MITOCHONDRIAL-DNA SEQUENCE VARIATION, AND THE GENETIC AFFINITY OF *PEROMYSCUS* FROM TRIANGLE ISLAND, BRITISH COLUMBIA, CANADA

A Senior Honors Thesis

by

ROOZBEH ARIANPOUR

Submitted to the Office of Honors Programs
& Academic Scholarships
Texas A&M University
in partial fulfillment of the requirements of the

UNIVERSITY UNDERGRADUATE
RESEARCH FELLOWS

April 2001

Group: Biochemistry

MITOCHONDRIAL-DNA SEQUENCE VARIATION, AND THE GENETIC AFFINITY OF *PEROMYSCUS* FROM TRIANGLE ISLAND, BRITISH COLUMBIA, CANADA

A Senior Honors Thesis

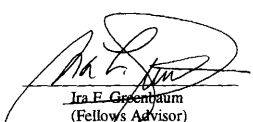
by

ROOZBEH ARIANPOUR

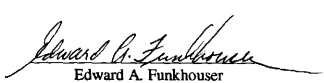
Submitted to the Office of Honors Programs
& Academic Scholarships
Texas A&M University
in partial fulfillment for the designation of

UNIVERSITY UNDERGRADUATE
RESEARCH FELLOW

Approved as to style and content by:



Ira F. Greenbaum
(Fellows Advisor)



Edward A. Funkhouser
(Executive Director)

April 2001

Group: Biochemistry

ABSTRACT

Mitochondrial-DNA Sequence Variation, and the Genetic

Affinity of *Peromyscus* from Triangle Island,

British Columbia, Canada. (April 2001)

Roozbeh Arianpour
Department of Biochemistry and Genetics
Texas A&M University

Fellows Advisor: Dr. Ira F. Greenbaum
Department of Biology

To determine the genetic and specific affinity of deer mice (*Peromyscus)* inhabiting Triangle Island, British Colombia, Canada, a 1,439 base-pair (bp) fragment of the mitochondrial genome (containing the ND3, ND4L, and a portion of the ND4 genes) was sequenced for twenty-seven individuals and compared to corresponding data from reference individuals of *P. maniculatus* and *P. keeni.* Analysis of sequences for the deer mice from Triangle Island revealed the existence of four sequences with a total percent sequence divergence of 0.093%. Among the individuals, one sequence predominated and characterized 14 of the 27 (52%) individuals. Each of the other sequences occurred in smaller, but similar, proportions (6, 22%; 4, 15%; 3, 11%) of the population sample. Phylogenetic analyses grouped the four sequences from the Triangle Island population and placed them within a cluster otherwise comprised of the *P. keeni* reference samples. The reference samples of *P. maniculatus* clustered outside and essentially equidistant from the *P. keeni*/Triangle Island deer mouse assemblage. These data clearly indicate

that the deer mice from Triangle Island are representative of *P. keeni*, and support the taxonomic relegation of these mice from *P. m. triangularis* to *P. k. triangularis*.

DEDICATION

This work is dedicated to those who have had the greatest impact on my growth and progress: to my sister, Rouzheen Arianpour, whose love and understanding has always sheltered me through the storms, whose brilliance and intellect continue to inspire me, and whose smile promises a golden tomorrow with a myriad of possibilities; to my mother, Farideh Arianpour, who has always been the pillar of strength in my life, never stopped believing in me and my goals, and nurtured me into the person I am today; and to my father, Reza Arianpour, for showing me the ugliness of improbity, fostering me with the confidence and ardor to succeed, and never allowing me to settle for mediocrity.

ACKNOWLEDGMENTS

I am foremost indebted to my advisor, Dr. Ira F. Greenbaum, as his guidance, advice, support, and friendship have immensely contributed to my education at Texas A&M University. Financial support was provided by NIH Grant GM27014 to Dr. Greenbaum. I am further indebted to Scott Chirhart, for introducing me to research and his continued support throughout the years. I am grateful to Dr. Rodney L. Honeycutt for giving me the opportunity to work in his laboratory. For his aid in analyses using PAUP I thank Dr. John C. Patton. For assistance in the laboratory I thank Tamara J. Gilbert, Katie Connell, Nickie Edstrom, Brandon Hill, Josh Weerasinghe, and Deepa Deshpande. Furthermore, I thank the administrative efforts of Dr. Donald R. Dickson and the Fellows committee.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

