTCTAGTAGCTGACCTCTTCATTCTTACCTGAGTAGGCA
GCCAACAGTCGAACACCCATTCATCATCATCGGACAACCTAGCTTCACTCGCCTACTTACAATCATCCTA

GTTCTGTTCCCACCTCGCATCCATCCTAGAAAATAAACTCCTAAAACCTCTAACCAACTCTAATAGTTTATAA
AAACATTGGTCTTTGTAAGCCAAAGATTGAAGACTACACCCCTTCTTAGAGTTTCAACACCTCAGAAAAGAAA
GGAATCAAACCCCTTATCGCCAACCTCCCAAAGCTGGAATTTTTATTAAACTACTTTCTGACCCCTCCCTTCCC
CCCTAAACAGCCCCGAATAGCCCCCGGTGATAACCCCGCACCAACTCCAGCACTACAAATAAAGTTAACAA
CAATCCCCAACCACCAATCAAAAACAGCCCCGCCCCCAGCAATAAAAAACTGCTACCCACCAAAAATCCG
TACGAGCCACAGACAACCCCTTACCATTCACTGTACCCTATTACCCACACACTCAACCCCTCCTCCAATT
ACCACCCCCACAGCAACAACCAACCCCATCCCTAACCCATACCCAACAACCCCTAAATCCCCCAGACTC
CGGATAAGGATCCGCCGCCAAAGAGACAGAATAAACAAACACCACCAACATCCCACCTAAATAAACTATAA
CCAACACCAAGGACACGAAAGAACTCCTAAACTTACCAATCAACCACACCCAGCAACAGATGCCAAAACC
AACCCAACCACCCATAATAAGGAGAAGGATTAGACGCCACTCCCAAAGCCCCAAAACAAAGCACAGCCC
TAAAAATAATACAAATTCCTATCATAATTCCCACCTGGCCTCTCTCCAGGATCAATGGCCTGAAAAACCATC
GTTATAAAAAATTTAACTACGGGAATCCTACCCCCCTTCCCCCCCCAGCATGTTTTCTCATGCTTTTTTCG
GGTATGTGGTCTGTCATCGCTCTCTTTGCCACATCAGACAACACTATAATGTAGGAAAACCTCACGCCATA
CGTATTGCCCTTGCCTCCAAAACCTCAAACATTATCGCCATAAGATGTTGTTCAAGCTATTACACCTCTAG
GTACTTTCTTGTTCAGGTACCATAAACCCAAGTGATCCTACCTCCAGCCGAAGCCGCAAGCGTTACTTAA
GATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACGGGTGAATGTCCAGCACACCTTTGC
ATTCCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTCTCCAACAGCCTTCAAGAACTCCCAA
GCCAGAGAACCTGGTTATCTATTAATCGTGTCTCTCACGAGAACCAGCTACTCAACGTCAAGTATACTTA
CGGTTATTTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAACGTGCCACCGGTTGTAATTTAGGA
ACGCTACTTGATTATTCCTACTCAATTTGTTCTTACAGATACATTTGCTTGGGGATGGACGCTCCTCATAT
ATATCCGTAATCGCGCTAGCCACAGGCCGTTTTTACTCTTCTTTTCTTTTGGGGTCTCTTCAATAAAC
CTTCAAGTGCGTAGCAGGAGTTATCTTCTCTTGACATGTCCATCACATGACCGGCCAACAATACGACTGCC
CGAAACCACCTAGAATGTATGGTTATTTGGATTAAGTAGACCAACTTCGGCACTGATGCACTTTGACCCCA
TTCATGAAACCCGCGTATTTACCTTTTAAAGTAGCCGATAATGTAATGGTTGGCAGACATAGTTAATTTT
TTCTCTTCTCTAGGAATTTACCCAAACCCCAATTTTTCGTCAATTTTTCGTTTCGTTTTTTTTTTTTTT
CTTTTTGTTTTGTCATTTTTCGTTTATTTTAGACAAACAAACAGCTATATAACCCCTACATTTATCAAAA
CTTAACCATCAATTCATTCACAACAACTTTCCCTCTAAAAATCACCCACTCACAAATCAATCACCATTCA
CATTTTAAACAACGTTTAAACAACAAACTACCATCATAAACATCACAAACCCAGCAACTTACACCCATCCC
TAAAAACCAACAAAAATACAAAACAACAATTCAGAAATAACCAATAACATCCTACCACCAAC

J1

CTTTGGATCACTCCTAGGTATCTGCTTAGTACCCCAAATCGTAACAGGCCCTGCTACTAGCAACACACTACA
CCGACAGACACCTCCCTAGCCTTTGCCTCAGTCGCCCATACCTGCCGAAACGTCCAGTTCGGCTGACTTATC
CGCAATCTCCACGCAACGGAGCCTCATCTCTTTATCTGCATCTACCTTCACATCGGCCGGGGAATCTA
CTACGGCTCATACTAAATAAAGAGACCTGAAACGTAGGAGTTATTTCTCTACTAACCCCTCATAGCAACTG
CTTTTCGTAGGGTACGTTCTACCATGAGGACAAATATCCTTCTGAGGGGCTACAGTCATCACCAACCTATTC
TCGGCAATCCCTACATCGGCCAAACACTAGTAGAATGAGCCTGAGGTGGATTCTCTGTGACAAACCTAC
ATTAACACGATTCTTCGCCCTGCACTTCTTACTTCCATTTCGTATCGCAGGCATTACATTAGTTACCTCA
CCTTCCCTACAGAAACAGGCTCAAACAACCCCTAGGAATCCCTCAGACTGTGACAAAATTCATTCAC
CCGTACTACTCCACAAAAGACATCCTAGGGTTTCGCTATCATGCTTATTTCTCCTAGTCACTCTAGCACTATT
CTCCCCAAACCTACTAGGAGACCCAGAAAACCTTACACCCAGCTAACCCCTAGTACACCCCCACACATCA
AACCCGAATGATATTTCCCTATTTCGCATACGCTATTTCTCCGATCCATCCCAAAACAACTAGGAGGAGTACTA
GCCCTTGGCCCTCCGTCCTAGTCTTATTCCTTATACCCCTTACTACACACATCCAAAACACGCTCAATAAC
CTTCGCCCACTCTCACAAATCCTCTTCTGAACCTTAGTAGCTGACCTCTTCATTTTACCTGAGTAGGCA
GCCAACCAGTCGAACACCCATTCATCATCATCGGACAACCTAGCTTCACTCGCCTACTTACAATCATCCTA
GTTCTGTTCCCCTCGCATCCATCCTAGAAAATAAACTCCTAAAACCTTAAACCAACTCTAATAGTTTATAA
AAACATTGGTCTTTGTAAGCCAAAGATTGAAGACTACACCCCTTCTTAGAGTTTCAACACCTCAGAAAAGAAA
GGAATCAAACCCCTTATCGCCAACTCCCAAAGCTGGAATTTTTATTAAACTACTTTCTGACCCCTCCTTCCC
CCCTAAACAGCCCCGAATAGCCCCCGGTGATAACCCCGCACCAACTCCAGCACTACAAAATAAAGTTAACAA
CAATCCCCAACCACCAATCAAAAACAGCCCCGCCCCCAGCAATAAAAAACTGCTACCCACCAAAAATCCG
TACGAGCCACAGACAACCCCTTACCATTCACTGTACCCTATTACCCACACACTCAACCCCTCCTCCAATTT
ACCACCCCTACAGCAACAACCAACCCCATCCCTAACCCATACCCAACAACCCCTAAATCCCCCAGACTC
CGGATAAGGATCCGCCGCCAAAGAGACAGAATAAACAAACACCACCAACATCCCACCTAAATAAACTATAA
CCAACACCAAGGACACGAAAGAAACCCCTAAACTTACCAATCAACCACACCCAGCAACAGATGCCAAAACC

AACCCAACCCTCCATAATAAGGAGAAGGATTAGACGCCACTCCCAAAGCCCCAAAACAAAGCACAGCCC
TAAAAATAATACAAATTCATCATAAATCCCACCTGGCCTCTCTCCAGGATCAATGGCCTGAAAAACCATC
GTTATAAAAAATTTAACTACGGGAATCCTACCCCCCCTTCCCCCCCAGCATGTTTTCTCATGCTTTTTTCG
GGTATGTGGTTCTGCATCGCTCTCTTTGCCACATCAGACAACTACTATAATGTAGGAAACTCCACGCCATA
CGTATTGCCTTGCCTCCAAAACCTCAAACATTATCGCCATAAGATGTTGTTCAAGCTATTACACCTCTAG
GTACTTTCTTGTTCAGGTACCATAAAACCCAAGTGATCCTACCTCCAGCCGAAGCCGCAAGCGTTACTTAA
GATCGACTATATCTCCCTGCGCATACCAATCCTCTCCTCGGATACGGGTGAATGTCCAGCACACCTTTGC
ATTCCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTCTCCAACAGCCTTCAAGAACTCCCAA
GCCAGAGAACCCTGGTTATCTATTAATCGTGTCTTCTCACGAGAACCAGCTACTCAACGTCAAGTTACTTAA
CGGTTATTGGTGTGAGGGGCATGCATCTAATAAACTTGATCTCCAACGTGCCACCGGTTGTAATTTAGGA
ACGCTACTTGATTATTCCTATTCAATTTCTTACAGATACATTTGCTTGGGGATGGACGCTCCTCATAAT
ATATCCGTAATCGCGCTAGCCACAGGCCGTTTTTACTCTCTTTCTTTTGGGGTCTCTCAATAAAACC
CTTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCACATGACCGGCGAACATACGACTGCC
CGAAACCACCTAGAATGTATGGTTATTGGATTAAGTAGACCAACTTCGGCACTGATGCACTTTGACCCCA
TTCATGAAACCCGCGCTATTTACCTTTTTAAGTAGCCGATAATGTAATGGTTGGCAGACATAGTTAATTTT
TTCTCTTCTCTAGGAATTATCACCCAAACCCCAATTTTTCTGTCATTTTTCTGTTGTTTTTTTTTTTT
CTTTTTTGTGTTCATTTTTCTGTTTCTTTTTAGACAAACAAACAGCTATATAACCTTACATTTATCAAAACA
CTTAACCATCAATTCATTCACAACAAACTTTCTCTAAAAATCACCCACTCACAAATCAATCACCATTTCAT
CATTTTAAACAACGTTTAAACAACAAACTACCATCATAAACATCACAAACCCAGCAACTTACACCCATCCC
TAAAAACCAACAAAAATACAAAACAACAATTCAGAATAACCAATAACATCCTACCACCAAC

L1

CTTTGGATCACTCCTAGGTATCTGCTTAGTCAACCAATCGTAACAGGCCTGCTACTAGCAACACACTACA
CCGAGACACCTCCCCTAGCCTTTGCCTCAGTCGCCCATACCTGCCGAAACGTCCAGTTTCGGCTGACTTATC
CGCAATCTCCACGCAAACGGAGCCTCATTTCTTTTATCTGCATCTACCTTCACATCGGCCGGGGAATCTA
CTACGGCTCATACCTAAATAAAGAGACCTGAAACGTAGGAGTTATTTCTCTACTAACCCCTCATAGCAACTG
CTTTCTGATAGGATACGTTCTACCATGAGGACAAATATCCTTCTGAGGGGCTACAGTCATCACCAACCTATTC
TCCGCAATCCCCTACATCGGCCAAACACTAGTAGAATGAGCCTGAGGTGGATTTCTCTGTGACAACCCCTAC
ATTAACAGATTTCTCGCCCTGCATTTCTTACTTCCATTCGTCATCGCAGGCATTACATAGTTTACCTCA
CCTTTCCCTACACGAAACAGGCTCAAACAACCCCTTAGGAATCCCCTCAGACTGTGACAAAATTCATTCAC
CCGTACTACTCCACAAAAGACATCCTAGGGTTCTGCTATCATGCTTATTTCTCTAGTCACTCTAGCACTATT
CTCCCCAAACCTACTAGGAGACCCAGAAAACCTTACACCCAGCTAACCCCTAGTCACACCCCCACACATCA
AACCCGAATGATATTTCCCTATTTCGCATACGCTATTCTCCGATCCATCCCAAAACAACTAGGAGGAGTACTA
GCCCTTGCCGCTCCGCTCTAGTCTTATTTCTTATACCCCTTACTACACACATCCAAAACACGCTCAATAAC
CTTCCGCCACTCTCACAAATCCTCTTCTGAACCTTAGTAGCTGACCTCTTCATTTCTTACCTGAGTAGGCA
GCCAACAGTCGAACACCCATTCATCATCATCGGACAACTAGCTTCACTCGCCTACTTACAATCATCCTA
GTTCTGTTCCCCTCGCATCCATCCTAGAAAATAAACTCCTAAAACTTTAACCAACTTAATAGTTTATAA
AAACATTTGGTCTTGTAAGCCAAAGATTGAAGACTACACCCCTTCTTAGAGTTTCAACACCTCAGAAAGAAA
GGAATCAAACCCCTTATCGCCAACCTCCCAAAGCTGGAATTTTTTATTAACTACTTTCTGACCCCTCCCTTCCC
CCCTAAACAGCCCGAATAGCCCCCGTGATAACCCCGCACCAACTCCAGCACTACAAATAAAGTTAACAA
CAATCCCCAACCACCAATCAAAAACAGCCCCGCCCCACGGAATAAAAAACTGCTACCCACCAAAAATCCG
TACGAGCCACAGACAACCCCTTACCATTCACTGTACCCTATTACCCACACACTCAACCCCTCCTCCAATTT
ACCACCCCCACAGCAACAACCAACCCCATCCCCTAACCCATACCCAAACCCCTAAATCTCCCCACGACTC
CGGATAAGGATCCGCCGCCAAAGAGACAGAATAAAACAAACACCACCAACATCCCACCTAAATAAACTATAA
CCAACACCAAGGACACGAAAGAAACCCCTAAACTTACCAATCAACCACACCCAGCAACAGATGCCAAAACC
AACCAACACCCCATATAAAGGAGAAGGATTAGACGCCACTCCCAAAGCCCCAAAACAAAGCACAGCCC
TAAAAATAATACAAATTCATCATAAATCCCACCTGGCCTCTCTCCAGGATCAATGGCCTGAAAAACCATC
GTTATAAAAAATTTAACTACGGGAATCCTACCCCCCCTTCCCCCCCAGCATGTTTTCTCATGCTTTTTTCG
GGTATGTGGTTCTGCATCGCTCTCTTTGCCACATCAGACAACTACTATAATGTAGGAAACTCCACGCCATA
CGTATTGCCTTGCCTCCAAAACCTCAAACATTATCGCCCAATAAGATGTTGTTCAAGCTATTACGCTCTAG
GTACTTTCTTGTTCAGGTACCATAAAACCCAAGTGATCCTACCTCCAGCCGAAGCCGCAAGCTTACTTAA
GATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACGGGTGAATGTCCAGCACACCTTTGC
ATTCCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTCTCCAACAGCCTTCAAGAACTCCCAA
GCCAGAGAACCCTGGTTATCTATTAATCGTGTCTTCTCACGAGAACCAGCTACTCAACGTCAAGTTACTTAA

CGGTTATTGGTGTGTCAGGGGCATACATCTAATAAACTTGATCTCCAACGTGCCACCGGTTGTAATTCAGGA
ACGCTACTTGATTATTCCTATTCAATGCTCTTCACAGATACATTTGCTTGGGGATGGACGCTCCTCATAT
ATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTTCTTTTGGGGTCTCTTCAATAAAACC
CTTCAAGTGCCTAGCAGGAGTTATCTTCCCTTTGACATGTCCATCACATGACCGGCGAACATACGACTGCC
CGAAACCACCTAGAATGTATGGTTATTGGATTAAGTAGACCAACTTCGGCACTGATGCACTTTGACCCCA
TTCATGAAACCCGCGCTATTTACCTTTTTAAGTAGCCGATAATGTAATGGTTGGCAGACATAGTTAATTTT
TTCTCTTCTTAGGAATTAFCACCCAAACCCCAATTTTTTCGTCAATTTTTTCGTTTCGTTTTTTTTTTTTT
CTTTTTTGTGTCATTTTTCGTTCATTTTAGACAAACAAACAGCTATAAACCCTACATTTATCAAAACA
CTTAACCATCAATTCATTCACAACAAACTTTCCCTCTAAAATTCACCCCTCACAAAATCAATCACCATTCA
CATTTTAAACAACGTTTAAACAACAAACTACCATCATAAACATCACAAACCCAGCAACTTACACCCATCCC
TAAAAACCAACAAAAATACAAAACAACAATTCAGAATAACCAATAACATCCTACCACCAAC

R1

CTTTGGATCCCTCCTAGGTATCTGCCTAATTACTCAAATGTAACAGGCCCTCTACTAGCAACACACTACA
CCGCAGACACCTCCCTAGCCTTCGCCTCAGTCGCCCATACCTGCCGAAATGTCCAATTCGGCTGACTTATC
CGCAACCTCCACGCAATGGGGCCTCATCTTCTTCATCTGCATCTACCTTCACATCGGCCGAGGAATCTA
CTACGGCTCATACTAAATAAAGAAACCTGAAACGTAGGAGTCAATCTATTAATAACCCTTATAGCAACTG
CTTTCGTAGGATATGTCTACCATGAGGACAAATATCCTTTTGGAGGGCTACAGTCAATACCAACCTATTT
TCCGCAATCCCTACATCGGCCAAACACTAGTAGAATGAGCCTGAGGCGGATTTCTGTAGACAACCCAC
GTTAACACGATTTCTCGCCCTTCACTTCCACTTCCATTCGTTCATCGCAGGCATTACACTAGTCCACCTCA
CCTTCCCTACACGAAACAGGCTCAAACAACCCACTAGGAATTCCTCAGACTGTGACAAAATTCATTCAC
CCATACTACTCCACAAAAGATATCCTAGGATTCGCCATCATACTAATTCCTTAGTACCCTGGCACTATT
CTCCCCTAACCTGCTAGGAGACCCAGAAAACCTCACACCAGCCAACCCACTAGTTCACACCCCCACACATTA
AACCTGAATGATACTTCCCTATTTGCATACGCTATTTCTCGATCGATCCCAAACAAACTGGGAGGAGTATTA
GCCCTTGCCGCTCTGTCTTAGTCTATTTCTCATACTCTACTACACACATCCAAACAGCGCTCAATAAC
CTTCCGCCACTCTCACAAATCCTCTTCTGAACCTTAGTAGCTGATCTCCTCATTTCTTACCTGAGTAGGCA
GCCAACAGTGAACATCCATTCATTCATCGGACAACACTAGCCTCACTTGTACTTCACAATCATCCWA
GTCTCTGTTCCCGCTCGCATCCCTAGAAAACAAGCTCCTAAAACCTTTAACCAACTTAATAGTTTATAA
AAACATTTGGTCTTGTAAAGCCAAAGATTGAAGACTACACCCCTTCTTAGAGTTTATACACCTCAGAAAGAAA
GGAATCAAACCTTATCGCCAACCTCCAAAGCTGGAATTTTWTAAACTACTTCTGACCTCCCTCCCT
CCCTAAACAGCCCGAATTGCCCCCGTGATAACCCCGCACCAACTCCAACACCACAAACAAAGTCAACAA
CAACCTCAACCACCAATTAAAAACAGCCCCGCCCCCACGAATAAAAAACTGCTACCCCACTAAAAATCCG
TAGGAGCCACAGACAACCCCTCACCATTACCCGTACCCTATTCACCACAAACTCAACCCCCCTCCAACA
ACTACCCCAACAATAATAACCAGCCCTATCCCCAATCCATACCCAAACACTCCTAAGCTCCCTCACGATTC
CGGATAAGGGTCCGCCGCCAGAGAAACAGAATAAACAAACACCACCAACATCCCACCTAAATAAACTATTA
CCAGCACCAGAGACACGAAAGAAACCCCTAAACTTACCAGTCAACCACACCAGCAACAGAAAGCCAAAACC
AACCCAAACCACCCATAATAAGGGGAAGGATTAGACGCYACCCCAAGCCCCAAAACAAAACACAGGCC
TAAAAATAATACAAATTCATCATAATTCACCTGGCTCTCTCCAGGATCAGTGGCTGAAAAACCACC
GTTACAAAATTTAACTATGGGAACCTCACCCCCCCCTTCCCCCCCCAGCATGTTTTCTCATGCTTTTTCG
GGTATGTGGTCTGTCATCGCTCTTTTGCCACATCAGACAACACTATAATGTAGGATACTCCACGCCATA
CGTATTGCCTTGCCCTCCAAAACCTCAAACATTAATCTCCATAAGATGTTGTTCAAGCTATCACACCTCTAG
GCACATTTCTGTTTACGGTACCATAAAACCCAGTATCCTACCTCCAGCCGAAGCCGCAAGCGTTACTTAA
GATCGACTACATCTCCCTGCACATACAAACCTCTCCTCGGATACGGGTGAATGTCCAGCACACCTTTGA
ATTCCTCAAGGTCACGAACCTTAGCCACCTCCTAGGAATTTATACTCTCCAACAGCTTTCAAGGACTCCCAA
GCCAGAGAACCCTGGTTATCTATTAATCGTGCCCTCACGAGAACCGAGCTACTCAACGTCAGTTATACTTT
CGGTTATTGGTGTGAGGGCATAACATCTAATAAACTTGATCTCAACGTCGCCACCGGTTGTAATTTGAGGA
ATACTGATTGATTGATTCCTACTCAATTTGCTCTTACAGATACATTTGCTTGGGGTCTCTCAATAAGCC
ATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTTTTTTCTTTTTTGGGGTCTCTCAATAAGCC
CTTCAAGTGCCTAGCAGGAGTTATCTTCCCTTTGACATGTCCATCACATGACCGGCGAACATACGACTGCC
CGAAACCACCTAGAATGTATGGTTATTGGATTAAGTAGACCAACTTCGGCACTGATGCACTTTGACCCCA
TTCATGAAACCCGCGCTATTTACCTCTTTAAGTAGCAGATAATGAAATGCTTGGCAGACATAGTTATTTAT
TTCTCCTTTTCTAGGATTTACCACCCAAACCCCAATTTTTTTCGTTCATTTTTTCGTTTCGTTTCGTTTTT
TTTTTTTTTCTTTTTTGTGTCATTTTTTGTTCATTTTAAACAAATAAACAGCTATAATAACCCCTACATTT
ATCAAACACTTAACCATCAATTCGTTTACAACAACTTTCCCTCTAAAATCCACTACCCACAAACCAATC

GTTACCCATCATTTTATTAACACTTAACAAACAAAACCATTAACATAAACACTTCAAATCCAGCAACTTAC
ACCCATCCCTAAAACCAAACAAAATACAAAACAACGACCAAAGAATAATCAATAACACCCTACCACTAA
T