INTRODUCTION

Mice of the genus *Peromyscus* are among the most successful small mammals native to North and Central America. The vast distribution and diversity of species and habitats in the genus have made it a favorite model organism in studies encompassing virtually all areas of organismal biology. In particular, mice of this genus have played a central role in the development of systematic mammalogy in North America with this impact being favorably compared to the importance of *Drosophila* in the growth of systematic biology and evolutionary genetics as a whole (Carleton 1989). Of the species of *Peromyscus*, none has been more intensively studied than *P. maniculatus*. This species ranges from the Atlantic to Pacific seaboards and from southern Alaska to central Mexico, occupies nearly all habitats within this range and exhibits phenotypic variation exemplified in the recognition of approximately 70 subspecies. While considerable attention has been paid to the partitioning of genetic variation within *P. maniculatus* and its closely related allies in the *P. maniculatus* species group, the complexity of *P. maniculatus* is both the main obstacle and major source of information relative to understanding the diversity and evolution of this group of mice.

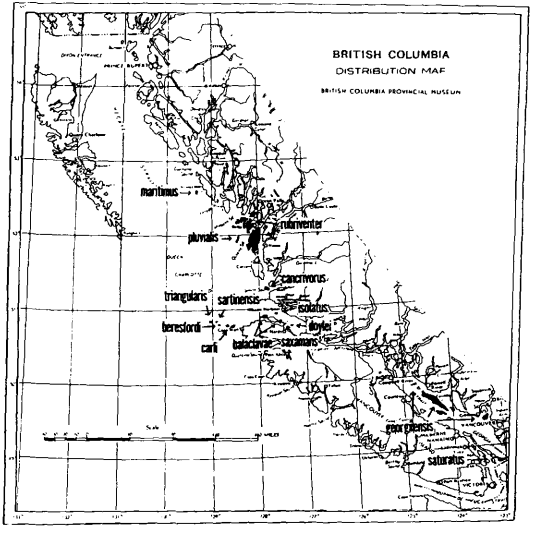
A variety of recent studies of genetic and morphological characters have raised questions concerning the appropriate circumscription of *P. maniculatus*. In particular, the taxonomy of deer mice in the Pacific Northwest has been revised by the recognition of the existence of a sibling species, *P. keeni*. As currently recognized, *P. keeni* is distributed west of the Coastal Mountains and north of about 50° N latitude (Greeenbaum 1999) and includes the previously recognized form *P. sitkensis* (from

several islands off the coast of southeastern Alaska, Hall 1981) and various northwestern subspecies which were previously assigned to *P. maniculatus* (Hogan et al. 1993). The key to the recognition of *P. keeni* was the discovery of its chromosomal distinction from *P. maniculatus*. Populations now recognized as *P. keeni* were found to have a fundamental number (FN, the number of arms in the autosomal karyotype) of 85-92; northwestern populations of *P. maniculatus* have a FN of 74-78, (Gunn and Greenbaum 1986; Gunn 1988). Although lacking in fixed allelic differences, allozymic analyses indicated a consistent dichotomy in allele frequencies between the two FN groups (Calhoun and Greenbaum 1991; Hogan et al. 1993). Multivariate morphological analyses ultimately revealed a single character (length of tail > 98 mm indicating *P. keeni*) that accurately classified greater than 95% of genetically identified specimens from mainland and insular populations from the Pacific Northwest (Allard and Greenbaum 1988). Analyses of restriction fragment length polymorphisms of the entire mitochondrial DNA (mtDNA) also separated these mice into their respective chromosomal groups. Further analyses of the mtDNA revealed species-specific fragments and sequences within the 1,439 bp region including the ND3, ND4L, and a portion of the ND4 genes (Hogan et al. 1997).

Despite the well-documented dichotomy of *P. maniculatus* and *P. keeni*, resolution of the taxonomy of the deer mice inhabiting the Pacific Northwest continues to be complicated by extensive insular endemism and the lack of critical data and difficulty of obtaining specimens from these remote locations. Based on morphological descriptions (Cowan and Guiguet 1965) and available karyotypic data (Thomas 1973),

however, Hogan et al. (1993) hypothesized that the deer mice currently recognized as *P. m. carli*, *P. m. doylei*, and *P. m. triangularis* (Fig. 1) are actually representative of *P. keeni*. The recent acquisition of frozen specimens of *P. m. triangularis* provided the opportunity to determine their genetic and taxonomic affinity. To this end the DNA sequence of the ND3/ND4L/ND4 region of the mtDNA was determined for the specimens of *P. m. triangularis* and compared to those of previously identified representatives of *P. maniculatus* and *P. keeni*.

Fig. 1. Distribution of subspecies of *P. maniculatus* on coastal islands of British Columbia, Canada (from Cowan and Guiguet, 1965).