1639-BASE PAIR HAPLOTYPE LIST

A

GTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGATACGTTCTACCATGAGGACAAAT
 ATCCTTCTGAGGGGCTACAGTCATACCAACCTATTCTCCGCAATCCCCTACATCGGCCAAAACACTAGTAG
 AATGAGCCTGAGGTGGATTCTCTGTGACAAACCTACATTAACACGATTCTTCGCCCTGCACTTCTTACTT
 CCATTCGTTCATCGCAGGCATTACATTAGTTCACCTCACCTTCTTACACGAAAACAGGCTCAAACAACCCCT
 AGGAATCCCCCTCAGACTGTGACAAAATCCATTCACCCCGTACTACTCCACAAAAGACATCCTAGGGTTCG
 CTATCATGCTTATTCTCCTAGTCACTCTAGCACTATTCTCCCCAAACCTACTAGGAGACCCAGAAAACCTC
 ACACCAGCTAACCCCTAGTACACCCCCACACATCAAACCCGAATGATATTTCCCTATTTCGCATACGCTAT
 TCTCCGATCCATCCCAAACAACACTAGGAGGAGTACTAGCCCTTGCCGCCTCCGTCTTAGTCTTATTCTTA
 TGCCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
 TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
 TGGGGTGGTTGGTTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGGGTTCCTTTTCG
 TGTCCTTGGTGTGGTTATAGTTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTATTCTGTCTCTTTGGCG
 GCGGATCCTTATCCGGAGTCGTGGGGAGATTTAGGGGTGTTGGGTATGGGTAGGGATGGGGTTGGTTGT
 TGCTGTGGGGTGGTAATTGGAGGAGGGGTTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
 TGCTGTGGCTCGTACGGATTTTGGTGGGGTAGCAGTTTTTTTATTTCGTGGGGGGCGGGCTGTTTTGATT
 GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCAGGGGTTATCAATCCTACCTCC
 AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAAATCCTCTCCTCGGATACG
 GGTGAATGTCCAGCACACCTTTGCATTCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTC
 TCCAACAGCCTTCAAGAATCCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTTCCTCACAGAGAACC
 AGCTACTCAACGTGATTATACCTACGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAA
 CGTGCCACCGGTTGTAATTTAGGAACGCTACTTGATTATTCCCTATTCAATTGTTCTTTCACAGATACATTT
 GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
 CTTTTTGGGGTCTCTTCAATAAGCCCTTCAAGTGCGTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
 CATGAC

B

GTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGATACGTTCTACCATGAGGACAAAT
 ATCCTTCTGAGGGGCTACAGTCATACCAACCTATTCTCCGCAATCCCCTACATCGGCCAAAACACTAGTAG
 AATGAGCCTGAGGTGGATTCTCTGTGACAAACCTACATTAACACGATTCTTCGCCCTGCACTTCTTACTT
 CCATTCGTTCATCGCAGGCATTACATTAGTTCACCTCACCTTCTTACACGAAAACAGGCTCAAACAACCCCT
 AGGAATCCCCCTCAGACTGTGACAAAATCCATTCACCCCGTACTACTCCACAAAAGACATCCTAGGGTTCG
 CTATCATGCTTATTCTCCTAGTCACTCTAGCACTATTCTCCCCAAACCTACTAGGAGACCCAGAAAACCTC
 ACACCAGCTAACCCCTAGTACACCCCCACACATCAAACCCGAATGATATTTCCCTATTTCGCATACGCTAT
 TCTCCGATCCATCCCAAACAACACTAGGAGGAGTACTAGCCCTTGCCGCCTCCGTCTTAGTCTTATTCTTA
 TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
 TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
 TGGGGTGGTTGGTTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGGGTTCCTTTTCG
 TGTCCTTGGTGTGGTTATAGTTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTATTCTGTCTCTTTGGCG
 GCGGATCCTTATCCGGAGTCGTGGGGAGATTTAGGGGTGTTGGGTATGGGTAGGGATGGGGTTGGTTGT
 TGCTGTGGGGTGGTAATTGGAGGAGGGGTTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
 TGCTGTGGCTCGTACGGATTTTGGTGGGGTAGCAGTTTTTTTATTTCGTGGGGGGCGGGCTGTTTTGATT
 GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCAGGGGTTATCAATCCTACCTCC
 AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAAATCCTCTCCTCGGATACG
 GGTGAATGTCCAGCACACCTTTGCATTCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTC
 TCCAACAGCCTTCAAGAATCCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTTCCTCACAGAGAACC
 AGCTACTCAACGTGATTATACCTACGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAA
 CGTGCCACCGGTTGTAATTTAGGAACGCTACTTGATTATTCCCTATTCAATTGTTCTTTCACAGATACATTT
 GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
 CTTTTTGGGGTCTCTTCAATAAACCCTTCAAGTGCGTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
 CATGAC

C

GTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGATACGTTCTACCATGAGGACAAAT
ATCCTTCTGAGGGGTACAGTCATACCAACCTATTCTCCGCAATCCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAAACCTACATTAACACGATTCCTCGCCCTGCACCTCTTACTT
CCATTCGTATCGCAGGCATTACATTAGTTCACCTCACCTTCTACACGAAACAGGCTCAAACAACCCCCCT
AGGAATCCCCCTCAGACTGTGACAAAATCCATTCACCCCGTACTACTCCACAAAAGACATCCTAGGGTTCG
CTATCATGCTTATTCTCCTAGTCACTCTAGCACTATTCTCCCCAAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCCCTAGTACACCCCCACACATCAAACCCGAATGATATTTCCCTATTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACACTAGGAGGAGTACTAGCCCTTGCCGCCTCCGTCTTAGTCTTATTCTCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGGGTGGTTGGGTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGGGTTCTTTCCG
TGTCCTTGGTGTGGTTATAGTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTATTCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCGTGGGGAGATTTAGGGGTGTTGGGTATGGGTAGGGATGGGGTGGTGT
TGCTGTGGGGTGGTAATTGGAGGAGGGGTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGCTGTGGCTCGTACGGATTTGGTGGGGTAGCAGTTTTTTTATTTCGTGGGGGGCGGGCTGTTTTTGATT
GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCAGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACG
GGTGAATGTCCAGCACACCTTTGCATTCCAAAGTCATAAAGTACTAGCCACCTCCTAGGATTTATTACTC
TCCAACAGCCTTCAAGAATCCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTTCCTCACGAGAACCG
AGCTACTCAACGTCAGTTATACTTACGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTAGGAACGCTACTTGATTATTCCCTATTCAATTTGTTCTTACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCCCTTCAAGTGCGTAGCAGGAGTTATCTTCCCTCTTGACATGTCCATCA
CATGAC

D

GTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGGTACGTTCTACCATGAGGACAAAT
ATCCTTCTGAGGGGTACAGTCATACCAACCTATTCTCCGCAATCCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAAACCTACATTAACACGATTCCTCGCCCTGCACCTCTTACTT
CCATTCGTATCGCAGGCATTACATTAGTTCACCTCACCTTCTACACGAAACAGGCTCAAACAACCCCCCT
AGGAATCCCCCTCAGACTGTGACAAAATCCATTCACCCCGTACTACTCCACAAAAGACATCCTAGGGTTCG
CTATCATGCTTATTCTCCTAGTCACTCTAGCACTATTCTCCCCAAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCCCTAGTACACCCCCACACATCAAACCCGAATGATATTTCCCTATTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACACTAGGAGGAGTACTAGCCCTTGCCGCCTCCGTCTTAGTCTTATTCTCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGAGTGGTTGGGTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGGGTTCTTTCCG
TGTCCTTGGTGTGGTTATAGTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTATTCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCGTGGGGGATTTAGGGGTGTTGGGTATGGGTAGGGATGGGGTGGTGT
TGCTGTAGGGGTGGTAATTGGAGGAGGGGTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGCTGTGGCTCGTACGGATTTGGTGGGGTAGCAGTTTTTTTATTTCGTGGGGGGCGGGCTGTTTTTGATT
GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCAGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACG
GGTGAATGTCCAGCACACCTTTGCATTCCAAAGTCATAAAGTACTAGCCACCTCCTAGGATTTATTACTC
TCCAACAGCCTTCAAGAATCCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTTCCTCACGAGAACCG
AGCTACTCAACGTCAGTTATACTTACGGTTATTGGTGTGAGGGGCATGATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTAGGAACGCTACTTGATTATTCCCTATTCAATTTGTTCTTACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
CTTTTTGGGGTCTCTTCAATAAGCCCTTCAAGTGCGTAGCAGGAGTTATCTTCCCTCTTGACATGTCCATCA
CATGAC

E

GTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGATACGTTCTACCATGAGGACAAAT
ATCCTTCTGAGGGGCTACAGTCATCACCAACCTATTCTCCGCAATCCCCACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAAACCCTACATTAACACGATTCTTCGCCCTGCACTTCTTACTT
CCATTTCGTATCGCAGGCATTACATTAGTTCACCTCACCTTCTTACACGAAACAGGCTCAAACAACCCCT
AGGAATCCCCTCAGACTGCGACAAAATCCATTCCACCCGTACTACTCCACAAAAGACATCCTAGGGTTCCG
CTATCATGCTTATTCTCCTAGTCACTCTAGCACATTTCTCCCCAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCTAGTCAACCCCCACACATCAAACCCGAATGATATTTCCCTATTTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACTGGGAGGAGTACTAGCCCTTCCGCCCTCCGTCTAGTCTTATTCTCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGGGTGGTTGGGTTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGGGTTTCTTTCCG
TGTCTTGGTGTGGTTATAGTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTTATTCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCGTGGGGGAGATTTAGGGGTTGTTGGGTATGGGTTAGGGATGGGGTTGGTTGT
TGCTGTGGGGGTGGTAATTGGAGGAGGGGTTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGTCTGTGGCTCGTACGGATTTTGGTGGGGTAGCAGTTTTTTTATTTCGTGGGGGGCGGGGCTGTTTTTATT
GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCGGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACG
GGTGAATGTCCAGCACACCTTTGCATTTCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTC
TCCAACAGCCTTCAAGAACTCCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTCTTCTCACGAGAACCG
AGCTACTCAACGTCAGTTATACTTACGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTAGGAACGCTACTTGATTTATTCCTATTCAATTGTTCTTTCACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCCCTTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

F

GTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGGTACGTTCTACCATGAGGACAAAT
ATCCTTCTGAGGGGCTACAGTCATCACCAACCTATTCTCCGCAATCCCCACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAAACCCTACACTAACACGATTCTTCGCCCTGCACTTCTTACTT
CCATTTCGTATCGCAGGCATTACATTAGTTCACCTCACCTTCTTACACGAAACAGGCTCAAACAACCCCT
AGGAATCCCCTCAGACTGTGACAAAATCCATTCCACCCGTACTACTCCACAAAAGACATCCTAGGGTTCCG
CTATCATGCTTATTCTCCTAGTCACTCTAGCACATTTCTCCCCAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCTAGTCAACCCCCACACATCAAACCCGAATGATATTTCCCTATTTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACTAGGAGGAGTACTAGCCCTTCCGCCCTCCGTCTAGTCTTATTCTCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGGGTGGTTGGGTTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGAGTTTCTTTCCG
TGTCTTGGTGTGGTTATAGTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTTATTCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCGTGGGGGAGATTTAGGGGTTGTTGGGTATGGGTTAGGGATGGGGTTGGTTGT
TGCTGTGGGGGTGGTAATTGGAGGAGGGGTTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGTCTGTGGCTCGTACGGATTTTGGTGGGGTAGCAGTTTTTTTATTTCGTGGGGGGCGGGGCTGTTTTTATT
GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCGGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACG
GGTGAATGTCCAGCACACCTTTGCATTTCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTC
TCCAACAGCCTTCAAGAACTCCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTCTTCTCACGAGAACCG
AGCTACTCAACGTCAGTTATACTTACGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTAGGAACGCTACTTGATTTATTCCTACTCAATTGTTCTTTCACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCCCTTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

G

GTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGATACGTTCTACCATGAGGACAAAT

ATCCTTCTGAGGGGCTACAGTCATCACCAACCTATTCTCCGCAATCCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAAACCTACATTAACACGATTCTTCGCCCTGCACTTCTTACTT
CCATTCGTCATCGCAGGCATTACATTAGTTACCTCACCTTCCCTACACGAAACAGGCTCAAACAACCCCT
AGGAATCCCCTCAGACTGTGACAAAATTCCATTCACCCCGTACTACTCCACAAAAGACATCCTAGGGTTCCG
CTATCATGCTTATTCTCCTAGTCACTCTAGCACTATTCTCCCCAAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCTAGTACACCCCCACACATCAAACCCGAATGATATTTCCCTATTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACCTAGGAGGAGTACTAGCCCTTGCCGCCTCCGTCTAGTCTTATTCCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGGGTGGTTGGTTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGGGTTTCTTTTCG
TGTCCTTGGTGTGGTTATAGTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTATTCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCGTGGGGGATTTGGGGGTTGTTGGGTATGGGTTAGGGATGGGGTTGGTTGT
TGCTGTGGGGGTGGTAATTGGAGGAGGGGTTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGCTGTGGCTCGTACGGATTTGGTGGGGTAGCAGTTTTTTTATTCTGTGGGGGGCGGGCTGTTTTGATT
GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCGGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAAATCCTCTCCTCGGATACG
GGTGAATGTCCAGCACACCTTTGCATTCCAAAGTCATGAACCTAGCCACCTCCTAGGATTTATTACTC
TCCAACAGCCTTCAAGAACCTCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTCTTACAGAGAACCG
AGCTACTCAACGTCAGTTATACTTACGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTAGGAACGCTACTTGATTATTCCTATTAATGTTCTTACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCCCTTCAAGTGCGTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