MATERIALS AND METHODS

Frozen specimens of *Peromyscus* (n=28) from Triangle Island, British Columbia (51° 52'N, 129° 05'W) were provided by Louise K. Blight, Department of Biological Sciences, Simon Fraser University, Burnaby BC, Canada V5A 1S6 under permit No. 85700-40/ST9910186, Ministry of Environment, Land and Parks, BC Parks, Strathcona District. Liver and spleen samples were removed from partially thawed carcasses and frozen (-80°) until use. The specimens were then prepared as vouchers (skin and skull or full skeleton) and deposited in the Texas Cooperative Wildlife Collection at Texas A&M University.

The Qiagen purification kit and procedure was used to isolate mtDNA from liver or spleen samples. Amplification and sequencing of the 1,439 bp region of mtDNA, from the 3' end of the glycine tRNA (^{gly}tRNA) through 672 bp of the 5' end of the ND4 gene, generally followed the techniques of Arevalo et al. (1994). The primers used in amplification of this region included: PI', Marg, MargRev, ND4L, Rataway, and Nap2 (Arevalo et al. 1994). Amplification (Perkin Elmer/Cetus DNA Thermal Cycler) proceeded in three stages with the following components and concentrations: 1 uL DNA (approximately 100 ng), 15 uL H_2O, 2.5 uL of 10X PCR Buffer II (PE Applied Biosystems), 2 uL of 25 mM $MgCl_2$, 0.5 uL BSA, 4 uL of 8 mM dNTPs (Amersham Pharmacia Biotech), 0.4 uL of forward and reverse primers, and 0.2 uL Taq (TaKaRa, Japan). Reaction conditions included an initial denaturation cycle at 95°C for five min, followed by 35 cycles including 45 sec at 95°C, 30 sec at 50°C, and 90 sec at 72°C, and

concluded with an extension time cycle spanning 10 min at 72°C. Amplified products were purified using the QIAquick PCR purification protocol (Qiagen).

The sequencing reaction included 1 uL of amplified/purified DNA, 2 uL sequencing standard (DNA sequencing kit, PE Applied Biosystems), 6 uL H_20, and 1 uL of desired primer with a concentration of 3 nm. The sequencing reaction began with one denaturation cycle of 3 min at 95°C, and continued with 30 cycles each including 25 sec at 95°C, 25 sec at 47°C, and 4 min at 60°C (Perkin Elmer/Cetus DNA Thermal Cycler). All sequences were determined by electrophoresis in an Applied Biosystems model 377 automated DNA sequencer. All sequences were aligned for analysis using the Sequencher 3.1.1 computer program.

As the ND3/ND4L/ND4 sequences previously reported for *P. maniculatus* and *P. keeni* were generated and read using manual techniques, DNA from individuals reported (Hogan et al. 1997) as characterized by each of the relevant sequence variants was re-isolated and sequenced as described above. The specimens and localities included in this analysis were: *P. k. interdictus*: Canada, British Columbia, Vancouver Island, Mt. Washington Ski Area (n=3); *P. k. oreas*: U.S.A., Washington, Graysharbor Co., 3 mi N, 1 mi E Grisdale, Satsop Workcamp (n=3); *P. m. austerus*: Canada, British Columbia, Vancouver Island, 35.7 km W Port Alberni (n=3); *P. m. rufinus*: U.S.A., Colorado, La Plata Co., Elk Park (n=1).

RESULTS AND DISCUSSION

Analysis of the ND3/ND4L/ND4 region for the twenty-seven deer mice from Triangle Island revealed the presence of four sequences with a total percent sequence divergence (Kimura 1980) of 0.093%. The variation within the coding region among the four sequences comprised a total of ten transitions of which seven occurred at third and three occurred at first positions within codons. Among the individuals, one of the four sequences (sequence A) characterized 14 of the 27 individuals (52%). The other three sequences (sequences B, C, and D) occurred in smaller, but similar, proportions (6, 22%; 4, 15%; 3, 11%, respectively) of the population sample.

The automated sequencing of the previously described ND3/ND4L/ND4 sequences for *P. maniculatus* and *P. keeni* yielded four subspecies-specific sequences identical to those reported by Hogan et al. (1997). Base-pair differences and percent sequence divergences for the *P. maniculatus* and *P. keeni* sequences (not reported by Hogan et al. 1997) and compared to the predominant sequence for the deer mice from Triangle Island are presented in Table 1. Of the total substitutions among the sequences from the reference specimens of *P. maniculatus* and *P. keeni*, 33% occurred at first, 8% at second, and 59% at third positions of codons and displayed a general 4:1 bias in favor of transitions over transversions.