H

GTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGGTACGTTCTACCATGAGGACAAA
ATCCTTCTGAGGGGCTACAGTCATCACCAACCTATTCTCCGCAATCCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAAACCTACATTAACACGATTCTTCGCCCTGCACTTCTTACTT
CCATTCGTCATCGCAGGCATTACATTAGTTACCTCACCTTCCCTACACGAAACAGGCTCAAACAACCCCT
AGGAATCCCCTCAGACTGTGACAAAATTCCATTCACCCCGTACTACTCCACAAAAGACATCCTAGGGTTCCG
CTATCATGCTTATTCTCCTAGTCACTCTAGCACTATTCTCCCCAAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCTAGTACACCCCCACACATCAAACCCGAATGATATTTCCCTATTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACCTAGGAGGAGTACTAGCCCTTGCCGCCTCCGTCTAGTCTTATTCCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGAGTGGTTGGTTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGGGTTTCTTTTCG
TGTCCTTGGTGTGGTTATAGTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTATTCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCGTGGGGGATTTAGGGGTTGTTGGGTATGGGTTAGGGATGGGGTTGGTTGT
TGCTGTAGGGGTGGTAATTGGAGGAGGGGTTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGCTGTGGCTCGTACGGATTTGGTGGGGTAGCAGTTTTTTTATTCTGTGGGGGGCGGGCTGTTTTGATT
GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCGGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAAATCCTCTCCTCGGATACG
GGTGAATGTCCAGCACACCTTTGCATTCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTC
TCCAACAGCCTTCAAGAACCTCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTCTTACAGAGAACCG
AGCTACTCAACGTCAGTTATACTTACGGTTATTGGTGTGAGGGGCATGCATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTAGGAACGCTACTTGATTATTCCTATTAATGTTCTTACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCCCTTCAAGTGCGTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

I

GTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGGTACGTTCTACCATGAGGACAAA
ATCCTTCTGAGGGGCTACAGTCATCACCAACCTATTCTCCGCAATCCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAAACCTACATTAACACGATTCTTCGCCCTGCACTTCTTACTT

CCATTCGTCATCGCAGGCATTACATTAGTTCACCTCACCTTCCTACACGAAACAGGCTCAAACAACCCCT
AGGAATCCCCCTCAGACTGTGACAAAATTCATTCACCCCGTACTACTCCACAAAAGACATCCTAGGGTTTCG
CTATCATGCTTATTCTCCTAGTCACTCTAGCACTATTCTCCCCAAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCTAGTACACCCCCACACATCAAACCCGAATGATATTTCTATTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACCTAGGAGGAGTACTAGCCCTTGCCGCCTCCGTCTAGTCTATTCCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGGGTGGTTGGGTTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGAGTTTCTTTCG
TGTCTTGGTGTGGTTATAGTTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTATCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCTGTTGGGGGATTTAGGGGTTGTTGGGTATGGGTAGGGATGGGGTTGGTTGT
TGCTGTGGGGTGGTAATTGGAGGAGGGTTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGTCTGTGGCTCGTACGGATTTTGGTGGGGTAGCAGTTTTTTATTCGTGGGGGGCGGGCTGTTTTGATT
GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCAGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAAACCTCTCCTCGGATACG
GGTGAATGTCCAGCACACCTTTGCATTCCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTC
TCCAACAGCCTTCAAGAACCTCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTCTTCTCACGAGAACCG
AGCTACTCAACGTCAGTTATACTTACGGTTATTTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCAA
CGTGCCACCGGTTGTAATTTCAAGAACGCTACTTGATTATTCCTACTCAATTGTTCTTTCACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTACTCTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCCCTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

J

GTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGGTACGTTCTACCATGAGGACAAAAT
ATCCTTCTGAGGGGCTACAGTCATCACCACCTATTCTCCGCAATCCCCTACATCGGCCAAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAAACCTACATTAACACGATTCTTCGCCCTGCACTTCTTACTT
CCATTCGTCATCGCAGGCATTACATTAGTTCACCTCACCTTCCTACACGAAACAGGCTCAAACAACCCCT
AGGAATCCCCCTCAGACTGTGACAAAATTCATTCACCCGTACTACTCCACAAAAGACATCCTAGGGTTTCG
CTATCATGCTTATTCTCCTAGTCACTCTAGCACTATTCTCCCCAAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCTAGTACACCCCCACACATCAAACCCGAATGATATTTCTATTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACCTAGGAGGAGTACTAGCCCTTGCCGCCTCCGTCTAGTCTTATTCTTAA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGAGTGGTTGGGTTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGGGTTTCTTTCG
TGTCTTGGTGTGGTTATAGTTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTATCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCTGTTGGGGGATTTAGGGGTTGTTGGGTATGGGTAGGGATGGGGTTGGTTGT
TGCTGTAGGGTGGTAATTGGAGGAGGGTTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGTCTGTGGCTCGTACGGATTTTGGTGGGGTAGCAGTTTTTTATTCGTGGGGGGCGGGCTGTTTTGATT
GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCAGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACCAATCCTCTCCTCGGATACG
GGTGAATGTCCAGCACACCTTTGCATTCCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTC
TCCAACAGCCTTCAAGAACCTCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTCTTCTCACGAGAACCG
AGCTACTCAACGTCAGTTATACTTACGGTTATTTGGTGTGAGGGGCATGCATCTAATAAACTTGATCTCAA
CGTGCCACCGGTTGTAATTTCAAGAACGCTACTTGATTATTCCTATTCATTTGTTCTTTCACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTACTCTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCCCTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

K

GTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGGTACGTTCTACCATGAGGACAAAAT
ATCCTTCTGAGGGGCTACAGTCATCACCACCTATTCTCCGCAATCCCCTACATCGGCCAAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAAACCTACATTAACACGATTCTTCGCCCTGCACTTCTTACTT
CCATTCGTCATCGCAGGCATTACATTAGTTCACCTCACCTTCCTACACGAAACAGGCTCAAACAACCCCT
AGGAATCCCCCTCAGACTGTGACAAAATTCATTCACCCGTACTACTCCACAAAAGACATCCTAGGGTTTCG

CTATCATGCTTATTCTCCTAGTCACTCTAGCACTATTCTCCCCAAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCTAGCCACACCCCCACACATCAAACCCGAATGATATTTCCCTATTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACTAGGAGGAGTACTAGCCCTGCCGCTCCGTCTAGTCTTATTCCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGAGTGGTTGGGTGGTTTTGGCATCTGTTGCTGGGTGGTGGTTGATTGGTAAGTTTAGGGGTTCTTTTCG
TGTCCTTGGTGTGGTTATAGTTTATTTAGGTGGATGTTGGTGGTGTGTTGTTTATTTCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCGTGGGGGATTTAGGGGTTGTTGGGTATGGGTAGGGATGGGGTTGGTTGT
TGCTGTAGGGTGGTAATTTGGAGGAGGGTTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGCTGTGGCTCGTACGGATTTTGGTGGGGTAGCAGTTTTTTTATTCGTGGGGGGCGGGCTGTTTTTGATT
GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCAGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAAGATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACG
GGTGAATGTCCAGCACACCTTTGCATTTCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTAATC
TCCAACAGCCTTCAAGAATCCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTCTTCTCAGAGAACCG
AGCTACTCAACGTCAGTTATACCTTACGGTTATTGGTGTGAGGGGCATGCATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTTCAAGAACGCTACTTGTATTATTCCTATTCAAATGTTCTTTCACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCCCTTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

L

GTAGGAGTTATTCTCCTACTAACCCCTCATAGCAACTGCTTTCGTAGGATACGTTCTACCATGAGGACAAAT
ATCCTTCTGAGGGGTACAGTCATACCAACCTATTCTCCGCAATCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAAACCTACATTAACACGATTCTTCGCCCTGCACTTCTTACTT
CCATTCGTATCGCAGGCATTACATTAGTTCACCTCACCTTCCCTACACGAAACAGGGCTCAAACAACCCCT
AGGAATCCCTCAGACTGTGACAAAATTCATTTCCACCCGTACTACTCCACAAAAGACATCCTAGGGTTTCG
CTATCATGCTTATTCTCCTAGTCACTCTAGCACTATTCTCCCCAAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCTAGTCACACCCCCACACATCAAACCCGAATGATATTTCCCTATTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACTAGGAGGAGTACTAGCCCTGCCGCTCCGTCTAGTCTTATTCCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
TTTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGGGTGGTTGGGTGGTTTTGGCATCTGTTGCTGGGTGGTGGTTGATTGGTAAGTTTAGGGGTTCTTTTCG
TGTCCTTGGTGTGGTTATAGTTTATTTAGGTGGATGTTGGTGGTGTGTTGTTTATTTCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCGTGGGGGATTTAGGGGTTGTTGGGTATGGGTAGGGATGGGGTTGGTTGT
TGCTGTGGGGTGGTAATTTGGAGGAGGGTTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGCTGTGGCTCGTACGGATTTTGGTGGGGTAGCAGTTTTTTTATTCGTGGGGGGCGGGCTGTTTTTGATT
GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCAGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAAGATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACG
GGTGAATGTCCAGCACACCTTTGCATTTCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTAATC
TCCAACAGCCTTCAAGAATCCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTCTTCTCAGAGAACCG
AGCTACTCAACGTCAGTTATACCTTACGGTTATTGGTGTGAGGGGCATGCATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTTCAAGAACGCTACTTGTATTATTCCTATTCAAATGCTCTTTCACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCCCTTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

M

GTAGGAGTTATTCTCCTACTAACCCCTCATAGCAACTGCTTTCGTAGGATACGTTCTACCATGAGGACAAAT
ATCCTTCTGAGGGGTACAGTCATACCAACCTATTCTCCGCAATCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAAACCTACATTAACACGATTCTTCGCCCTGCACTTCTTACTT
CCATTCGTATCGCAGGCATTACATTAGTTCACCTCACCTTCCCTACACGAAACAGGGCTCAAACAACCCCT
AGGAATCCCTCAGACTGTGACAAAATTCATTTCCACCCGTACTACTCCACAAAAGACATCCTAGGGTTTCG
CTATCATGCTTATTCTCCTAGTCACTCTAGCACTATTCTCCCCAAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCTAGTCACACCCCCACACATCAAACCCGAATGATATTTCCCTATTCGCATACGCTAT

TCTCCGATCCATCCCAAACAACCTAGGAGGAGTACTAGCCCTTGCCGCTCCGTCCCTAGTCTTATTCCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGGGTGGTTGGGTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGGGTTCTTTTCG
TGTCTTGGTGTGGTTATAGTTTATTTAGGTGGGATGTTGGTGGTGTGTTTTATTTCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCGTGGGAGATTTAGGGGTTGTTGGGTATGGGTAGGGATGGGGTTGGTTGT
TGCTGTGGGGTGGTAATTGGAGGAGGGGTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGTCTGTGGCTCGTACGGATTTTGGTGGGTAGCAGTTTTTTATTCGTGGGGGGCGGGCTGTTTTGATT
GGTGGTTGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACG
GGTGAATGTCCCAGCACACCTTTGCATTCCCAAAGTCATAAACTTAGCCACCTCTAGGATTTATTACTC
TCCAACAGCCTTCAAGAACTCCCAAGCCAGAGAACCCTGGTTATCTATTAATCGTGTCTTCTCAGGAAACC
AGCTACTCAACGTGATTATACTTACGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTAGGAACACTACTTGGATTATTCCTATTCAATTGTTCTTCCAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCCCTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