The relationship of the deer mice from Triangle Island to the reference samples of *P. maniculatus* and *P. keeni* was interpreted from sequence-distance data (Table 1) employing the two-parameter model of Kimura 1980. The distance data were analyzed

Table 1. Number of base-pair substitutions in the coding region and total percent sequence divergence (in parentheses) between the sequences for the ND3/ND4L/ND4 region of the mtDNA for the deer mice from Triangle Island (*"triangularis"*) and each of the reference subspecies examined.

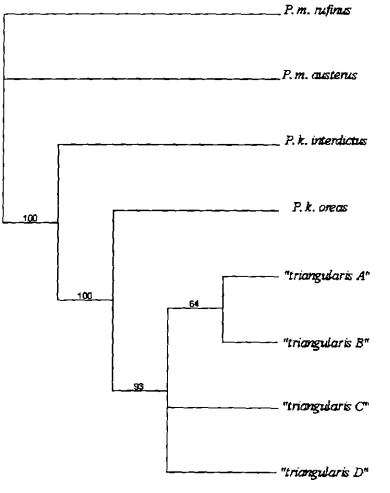
Subspecies	2	3	4	5
1. *P. m. austerus*	19 (1.46)	55 (4.25)	66 (4.80)	63 (4.73)
2. *P. m. rufinus*		57 (4.18)	60 (4.53)	61 (4.46)
3. *P. k. interdictus*			25 (1.95)	24 (1.88)
4. *P. k. oreas*				7 (0.49)
5. *"triangularis"*[1]				

[1]Computations based on the predominant (A) sequence for the deer mice from Triangle Island.

using maximum parsimony (exhaustive search) and neighbor-joining (1000 bootstrap replications) methods of the Phylogenetic Analysis Using Parsimony (PAUP) computer program (Swofford 1993, version 4.0b). Similar to the findings of Hogan et al. (1997) alternative weighting schemes did not alter the topology of the most parsimonious tree obtained from equal weighting of characters (transitions versus transversions) for the ND3/ND4L/ND4 sequences. All analyses produced the single most parsimonious tree (branch length = 95) retained by the exhaustive search and presented in Figure 2. The phylogenetic analysis grouped the four sequences from the Triangle Island population and placed these within a cluster otherwise comprised of the *P. keeni* reference samples. The reference samples of *P. maniculatus* clustered outside and essentially equidistant from the *P. keeni*/Triangle Island deer mouse assemblage. These data clearly indicate that the deer mice from Triangle Island are representative of *P. keeni* and support the suggestion (Hogan et al. 1993) that the populations of *Peromyscus* inhabiting the islands north of Vancouver Island be referred to *P. keeni*.

Although Hogan et al. (1997) concluded that the ND3/ND4L/ND4 region of the mtDNA of deer mice is characterized by little or no intraspecific variation, this conclusion was based primarily on restriction-enzyme analyses of small samples ($n<20$) and sequences from one or two individuals per species. Using an increased number of restriction enzymes and larger numbers of individuals, Walpole et al. (1997) reported more extensive variation in this region of the mtDNA both among and within populations of *P. eremicus*. The existence of at least four sequences maintained in the ND3/ND4L/ND4 region of the deer mice from Triangle Island, supports the conclusion

Fig. 2. The single tree retained from the maximum parsimony (exhaustive search) analysis of the ND3/ND4L/ND4 sequence distances among the taxa included in this study. Bootstrap values are those obtained for the analysis with characters equally weighted.



that this region of the mtDNA is subject to relatively rapid evolution and high levels of sequence polymorphism. Indeed, considering the likelihood of genetic drift associated with the insularity of the *Peromyscus* from Triangle Island, considerably higher levels of ND3/ND4L/ND4 variation can be expected for mainland populations of *P. keeni* and *P. maniculatus*.