N

GTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGGTACGTTCTACCATGAGGACAAAT
ATCCTTCTGAGGAGCTACAGTCATACCAACCTATTTCTCCGCAATCCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAACCCCTACATTAACACGATTCCTTCGCCCTGCACCTCTTACTT
CCATTCGTGATCGCAGGCATTACATTAGTTACCTCACCTTCCCTACACGAAACAGGCCTCAAACAACCCCT
AGGAATCCCCTCAGACTGTGACAAAATTCATTCACCCCGTACTACTCCACAAAAGACATCCTAGGGTTCCG
CTATCATGCTTATTCTCCTAGTCACTCTAGCACTATTTCTCCCAAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCTAGTCACACCCCCACACATCAAACCCGAATGATATTTCCCTATTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACCTAGGAGGAGTACTAGCCCTTGCCGCTTCCGTCCCTAGTCTTATTCCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGGGTGGTTGGGTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGGGTTCTTTTCG
TGTCTTGGTGTGGTTATAGTTTATTTAGGTGGGATGTTGGTGGTGTGTTTTATTCGTGTTCTTTGGCG
GCGGATCCTTATCCGGAGTCGTGGGGGATTTAGGGGTTGTTGGGTATGGGTAGGGATGGGGTTGGTTGT
TGCTGTGGGGTGGTAATTGGAGGAGGGGTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGTCTGTGGCTCGTACGGATTTTGGTGGGTAGCAGTTTTTTATTCGTGGGGGGCGGGCTGTTTTGATT
GGCGGTTGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACG
GGTGAATGTCCCAGCACACCTTTGCATTCCCAAAGTCATAAACTTAGCCACCTCTAGGATTTATTACTC
TCCAACAGCCTTCAAGAACTCCCAAGCCAGAGAACCCTGGTTATCTATTAATCGTGTCTTCTCAGGAAACC
AGCTACTCAACGTGATTATACTTACGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTAGGAACGCTACTTGGATTATTCCTATTCAATTGTTCTTCCAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCCCTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

O

GTAGGAGTTATTCTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGGTACGTTCTACCATGAGGACAAAT
ATCCTTCTGAGGGGCTACAGTCATACCAACCTATTTCTCCGCAATCCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAACCCCTACATTAACACGATTCCTTCGCCCTGCACCTCTTACTT
CCATTCGTGATCGCAGGCATTACATTAGTTACCTCACCTTCCCTACACGAAACAGGCCTCAAACAACCCCT
AGGAATCCCCTCAGACTGTGACAAAATTCATTCACCCCGTACTACTCCACAAAAGACATCCTAGGGTTCCG
CTATCATGCTTATTCTCCTAGTCACTCTAGCACTATTTCTCCCAAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCTAGTCACACCCCCACACATCAAACCCGAATGATATTTCCCTATTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACCTAGGAGGAGTACTAGCCCTTGCCGCTTCCGTCCCTAGTCTTATTCCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA

TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGGGTGGTTGGGTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGGTTCCTTTTCG
TGTCCTTGGTGTGGTTATAGTTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTATTCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCGTGGGGGGATTTAGGGGTTGTTGGGTATGGGTTAGGGATGGGGTTGGTTGT
TGCTGTGGGGTGGTAATTGGAGGAGGGGTTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGCTGTGGCTCGTACGGATTTTGGTGGGGTAGCAGTTTTTTATTCGTGGGGGGCGGGCTGTTTTTGATT
GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCAGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACG
GGTGAATGTCCAGCACACCTTTGCATTTCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTC
TCCAACAGCCTTCAAGAACCTCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTTCCTCAGGAGAACC
AGCTACTCAACGTGAGTTATACTTACGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTTCAAGAACGCTACTTGAATTTTCTACTCAATTTGTTCTTTCACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCTTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

P

GTAGGAGTTATTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGATACGTTCTACCATGAGGACAAAT
ATCCTTCTGAGGGGCTACAGTCATACCAACCTATTCCTCCGCAATCCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAACCTTACATTAACACGATTCTTCGCCCTGCACTTCTTACTT
CCATTCGTTCATCGCAGGCATTACATTTAGTTCACCTCACCTTCTTACAGAAACAGGCTCAAACAACCCCT
AGGAATCCCCTCAGACTGTGACAAAATTCATTTCCACCCGTACTACTCCACAAAAGACATCCTAGGGTTTCG
CTATCATGCTTATTCCTTAGTCACTCTAGCACTATTCCTCCCAAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCTAGTCACACCCCCACACATCAAACCCGAATGATATTTCTTATTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACCTAGGAGGAGTACTAGCCCTTGCCGCCTCCATCCTAGTCTTATTCCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATATGATAGAA
TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGGGTGGTTGGGTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGGTTCCTTTTCG
TGTCCTTGGTGTGGTTATAGTTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTATTCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCGTGGGGAGATTTAGGGGTTGTTGGGTATGGGTAGGGATGGGGTTGGTTGT
TGCTGTGGGGTGGTAATTGGAGGAGGGGTTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGCTGTGGCTCGTACGGATTTTGGTGGGGTAGCAGTTTTTTTATTTCGTGGGGGGCGGGCTGTTTTTGATT
GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGTGGAGTTGGTGCAGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACG
GGTGAATGTCCAGCACACCTTTGCATTTCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTC
TCCAACAGCCTTCAAGAACCTCCAAGCCAGAGAACCTGGTTATCTATTAATCGTGTTCCTCAGGAGAACC
AGCTACTCAACGTGAGTTATACTTACGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTTCAAGAACGCTACTTGAATTTTCTTATTCATTTGTTCTTTCACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCTTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

Q

GTAGGAGTTATTCCTACTAACCCTCATAGCAACTGCTTTCGTAGGGTACGTTCTACCATGAGGACAAAT
ATCCTTCTGAGGGGCTACAGTCATACCAACCTACTCTCCGCAATCCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGTGGATTCTCTGTGACAACCTTACATTAACACGATTCTTCGCCCTGCACTTCTTACTT
CCATTCGTTCATCGCAGGCATTACATTTAGTTCACCTCACCTTCTTACAGAAACAGGCTCAAACAACCCCT
AGGAATCCCCTCAGACTGTGACAAAATTCATTTCCACCCGTACTACTCCACAAAAGACATCCTAGGGTTTCG
CTATCATGCTTATTCCTTAGTCACTCTAGCACTATTCCTCCCAAACCTACTAGGAGACCCAGAAAACCTC
ACACCAGCTAACCCCTAGTCACACCCCCACACATCAAACCCGAATGATATTTCTTATTCGCATACGCTAT
TCTCCGATCCATCCCAAACAACCTAGGAGGAGTACTAGCCCTTGCCGCCTCCGTCCTAGTCTTATTCCTTA
TACCCTTACTACACACATCCAAACAACGCTCAATAACCTTCCGCCACTCTCAGTGGGAATATGATAGAA
TTTGTATTATTTTTAGGGCTGTGCTTTGTTTTGGGGGCTTTGGGAGTGGCGTCTAATCCTTCTCCTTATTA
TGGAGTGGTTGGGTGGTTTTGGCATCTGTTGCTGGGTGTGGTTGATTGGTAAGTTTAGGGTTCCTTTTCG

TGTCCTTGGTGTGGTTATAGTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTATTCTGTCTCTTTGGCG
GCGGATCCTTATCCGGAGTCGTGGGGGGATTTAGGGGTTGTTGGGTATGGGTTAGGGATGGGGTTGGTTGT
TGCTGTAGGGGTGGTAATTGGAGGAGGGGTTGAGTGTGTGGTGAATAGTGGTACAGTGAATGGTGAAGGGT
TGTCTGTGGCTCGTACGGATTTTGGTGGGGTAGCAGTTTTTTTATTTCGTGGGGGGCGGGGCTGTTTTTATT
GGTGGTTGGGGATTGTTGTTAACTTTATTTGTAGTGCCTGGAGTTGGTGCGGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTATATCTCCCTGCGCATACAAATCCTCTCCTCGGATACG
GGTGAATGTCCCAGCACACCTTTGCATTTCCCAAAGTCATAAACTTAGCCACCTCCTAGGATTTATTACTC
TCCAACAGCCTTCAAGAACTCCCAAGCCAGAGAACCCTGGTTATCTATTAATCGTGTCTTACAGAGAACCG
AGCTACTCAACGTCAGTTATACTTACGGTTATTGGTGTGAGGGGCATGCATCTAATAAACTTGATCTCCAA
GCTGCCACCGGTTGTAATTTACAGGAACGCTACTTGATTATTCCTATTCAATTGTTCTTACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTCTTTT
CTTTTTGGGGTCTCTTCAATAAACCCCTTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

R

GTAGGAGTCATTCTATTACTAACCCTTATAGCAACTGCTTTTCGTAGGATATGTCCTACCATGAGGACAAAT
ATCCTTTTGGGGGCTACAGTCATTACCAACCTATTTTCCGCAATCCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGCGGATTTTCTGTAGACAACCCACGTTAACACGATTCTTCGCCCTTCACTTCTACTT
CCATTCGTATCGCAGGCATTACACTAGTCCACCTCACCTTCCACACGAAACAGGCTCAAACAACCCACT
AGGAATTCCTCAGACTGTGACAAAATCCATTCACCCATACTACTCCACAAAAGATATCCTAGGATTCG
CCATCATACTAATTCTCCTAGTTACCCCTGGCCTATTTCTCCCTAACCTGCTAGGAGACCCAGAAAACCTC
ACACCAGCCAACCCACTAGTCACACCCCCACACATTAACCTGAATGATACCTTCTATTTGCATACGCTAT
TCTCCGATCGATCCCAAACAACTGGGAGGAGTATTAGCCCTTGCCGCCTCTGTCCTAGTCTTCTCTCA
TACCTCTACTACACACATCCAAAACAGCGCTCAATAACCTTCGCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTAGGCCGTGTGTTTTGTTTTGGGGGCTTTGGGGGTAGCGTCTAATCCTTCCCCTTATTA
TGGGGTGGTTGGGTTGGTTTGGCTTCTGTTGCTGGGTGTTGGTTGACTGGTAAGTTAGGGGTTCTTTTCG
TGTCTCTGGTGTGGTAATAGTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTATTCTGTTTCTCTGGCG
GCGGACCCCTATCCGGAATCGTGAGGGAGCTTAGGAGTTGTTGGGTATGGATTGGGATAGGGCTGGTTAT
TATTGTGGGGTAGTTGTTGGAGGGGGGTTGAGTTTGTGGTGAATAGTGGTACGGTGAATGGGTGAGGGGT
TGTCTGTGGCTCGTACGGATTTTACTGGGGTAGCAGTTTTTTTATTTCGTGGGGGGCGGGGCTGTTTTTAATT
GGTGGTTGAGGGTTGTTGTTGACTTTGTTTGTGGTGTGGAGTTGGTGCGGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTACATCTCCCTGCACATACAAACCCCTCCTCCTCGGATACG
GGTGAATGTCCCAGCACACCTTTGAATTTCCCAAGGTCACGAACCTTAGCCACCTCCTAGGAATTTATACTC
TCCAACAGCTTTCAAGGACTCCCAAGCCAGAGAACCCTGGTTATCTATTAATCGTGCCTCTCACGAGAACCG
AGCTACTCAACGTCAGTTATACTTTCGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTACAGGAATACTGATTGATTGTTTCCCTACTCAATTGCTCTTACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTTTTTT
CTTTTTGGGGTCTCTTCAATAAGCCCTTCAAGTGCCTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

S

GTAGGAGTCATTCTATTACTAACCCTTATAGCAACTGCTTTTCGTAGGATATGTCCTGCCATGAGGACAAAT
ATCCTTTTGGGGGCTACAGTCATTACCAACCTATTTTCCGCAATCCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGCGGATTTTCTGTAGACAACCCACGTTAACACGATTCTTCGCCCTTCACTTCTACTT
CCATTCGTATCGCAGGCATTACACTAGTCCACCTCACCTTCCACACGAAACAGGCTCAAACAACCCACT
AGGAATTCCTCAGACTGTGACAAAATCCATTCACCCATACTACTCCACAAAAGATATCCTAGGATTCG
CCATCATACTAATTCTCCTAGTTACCCCTGGCCTATTTCTCCCTAACCTGCTAGGAGACCCAGAAAACCTC
ACACCAGCCAACCCACTAGTCACACCCCCACACATTAACCTGAATGATACCTTCTATTTGCATACGCTAT
TCTCCGATCGATCCCAAACAACTGGGAGGAGTATTAGCCCTTGCCGCCTCTGTCCTAGTCTTCTCTCA
TGCTCTACTACACACATCCAAAACAGCGCTCAATAACCTTCGCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTAGGCCGTGTGTTTTGTTTTGGGGGCTTTGGGGGTAGCGTCTAATCCTTCCCCTTATTA
TGGGGTGGTTGGGTTGGTTTGGCTTCTGTTGCTGGGTGTTGGTTGACTGGTAAGTTAGGGGTTCTTTTCG
TGTCTCTGGTGTGGTAATAGTTTATTTAGGTGGGATGTTGGTGGTGTGTTGTTTATTCTGTTTCTCTGGCG
GCGGACCCCTATCCGGAATCGTGAGGAACTTAGGAGTTGTTGGGTATGGATTGGGGATAGGGCTGGTTAT

TATTGTGGGGTAGTTGTTGGAGGAGGGTTGAGTTTGTGGTGAATAGTGGTACGGTGAATGGTGAGGGGT
TGTCTGTGGCTCGTACGGATTTTAGTGGGGTAGCAGTTTTTTTATTTCGTGGGGGGCGGGGCTGTTTTTAATT
GGTGGTTGAGGGTTATTGTTGACTTTGTTTGTGGTGTGGAGTTGGTACGGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTACATCTCCCTGCACATACAAACCCTCTCCTCGGATACG
GGTGAATGTCCCAGCACACCTTTGAATTTCCAAGGTCACGAACTTAGCCACCTCCTAGGGATTTATACTC
TCCAACAGCTTTCAAGGACTCCCAAGCCAGAGAACCCTGGTTATCTATTAATCGTGCCTCTCACGAGAACCG
AGCTACTCAACGTCAGTTATACTTTTCGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTTCAGGAATACTGATTGATTGTTTCCTACTCAATTGCTCTTCACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTTTTTT
CTTTTTGGGGTCTCTTCAATAAGCCCTTCAAGTGGTGTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

T

GTAGGAGTCATTCTATTACTAACCCTTATAGCAACTGCTTTCGTAGGATATGTCCTGCCATGAGGACAAAT
ATCCTTTTGGAGGGCTACAGTCATTACCAACCTATTTTCCGCAATCCCCTACATCGGCCAAACACTAGTAG
AATGAGCCTGAGGCGGATTTTCTGTAGACAACCCACGTTAACACGATTCTTCGCCCTTCACTTCTACTT
CCATTGTCATCGCAGGCATTACACTAGTCCACCTCACCTTCTACACGAAACAGGCTCAAACAACCCACT
AGGAATTCCTCAGACTGTGACAAAATTCATTCCACCCATACTACTCCACAAAAGATATCCTAGGATTTCG
CCATCATACTAATTCTCCTAGTTACCCTGGCACTATTTCTCCCCTAACCTGCTAGGAGACCCAGAAAACCTC
ACACCAGCCAACCCACTAGTCACACCCCCACACATTAACCTGAATGATACTTCTTATTTGCATACGCTAT
TCTCCGATCGATCCCAAACAAACTGGGAGGAGTATTAGCCCTTGCCGCCTCTGTCCTAGTCTTATCCTCA
TACCTCTACTACACACATCCAAACAGCGCTCAATAACCTTCCGCCACTCTCAGTGGGAATTATGATAGAA
TTTGTATTATTTTTAGGCCCTGTGTTTTGTTTTGGGGGCTTTGGGGGTAGCGTCTAATCCTTCCCCTTATTA
TGGGGTGGTTGGGTTGGTTTTGGCTTCTGTTGCTGGGTGTGGTTGACTGGTAAGTTTAGGGGTTTTCTTTTCG
TGTCTCTGGTGCTGGTAATAGTTTATTAGGTGGGATGTTGGTGGTGTGTTGTTTATTCTGTTTCTCTGGCG
GCGGACCCTTATCCGGAATCGTGAGGGAACCTTAGGAGTTGTTGGGTATGGATTGGGGATAGGGCTGGTTAT
TATTGTGGGGTAGTTGTTGGAGGAGGGTTGAGTTTGTGGTGAATAGTGGTACGGTGAATGGTGAGGGGT
TGTCTGTGGCTCGTACGGATTTTAGTGGGGTAGCAGTTTTTTTATTTCGTGGGGGGCGGGGCTGTTTTTAATT
GGTGGTTGAGGGTTATTGTTGACTTTGTTTGTGGTGTGGAGTTGGTACGGGGGTTATCAATCCTACCTCC
AGCCGAAGCCGCAAGCGTTACTTAAGATCGACTACATCTCCCTGCACATACAAACCCTCTCCTCGGATACG
GGTGAATGTCCCAGCACACCTTTGAATTTCCAAGGTCACGAACTTAGCCACCTCCTAGGAATTTATACTC
TCCAACAGCTTTCAAGGACTCCCAAGCCAGAGAACCCTGGTTATCTATTAATCGTGCCTCTCACGAGAACCG
AGCTACTCAACGTCAGTTATACTTTTCGGTTATTGGTGTGAGGGGCATACATCTAATAAACTTGATCTCCAA
CGTGCCACCGGTTGTAATTTTCAGGAATACTGATTGATTGTTTCCTACTCAATTGCTCTTCACAGATACATTT
GCTTGGGGATGGACGCTCCTCATATATATCCGTAATCGCGGCTAGCCACAGGCCGTTTTTACTCTTTTTTT
CTTTTTGGGGTCTCTTCAATAAGCCCTTCAAGTGGTGTAGCAGGAGTTATCTTCTCTTGACATGTCCATCA
CATGAC

APPENDIX B
DISTANCE MATRICES

DISTANCE MATRICES

	1	2	3	4	5	6	7	8	9	10	11
1 Piute Mountain		0.0457	0.2209	0.1413	0.1367	0.1229	*0.2270	*0.1777	0.0917	*0.1882	*0.2023
2 Cerro Noroeste	102.1		0.0427	0.0000	0.0225	0.1185	*0.1905	*0.1429	0.0000	0.1104	*0.1493
3 Frazier Mountain	92.0	23.4		0.0375	*0.3095	0.1318	*0.2381	*0.1905	*0.1377	0.1864	*0.1630
4 Winston Peak	128.2	129.4	106.2		*0.1704	0.1312	*0.1667	0.1007	0.0616	0.1290	0.1250
5 Blue Ridge	138.2	151.6	128.3	23.8		*0.2138	*0.2619	*0.1809	0.0128	0.1227	0.1884
6 Delamar Mountain	185.2	218.7	195.2	92.4	68.7		*0.1905	0.0233	0.0233	0.0000	0.0000
7 Onyx Peak	206.0	240.5	217.2	113.5	89.9	22.0		*0.1064	*0.1250	*0.1951	0.1312
8 Grinnell Mountain	206.5	237.1	213.8	109.0	85.6	21.0	9.2		0.0148	0.0040	0.0000
9 Black Mountain	234.1	253.1	230.1	124.0	103.1	53.5	42.0	34.9		0.0000	0.0000
10 Thomas Mountain	256.2	270.2	247.6	142.1	122.9	77.8	65.3	58.9	24.4		0.0000
11 Laguna Mountain	339.5	338.5	317.3	217.0	202.0	165.6	152.3	146.6	112.2	87.9	

Values below the diagonal are pairwise geographic distances in kilometers for *Sitta pygmaea* collecting locations. Values above the diagonal are pairwise F_{ST} values for *Sitta pygmaea* collecting locations. * indicates significance at $p=0.05$. Negative values were changed to zero.

APPENDIX C
INPUT FILES

ARLEQUIN INPUT FILE

[Profile]

Title="Sitta pygmaea AMONG MOUNTAIN RANGES AMOVA"
NbSamples= 7
DataType= DNA
GenotypicData= 0
LocusSeparator= NONE
GameticPhase= 0
RecessiveData= 0
RecessiveAllele= null
MissingData= '?'

[Data]

[[HaplotypeDefinition]]

HaplListName="Sitta pygmaea Haplotype List"
HaplList={
[Insert 1639-base pair haplotype list without haplotypes R, S and T
here]
}

[[Samples]]

SampleName="PIUTE"
SampleSize=6
SampleData= {
A 2
B 3
C 1
}
SampleName="NOROESTE"
SampleSize=7
SampleData= {
A 3
C 2
D 1
E 1
}
SampleName="FRAZIER"
SampleSize=7
SampleData= {
C 3
D 3
E 1
}
SampleName="SAN GABRIEL"
SampleSize=14
SampleData= {

```

    A 5
    C 1
    D 1
    E 3
    F 1
    G 2
    H 1
}

SampleName="SAN BERNARDINO"
SampleSize=21
SampleData= {
    B 1
    C 2
    F 1
    H 5
    I 1
    J 3
    K 1
    L 3
    M 1
    N 2
    O 1
}

SampleName="SAN JACINTO"
SampleSize=13
SampleData= {
    A 3
    D 2
    H 5
    K 1
    P 1
    Q 1
}

SampleName="LAGUNA"
SampleSize=7
SampleData= {
    D 1
    H 3
    J 1
    N 1
    P 1
}

[[Structure]]

StructureName = "AMONG MOUNTAIN RANGES"
NbGroups=1
Group={
    "PIUTE"
}

```

```
"NOROESTE"  
"FRAZIER"  
"SAN GABRIEL"  
"SAN BERNARDINO"  
"SAN JACINTO"  
"LAGUNA"  
}
```

[Profile]

```
Title="Sitta pygmaea subspecies AMOVA"  
NbSamples= 2  
DataType= DNA  
GenotypicData= 0  
LocusSeparator= NONE  
GameticPhase= 0  
RecessiveData= 0  
RecessiveAllele= null  
MissingData= '?'
```

[Data]

[[HaplotypeDefinition]]

```
HaplListName="Sitta pygmaea Haplotype List"  
HaplList={  
[Insert 1639-base pair haplotype list without haplotypes R, S and T  
here]  
}
```

[[Samples]]

```
SampleName="melanotis"  
SampleSize=68  
SampleData= {  
A 13  
B 4  
C 9  
D 7  
E 5  
F 2  
G 2  
H 11  
I 1  
J 3  
K 2  
L 3  
M 1  
N 2  
O 1  
P 1  
Q 1  
}
```



```

SampleName="leuconucha"
SampleSize=7
SampleData= {
  D 1
  H 3
  J 1
  N 1
  P 1
}

[[Structure]]

StructureName = "between subspecies san jacinto=melanotis"
NbGroups=1
Group={
  "melanotis"
  "leuconucha"
}

```

```

[Profile]

Title="Sitta pygmaea subspecies AMOVA"
NbSamples= 2
DataType= DNA
GenotypicData= 0
LocusSeparator= NONE
GameticPhase= 0
RecessiveData= 0
RecessiveAllele= null
MissingData= '?'

[Data]

[[HaplotypeDefinition]]
HaplListName="Sitta pygmaea Haplotype List"
HaplList={
[Insert 1639-base pair haplotype list without haplotypes R, S and T
here]
}

[[Samples]]

SampleName="melanotis"
SampleSize=55
SampleData= {
  A 10
  B 4
  C 9
  D 5
  E 5
  F 2
}