Current geographic considerations would appear to favor the hypothesis that the *Peromyscus* of Triangle Island originated from source populations dispersed from Vancouver Island. In addition to Triangle Island being only 30 miles from northeastern-most Vancouver Island, the distance between Triangle and Vancouver Islands is interspersed with a series of islands (the Scott Islands) that are inhabited by *Peromyscus*. However, various studies have suggested that a Pleistocene glacial refugium existed on the now submerged continental shelf beneath the waterways (Hecate Strait and Queen Charlotte Sound) connecting the Queen Charlotte Islands, Vancouver Island, and the coastal fringe of mainland British Columbia (Byun et al. 1997). If *P. keeni* was initially isolated from *P. maniculatus* in such a Hecate refugium, a close phylogeographic association of modern populations on the islands in Queen Charlotte Sound and Vancouver Island might not be expected and a Vancouver Island origin for the deer mice of Triangle Island (and the other Scott Islands) would be highly unlikely. This study was not, however, designed to address the biogeographic history of the complex distribution of *P. keeni*. A critical evaluation of the alternative biogeographic hypotheses for *P. keeni* will require extensive sampling and analyses of genetic variation

within and among populations from across the insular and mainland range of this species.

REFERENCES

Allard, M.W., and Greenbaum, I.F. 1988. Morphological variation and taxonomy of chromosomally differentiated *Peromyscus* from the Pacific Northwest. Can. J. Zool. **66**: 2734-2739.

Arevalo, E., Davis, S.K., and Sites, J.W. 1994. Mitochondrial DNA sequence divergence and phylogenetic relationships among eight chromosome races of the *Sceloporus grammicus* complex (Phrynosomatidae) in Central Mexico. Sys. Biol. **43**: 387-418.

Byun, S.A., Koop, B.F., and Reimchen T.E. 1997. North American black bear mtDNA phylogeography: implications for morphology and the Haida Gwaii glacial refugium controversy. Evolution. **51**: 1647-1653.

Calhoun, S.W., and Greenbaum, I.F. 1991. Evolutionary implications of genic variation among insular populations of *Peromyscus maniculatus* and *Peromyscus oreas*. J. Mammal. **72(2)**: 248-262.

Carlton, M.D. 1989. Systematics and Evolution. In advances in the study of *Peromyscus* (Rodentia). Edited by G.L. Kirkland and J.N. Layne. Texas Tech University Press, Lubbock. pp. 7-141.

Cowan, I. McT., and Guiguet, C.J. 1965. The mammals of British Columbia. Revised ed. B.C. Prov. Mus. Nat. His. Anthropol. Handb. No. 11.

Greenbaum, I.F. 1999. Northwestern deer mouse/*Peromysucs keeni*. In The Smithsonian book of North American mammals. Edited by D.E. Wilson and S. Ruff. Smithsonian Institution Press, Washington, D.C. pp. 571-572.

Gunn, S.J. 1988. Chromosomal variation and differentiation among insular populations Of *Peromyscus* from the Pacific Northwest. Can. J. Zool. **66**: 2726-2733.

Gunn, S.J., and Greenbaum, I.F. 1986. Systematic implications of karyotypic and morphological variation in mainland *Peromyscus* from the Pacific Northwest. J. Mammal. **67**:294-304.

Hall, E.R. 1981. The mammals of North America. John Wiley and Sons, New York.

Hogan, K.M., Davis, S.K., and Greenbaum I.F. 1997. Mitochondrial-DNA analysis of the systematic relationship within the *Peromyscus maniculatus* species group. J. Mammal. **78(3)**:733-743.

Hogan, K.M., Hedin, M.C., Koh, H.S., Davis, S.K., and Greenbaum, I.F. 1993. Systematic and taxonomic implications of karyotypic, electrophoretic, and mitochondrial-DNA variation in *Peromyscus* from the Pacific Northwest. J. Mammal. **74**: 819-831.

Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. J. Mol. Evol. **16**: 111-120.

Swofford, D.L. 1993. PAUP: phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign. pp. 98.

Thomas, B. 1973. Evolutionary implications of karyotypic variation in some insular *Peromyscus* from British Columbia, Canada. Cytologia. **38**: 485-495.

Walpole, D.K., Davis, S.K., and Greenbaum, I.F. 1997. Variation in mitochondrial DNA in populations of *Peromyscus eremicus* from the Chihuahuan and Sonoran Desert. J. Mammal. **78(2)**: 397-404

VITA

Roozbeh Arianpour was born on March 25, 1980 to Reza and Farideh Arianpour in Ahwaz, Iran. He graduated from Robert E. Lee High School, Tyler, Texas, in 1998. He then attended Texas A&M University and pursued a Bachelor of Science degree in Genetics. In addition to the Texas A&M University Honors program other academic honors include: Phi Eta Sigma, Sigma Xi, Gamma Sigma Delta, The National Collegiate Honor Society, and Golden Key National Honor Society; in 1999 he was the awarded the Edward A. Funkhouser Scholarship in accord with the University Undergraduate Research Fellows.

His permanent mailing address is 2840 Woodlake Dr., Tyler, Texas, 75701.