```

```
G 2
H 6
I 1
J 3
K 1
L 3
M 1
N 2
O 1
```

```
}
```

```
SampleName="leuconucha"
SampleSize=20
SampleData= {
  A 3
  D 3
  H 8
  J 1
  K 1
  N 1
  P 2
  Q 1
```

```
}
```

```
[[Structure]]
```

```
StructureName = "between subspecies san jacinto=leuconucha"
NbGroups=1
Group={
  "melanotis"
  "leuconucha"
}
```

```
[Profile]
```

```
Title="Sitta pygmaea subspecies AMOVA"
NbSamples= 2
DataType= DNA
GenotypicData= 0
LocusSeparator= NONE
GameticPhase= 0
RecessiveData= 0
RecessiveAllele= null
MissingData= '?'
```

```
[Data]
```

```
[[HaplotypeDefinition]]
```

```
HaplListName="Sitta pygmaea Haplotype List"
HaplList={
[Insert 1639-base pair haplotype list without haplotypes R, S and T
here]
```

```

}

[[Samples]]

  SampleName="melanotis"
  SampleSize=34
  SampleData= {
    A 10
    B 3
    C 7
    D 5
    E 5
    F 1
    G 2
    H 1
  }

  SampleName="leuconucha"
  SampleSize=41
  SampleData= {
    A 3
    B 1
    C 2
    D 3
    F 1
    H 13
    I 1
    J 4
    K 2
    L 3
    M 1
    N 3
    O 1
    P 2
    Q 1
  }

[[Structure]]

  StructureName = "between subspecies san bernardino=leuconucha"
  NbGroups=1
  Group={
    "melanotis"
    "leuconucha"
  }

```

```

[Profile]

Title="Sitta pygmaea AMOVA"
NbSamples= 11
DataType= DNA
GenotypicData= 0

```

```
LocusSeparator= NONE
GameticPhase= 0
RecessiveData= 0
RecessiveAllele= null
MissingData= '?'
```

```
[Data]
```

```
[[HaplotypeDefinition]]
```

```
  HaplListName="Sitta pygmaea Haplotype List"
```

```
  HaplList={
```

```
[Insert 1639-base pair haplotype list without haplotypes R, S and T here]
```

```
  }
```

```
[[Samples]]
```

```
  SampleName="PIUTE"
```

```
  SampleSize=6
```

```
  SampleData= {
```

```
    A 2
```

```
    B 3
```

```
    C 1
```

```
  }
```

```
  SampleName="NOROESTE"
```

```
  SampleSize=7
```

```
  SampleData= {
```

```
    A 3
```

```
    C 2
```

```
    D 1
```

```
    E 1
```

```
  }
```

```
  SampleName="FRAZIER"
```

```
  SampleSize=7
```

```
  SampleData= {
```

```
    C 3
```

```
    D 3
```

```
    E 1
```

```
  }
```

```
  SampleName="WINSTON"
```

```
  SampleSize=7
```

```
  SampleData= {
```

```
    A 1
```

```
    C 1
```

```
    D 1
```

```
    E 3
```

```
    F 1
```

```
  }
```

```
  SampleName="BLUE RIDGE"
```

```
  SampleSize=7
```

```

SampleData= {
  A 4
  G 2
  H 1
}

SampleName="DELAMAR"
SampleSize=7
SampleData= {
  B 1
  C 2
  H 3
  I 1
}

SampleName="ONYX"
SampleSize=7
SampleData= {
  J 2
  K 1
  L 3
  M 1
}

SampleName="GRINNELL"
SampleSize=7
SampleData= {
  F 1
  H 2
  J 1
  N 2
  O 1
}

SampleName="BLACK"
SampleSize=7
SampleData= {
  A 2
  D 1
  H 2
  K 1
  P 1
}

SampleName="THOMAS"
SampleSize=6
SampleData= {
  A 1
  D 1
  H 3
  Q 1
}

SampleName="LAGUNA"

```

```

SampleSize=7
SampleData= {
  D 1
  H 3
  J 1
  N 1
  P 1
}

[[Mantel]]
MatrixSize=11
MatrixNumber=2
Ymatrix=Fst
YmatrixLabels= {
  "PIUTE"
  "NOROESTE"
  "FRAZIER"
  "WINSTON"
  "BLUE RIDGE"
  "DELAMAR"
  "ONYX"
  "GRINNELL"
  "BLACK"
  "THOMAS"
  "LAGUNA"
}
DistMatMantel={
  0.00
  102.1 0.00
  92.0 23.4 0.00
  128.2 129.4 106.2 0.00
  138.2 151.6 128.3 23.8 0.00
  185.2 218.7 195.2 92.4 68.7 0.00
  206.0 240.5 217.2 113.5 89.9 22.0 0.00
  206.5 237.1 213.8 109.0 85.6 21.0 9.2 0.00
  234.1 253.1 230.1 124.0 103.1 53.5 42.0 34.9 0.00
  256.2 270.2 247.6 142.1 122.9 77.8 65.3 58.9 24.4 0.00
  339.5 338.5 317.3 217.0 202.0 165.6 152.3 146.6 112.2 87.9 0.00
}
UsedYMatrixLabels={
  "PIUTE"
  "NOROESTE"
  "FRAZIER"
  "WINSTON"
  "BLUE RIDGE"
  "DELAMAR"
  "ONYX"
  "GRINNELL"
  "BLACK"
  "THOMAS"
  "LAGUNA"
}

```

MRBAYES INPUT FILE

```
#NEXUS
BEGIN DATA;
DIMENSIONS NTAX=20 NCHAR=1639;
FORMAT DATATYPE=DNA MISSING=? GAP=- ;

MATRIX
[Insert 1639-base pair haplotype list here]
;
END;

begin mrbayes;
log start filename = pynuhap.2005e.log;
outgroup T;
charset molecules = 1-1639;
charset Cytb = 1-621;
charset ND6 = 622-1125;
charset Dloop = 1126-1639;
charset 1st_pos = 1-1125\3;
charset 2nd_pos = 2-1125\3;
charset 3rd_pos = 3-1125\3;
charset non_coding = 1126-1639;
partition by_codon = 4: 1st_pos, 2nd_pos, 3rd_pos, non_coding;
partition genes = 3: Cytb, ND6, Dloop;
partition all = 1: molecules;
set partition = by_codon;
databreaks 621 1125;
lset nst=2 code=vertmt;
prset ratepr=variable;
set autoclose=yes;
mcmc ngen=30000000 printfreq=10000 samplefreq=1000
nchains=4 savebrlens=yes filename=pynuhap.2005e.bay;
mcmc;
plot filename=pynuhap.2005e.bay.p;
sumt filename=pynuhap.2005e.bay.t burnin=3000 contype=halfcompat;
log stop;
end;
```

GEODIS INPUT FILE

Sitta pygmaea

11

1 PIUTE

6 35 36 21 N 118 23 37 W

2 NOROESTE

7 34 49 41 N 119 13 25 W

3 FRAZIER

7 34 46 10 N 118 58 39 W

4 WINSTON

7 34 21 10 N 117 56 06 W

5 BLUE

7 34 20 52 N 117 40 33 W

6 DELAMAR

7 34 17 05 N 116 55 58 W

7 ONYX

7 34 12 19 N 116 42 53 W

8 GRINNELL

7 34 08 30 N 116 46 41 W

9 BLACK

7 33 49 42 N 116 45 04 W

10 THOMAS

6 33 37 28 N 116 41 11 W

11 LAGUNA

7 32 51 08 N 116 26 06 W

6

Clade 1.1

3

F I O

1 0 1

3

4 6 8

1 0 1

0 1 0

0 0 1

Clade 1.5

3

C G P

0 1 1

8

1 2 3 4 5 6 9 11

1 2 3 4 0 2 0 0

0 0 0 0 2 0 0 0

0 0 0 0 0 0 1 1

Clade 1.6

3

B L M

0 1 1

3

1 6 7

3 1 0

0 0 3


```

0 0 1
Clade 1.7
5
D H J K Q
1 0 1 1 1
10
2 3 4 5 6 7 8 9 10 11
1 3 1 0 0 0 0 1 1 1
0 0 0 1 3 0 2 2 3 3
0 0 0 0 0 2 1 0 0 1
0 0 0 0 0 1 0 1 0 0
0 0 0 0 0 0 0 0 1 0
Clade 2.3
4
1.3 1.4 1.5 1.6
1 1 1 0
10
1 2 3 4 5 6 7 9 10 11
0 1 1 3 0 0 0 0 0 0
2 3 0 1 4 0 0 2 1 0
1 2 3 4 2 2 0 1 0 1
3 0 0 0 0 1 4 0 0 0
Clade 3.1
4
2.1 2.2 2.3 2.4
1 1 1 0
11
1 2 3 4 5 6 7 8 9 10 11
0 0 0 1 0 1 0 2 0 0 0
0 0 0 0 0 0 0 2 0 0 1
6 6 4 8 6 3 4 0 3 1 1
0 1 3 1 1 3 3 3 4 5 5

```

REFERENCES

- Alexander, M.P., and K.J. Burns. (unpublished).
Intraspecific phylogeography and adaptive divergence
in the white-headed woodpecker.
- Aquadro, C.F., and B.D. Greenberg. 1983. Human
mitochondrial DNA variation an evolution: analysis of
nucleotide sequences from seven individuals. *Genetics*.
103:287-312.
- Avise, J.C. 2000. *Phylogeography: The History and Formation
of Species*. Harvard University Press. Cambridge, MA.
- Avise, J.C. 1998. The history and purview of
phylogeography: a personal reflection. *Molecular
Ecology*. 7:371-379.
- Avise, J.C., and D. Walker. 1998. Pleistocene effects on
avian populations and the speciation process. *Proc R
Soc London B*. 265:457-463.
- Avise, J.C., J. Arnold, R.M. Ball, E. Bermingham, T. Lamb,
J.E. Neigel, C.A. Reeb, and N.C. Saunders. 1987.
Intraspecific phylogeography: the mitochondrial DNA
bridge between population genetics and systematics.
Ann Rev Ecol Syst. 18:489-522.
- Baker, A.J., and H.D. Marshall. 1997. Mitochondrial control
region sequences as tools for understanding evolution.

- In Avian Molecular Systematics and Evolution. D.P. Mindell (Ed.). pp. 51-82.
- Baker, A.J., T. Piersma, and L. Rosenmeier. 1994. Unraveling the intraspecific phylogeography of knots *Calidris canutus*: a progress report on the search for genetic markers. *Journal für Ornithologie*. 135:599-608.
- Barrowclough, G.F., J.G. Groth, L.A. Mertz, and R.J. Gutierrez. 2004. Phylogeographic structure, gene flow and species status in blue grouse (*Dendragapus obscurus*). *Molecular Ecology*. 13:1911-1922.
- Bent, A.C. 1948. Life histories of North American nuthatches, wrens, thrashers, and their allies. U.S. National Museum Bulletin 195. pp. 44-55.
- Berlin, S., and H. Ellegren. 2001. Clonal inheritance of avian mitochondrial DNA. *Nature*. 413:37-38.
- Berlin, S., N.G.C. Smith, and H. Ellegren. 2004. Do avian mitochondria recombine? *Journal of Molecular Evolution*. 58:163-167.
- Bermingham, E. and J.C. Avise. 1986. Molecular zoogeography of freshwater fishes in the southeastern United States. *Genetics*. 113:939-965.
- Birky, C.W.J., T. Muryama, and P. Fuerst. 1983. An approach to population and evolutionary genetic theory for

- genes in mitochondria and chloroplasts, and some results. *Genetics*. 103:513-527.
- Brown, W.M., G. Matthew, and A.C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci USA*. 76:1967-1971.
- Buerkle, C.A. 1999. The historical pattern of gene flow among migratory and nonmigratory populations of prairie warblers (Aves: Parulinae). *Evolution*. 53:1915-1924.
- Burns, K.J., and D.N. Barhoum. 2006. Population-level history of the wrentit (*Chamaea fasciata*): implications for comparative phylogeography in the California Floristic Province. *Molecular Phylogenetics and Evolution*. 38:117-129.
- Burton, R.S., and B.-N. Lee. 1994. Nuclear and mitochondrial gene genealogies and allozyme polymorphism across a major phylogeographic break in the copepod *Tigris californicus*. *Proc Natl Acad Sci USA*. 91:5197-5201.
- Calsbeek, R., J.N. Thompson, and J.E. Richardson. 2003. Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. *Molecular Ecology*. 12:1021-1029.

- Cann, R.L., W.M. Brown, and A.C. Wilson. 1984. Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics*. 106:479-499.
- Clement, M., D. Posada, and K.A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology*. 9:1657-1659.
- DeChaine, E.G., and A.P. Martin. 2004. Historical cycles of fragmentation and expansion in *Parnassius smintheus* (Papilionidae) inferred using mitochondrial DNA. *Evolution*. 58:113-127.
- Dorf, E. 1976. Climatic changes of the past and present. In *Paleobiogeography*. C.A. Ross (Ed.). pp. 384-412.
- Excoffier, L., P.E. Smouse, and J.M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. 131:479-491.
- Farris, J.S., M Källersjö, A.G. Kluge, and C. Bult. 1995. Testing significance of incongruence. *Cladistics*. 10:315-319.
- Fleischer, R.C., C.E. McIntosh, and C.L. Tarr. 1998. Evolution on a volcanic conveyor belt: Using phylogenetic reconstructions and K-Ar-based ages of

- the Hawaiian Islands to estimate molecular evolutionary rates. *Molecular Ecology*. 7:533-545.
- Fok, K.W., C.M. Wade, and D.T. Parkin. 2002. Inferring the phylogeny of disjunct populations of the azure-winged magpie *Cyanopica cyanus* from mitochondrial control region sequences. *Proc R Soc London B*. 269:1671-1679.
- Garrett, K., and J. Dunn. 1981. Birds of Southern California: Status and Distribution. Los Angeles Audubon Society. Los Angeles, CA. p 271.
- Gill, F. 1999. Speciation in North American Chickadees: II. Geography of mtDNA Haplotypes in *Poecile carolinensis*. *The Auk*. 116:274-277.
- Graves, J.E., S.D. Ferris, and A.E. Dizon. 1984. High genetic similarity of Atlantic and Pacific skipjack tuna demonstrated with restriction endonuclease analysis of mitochondrial DNA. *Marine Biology*. 79:315-319.
- Grinnell, J., and A.H. Miller. 1944. The Distribution of the Birds of California. Cooper Ornithological Club. Pacific Coast Avifauna: Number 27. pp. 318-321.
- Gyllensten, U. 1985. The genetic structure of fish: differences in the intraspecific distribution of biochemical genetic variation between marine,

- anadromous, and freshwater species. *Journal of Fish Biology*. 26:691-699.
- Hartl, D.L., and A.G. Clark. 1997. Principles of Population Genetics, Third Edition. Sinauer Associates, Inc. Sunderland, MA.
- Hasegawa, M., H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*. 21:160-174.
- Hillis, D.M., B.K. Mable, A. Larson, S.K. Davis, and E.A. Zimmer. 1996. Nucleic Acids IV: Sequencing and Cloning. In *Molecular Systematics*. D.M. Hillis, C. Moritz, and B.K. Mable (Eds). Sinauer Associates, Inc. Sunderland, MA. pp. 321-381.
- Hoeh, W.R., K.H. Blakley, and W.M. Brown. 1991. Heteroplasmy suggests limited biparental inheritance of *Mytilus* mitochondrial-DNA. *Science*. 251:1488-1490.
- Hoelzer, G.A. 1997. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees revisited. *Evolution*. 51:622-626.
- Howell, S.N.G., and S. Webb. 1995. A Guide to the Birds of Mexico and Northern Central America. Oxford University Press. New York, NY. p. 554.

- Hudson, R.R., D.D. Boos, and N.L. Kaplan. 1992. A statistical test for detecting geographic subdivision. *Molecular Biology and Evolution*. 9:138-151.
- Huelsenbeck, J.P., and F.R. Ronquist. 2001. MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics*. 17:754-755.
- Idaghdour, I., D. Broderick, A. Korrida, and F. Chbel. 2004. Mitochondrial control region diversity of the houbara bustard *Chlamydotis undulata* complex and genetic structure along the Atlantic seaboard of North Africa. *Molecular Ecology*. 13:43-54.
- Johnson, N.K., and C. Cicero. 2004. New mitochondrial DNA data affirm the importance of Pleistocene speciation in North American birds. *Evolution*. 58:1122-1130.
- Kingery, H.E., and C.K. Ghalambor. 2001. Pygmy Nuthatch (*Sitta pygmaea*). In *The Birds of North America*, No. 567. The Birds of North America, Inc. Philadelphia, PA.
- Kircham, J.J., L.A. Whittingham, and F.H. Sheldon. 2000. Relationships among cave swallow populations (*Petrochelidon fulva*) determined by comparisons of microsatellite and cytochrome *b* data. *Molecular Phylogenetics and Evolution*. 14:107-121.

- Kondo, R., Y. Satta, E.T. Matsuura, H. Ishiwa, N. Takahata, and S.I. Chigusa. 1990. Incomplete maternal transmission of mitochondrial DNA in *Drosophila*. *Genetics*. 126:657-663.
- Kumar, S., K. Tamura, I.B. Jakobsen, and M. Nei. 2001. MEGA2: Molecular Evolutionary Genetics Analysis software. *Bioinformatics*. 17:1244-1245.
- Kvist, L., J. Martens, A.A. Nazarenko, and M. Orell. 2003. Paternal leakage of mitochondrial DNA in the great tit (*Parus major*). *Molecular Biology and Evolution*. 20:243-247
- Kvist, L., J. Martens, A. Ahola, and M. Orell. 2001. Phylogeography of a palearctic sedentary passerine, the willow tit (*Parus montanus*). *J Evolutionary Biology*. 14:930-941.
- Lamb, T., T.R. Jones, and P.J. Wettstein. 1997. Evolutionary genetics and phylogeography of tassel-eared squirrels (*Sciurus aberti*). *J Mammalogy*. 78:117-133.
- Little, E.L., Jr. 1971. Atlas of United States trees, volume 1: Conifers and Important Hardwoods. Misc. pub. 1146. U.S. Dept. of Agriculture. Washington, D.C.

- Lovette, I. 2004. Mitochondrial dating and mixed support for the "2% rule" in birds. *The Auk*. 121:1-6.
- Lynch, M., and T.J. Crease. 1990. The analysis of population survey data on DNA sequence variation. *Molecular Biology and Evolution*. 7:377-394.
- Maddison, D.R., and W.P. Maddison. 2000. MacClade 4: Analysis of phylogeny and character evolution. Version 4.0. Sinauer Associates, Inc. Sunderland, MA.
- Masta, S. 2000. Phylogeography of the jumping spider *Habronattus pugillus* (Araneae: Salticidae): recent vicariance of sky island populations. *Evolution*. 54:1699-1711.
- Metcalf, A.E., L. Nunney and B.C. Hyman. 2001. Geographic patterns of genetic differentiation within the restricted range of the endangered species Stephen's kangaroo rat, *Dipodomys stephensi*. *Evolution*. 55:1233-1244.
- Moore, W.S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution*. 49:718-726.
- Moore, W.S. 1997. Mitochondrial-gene trees versus nuclear-gene trees, a reply to Hoelzer. *Evolution*. 51:627-629.

- Moore, W.S., and V.R. DeFilippis. 1997. The window of taxonomic resolution for phylogenies based on mitochondrial cytochrome *b*. In *Avian Molecular Evolution and Systematics*. D.P. Mindell (Ed.). pp. 83-119.
- National Geographic Society. 2002. *National Geographic Society Field Guide to the Birds of North America*, fourth edition. National Geographic Society. Washington, D.C. p. 332.
- Norris, R.A. 1958. Comparative biosystematics and life history of the nuthatches *Sitta pygmaea* and *Sitta pusilla*. *University of California Publications in Zoology*. 56:119-300.
- Patten, M.A., and P. Unitt. 2002. Diagnosability versus mean differences of sage sparrow subspecies. *The Auk*. 119:26-35.
- Pearce, R.L., J.J. Wood, Y. Arthukin, T.P. Birt, M. Damus, and V.L. Friesen. 2002. Mitochondrial DNA suggests high gene flow in ancient murrelets. *The Condor*. 104:84-91.
- Pielou, E.C. 1991. *After the Ice Age: the return of life to glaciated North America*. University of Chicago Press. Chicago, IL.

- Posada, D., and K.A. Crandall. 1998. Modeltest: testing the model of DNA substitution. *Bioinformatics*. 14:817-818.
- Posada, D., K.A. Crandall, and A.R. Templeton. 2000. GeoDis: a program for the cladistic analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*. 9:487-488.
- Prychitko, T.M., and W.S. Moore. 2000. Comparative evolution of the mitochondrial cytochrome b gene and nuclear β -fibrinogen intron 7 in woodpeckers. *Molecular Biology and Evolution*. 17:1101-1111.
- Rising, J.D. 2005. Subspecies and their significance in modern ornithology. In Program & Abstracts: 123rd Stated Meeting of the American Ornithologists' Union. p. 55.
- Rokas, A., E. Ladoukakis, and E. Zouros. 2003. Animal mitochondrial DNA recombination revisited. *Trends in Ecology and Evolution*. 18:411-417.
- Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin v2.0: a software for population genetics data analysis. Genetics and Biometry Laboratory. University of Geneva, Switzerland.
- Sgariglia, E.A., and K.J. Burns. 2003. Phylogeography of the California thrasher (*Toxostoma redivivum*) based on

- nested-cleade analysis of mitochondrial-DNA variation.
The Auk. 120:346-361.
- Small, A. 1994. California Birds: Their Status and
Distribution. Ibis Publishing Company. Vista, CA. p.
182.
- Sorenson, M.D., and T.W. Quinn. 1998. Numts: a challenge
for avian systematics and population biology. *The Auk*.
115:214-221.
- South Coast Wildlands. 2005. South Coast Missing Linkages
Project: A Linkage Design for the San Bernardino-San
Jacinto Connection (Draft Report). South Coast
Wildlands, Idyllwild, CA.
- Swofford, D.L. 2003. Paup* Phylogenetic Analysis Using
Parsimony (*and Other Methods). Version 4. Sinauer
Associates, Inc. Sunderland, MA.
- Swofford, D.L., G.J. Olsen, P.J. Waddell, and D.M. Hillis.
1996. Phylogenetic Inference. *In* Molecular
Systematics. D.M. Hillis, C. Moritz, and B.K. Mable
(Eds). Sinauer Associates, Inc. Sunderland, MA. pp.
407-514.
- Tarr, C.L. 1995. Primers for amplification and
determination of mitochondrial control-region

sequences in oscine passerines. *Molecular Ecology*.
4:527-529.

Templeton, A.R. 1998. Nested clade analyses of
phylogeographic data: testing hypotheses about gene
flow and population history. *Molecular Ecology*. 7:381-
397.

Templeton, A.R., E. Boerwinkle, and C.F. Sing. 1987. A
cladistic analysis of phenotypic associations with
haplotypes inferred from restriction endonuclease
mapping. I. Basic theory and an analysis of alcohol
dehydrogenase activity in *Drosophila*. *Genetics*.
117:343-351.

Templeton, A.R., and C.F. Sing. 1993. A cladistic analysis
of phenotypic associations with haplotypes inferred
from restriction endonuclease mapping. IV. Nested
analyses with cladogram uncertainty and recombination.
Genetics. 134:659-669.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin,
and D.G. Higgins. 1997. The ClustalX windows
interface: flexible strategies for multiple sequence
alignment aided by quality analysis tools. *Nucleic
Acids Research*. 25:4876-4882.

- Ward, R.D., M. Woodmark, and D.O.F. Skibinski. 1994. A comparison of genetic diversity in marine, freshwater, and anadromous fishes. *Journal of Fish Biology*. 44:213-232.
- Wright, S. 1931. Evolution in Mendelian populations. *Genetics*. 16:97-159.
- Wright, S. 1943. Isolation by distance. *Genetics*. 28:114-138.
- Zink, R.M. 2004. The role of subspecies in obscuring avian biological diversity and misleading conservation policy. *Proc R Soc London B*. 271:561-564.
- Zouros, E. K.R. Freeman, A. Oberhauser-Ball, and G.H. Pogson. 1992. Direct evidence for extensive paternal mitochondrial DNA inheritance in the mussel *Mytilus*. *Nature*. 359:412-414.