Phylogeny, phylogeography and genetic structure of the North American diaptomid copepod, *Hesperodiaptomus shoshone* (Copepoda: Calanoida).

Malgorzata A. Marszalek

A Thesis
in
The Department
of
Biology

Presented in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy at
Concordia University
Montreal, Quebec, Canada
February 2007

Your file Votre référence
ISBN: 978-0-494-30115-9
Our file Notre référence
ISBN: 978-0-494-30115-9

ABSTRACT

Phylogeny, phylogeography and genetic structure of the North American diaptomid copepod, *Hesperodiaptomus shoshone* (Copepoda: Calanoida).

Malgorzata A. Marszalek, Ph.D.
Concordia University, 2007

The freshwater diaptomid copepod *Hesperodiaptomus shoshone* (S.A. Forbes, 1882) occurs in high altitudes of Western North America, ranging from British Columbia to Colorado in the Rocky Mountains, and in the Sierra Nevada of California. I examined the genetic variation of *H. shoshone* at different spatial scales to determine the extent of current demographic processes and historical events shaping the distribution and dispersal of this species. First, I examined the phylogenetic relationships within the genus *Hesperodiaptomus* based upon nuclear ribosomal DNA gene sequences. Phylogenetic reconstructions supported the monophyletic origin of the genus and revealed two clades supported by morphological characters. Second, I carried out a phylogeographic study using partial DNA sequences of the cytochrome oxidase I gene and from nuclear ribosomal DNA gene sequences from 51 populations representing the current distribution range of the species. The genetic divergence among populations was high, and indicates a low level of dispersal among populations. The data revealed the presence of two geographically distinct deep lineages (North/South), with further genetic sub-structuring within each region. I hypothesize that the observed patterns of genetic diversity and structure reflect historical dispersal events and episodic range displacement due to glaciations. Finally, I characterized three microsatellite markers for this

species and carried out a population genetic study to examine the genetic structure among 12 neighbouring populations in the Gunnison area of Colorado. The results indicated that *H. shoshone* have significant population differentiation and high genetic divergence suggesting a limited gene flow even at relatively small spatial scales.

ACKNOWLEDGEMENTS

I wish to express my sincere appreciation to my co-supervisors, Dr. Dayanandan and Dr. Maly, for their support, patience, and guidance to pursue a great project. They have influenced my life from both a professional and personal standpoint.

I would also like to thank my committee members and examiners for their invaluable comments that improved this thesis immensely, Dr. McLaughlin, Dr. Widden, Dr. Joyce, Dr. Brown and Dr. Turgeon.

I would like to thank Emma Rocke my field assistant who braved the mountains and wild animals in search of big red copepods. I could have not done it without you.

This thesis could not be possible without all the aquatic biologists and extremely kind individuals who provided me with necessary copepod samples and directions to my sampling sites. A warm thanks to: Andrew Kramer for *H. shoshone* Sierra Nevada populations, Stan Dodson for Hesperodiaptomid samples and for opening your home and lab to Emma and I, Stewart Anderson for providing exceptional directions to all Alberta populations, Dr. Suarez-Morales for *H. morelenses*, John Salinas and Allan Vogel for Lake Waldo samples, Barbara Taylor for *H. augustaensis* samples, Blake Matthews for British Columbia populations, Brian Parker for *H. arcticus* samples, Brenda Dobson from Parks Canada for help with the Parks Canada permit, Cameron Goater for Waterton Lakes samples, Chris Darimont for British Columbia coastal island samples, Daniel Schindler for *H. kenai* samples, Dr. Pfrender for California samples, Melanie McFarland of the Lassen National Forest for learning how to sample with a plankton net, Janet Reid for all the taxonomy help.

I would also like to thank all the great minds of the Rocky Mountain Biological Laboratory who willingly devoted their time and knowledge especially Dr. Scott Wissinger and Dr. Bobbi Peckarsky who are an inspiration to all students that encounter them. RMBL is a remarkable research laboratory and I feel very privileged to have been able to learn and conduct my research there. I am thankful to the Colorado Nature Conservancy for the chance to study an unspoiled habitat, The Mexican Cut Nature Preserve.

I would like to thank Concordia University, Natural Sciences and Engineering Research Council of Canada, and the Snyder Memorial Fellowship of RMBL for their generous financial support that made this study possible.

I thank my fellow undergraduate and graduate students at the Forest and Evolutionary Genomics Lab and the aquatic ecology lab.

Finally, I extend a warm and special acknowledgment to my parents and my husband Trevor Lane, who continuously encouraged me and kept me sane throughout this process.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1: Phylogeny of the genus *Hesperodiaptomus* (Copepoda) based on nucleotide sequence data of the nuclear ribosomal gene.

Chapter 2: Phylogeography of *Hesperodiaptomus shoshone.*

Chapter 4: Genetic structure of *Hesperodiaptomus shoshone* (Copepoda) in aestival ponds in the Rocky Mountains of Colorado based on nuclear microsatellite data.

LIST OF TABLES

Chapter 1: Phylogeny of the genus *Hesperodiaptomus* (Copepoda) based on nucleotide sequence data of the nuclear ribosomal gene.

Chapter 2: Phylogeography of *Hesperodiaptomus shoshone*

Chapter 3: Isolation and characterization of microsatellite markers for *Hesperodiaptomus shoshone.*

Chapter 4: Genetic structure of *Hesperodiaptomus shoshone* (Copepoda) in aestival ponds in the Rocky Mountains of Colorado based on nuclear microsatellite data.

General Introduction

My doctoral research project examined the molecular evolution of the North American freshwater diaptomid copepod, *Hesperodiaptomus shoshone* (S.A. Forbes, 1882) to gain insights into the biological diversity of this species in high altitudinal mountainous areas. Knowledge of the partitioning of genetic diversity at various hierarchical levels ranging from individual populations to species is invaluable for understanding the micro- and macro-evolutionary forces influencing the evolution of biological diversity.

Copepods are the most important primary consumers in aquatic communities playing an important role in virtually all pelagic food chains (Huys and Boxshall, 1991). There are currently 11 500 described taxa that may represent only 15 % of the total number of species (Humes, 1994). This group of arthropods is also spectacularly abundant as an estimated $1.37x10^{21}$ planktonic copepods inhabit the pelagic realm (Boxshall and Halsey, 2004). Needless to say, copepods are considered to be the most abundant metazoans on earth. Additionally, they are morphologically diverse and include numerous groundwater, benthic, symbiotic and parasitic forms. During their long evolutionary history, starting in the Lower Cretaceous, copepods have successfully colonized virtually all aquatic habitats with an extensive range of salinity regimes; from freshwater to marine and hypersaline inland waters, and all temperature regimes; from polar waters to hydrothermal vents (Huys and Boxshall, 1991). Furthermore, they have an incredible vertical range occurring from depths of 10 000 meters in the Philippine Trench to an altitude of 5500 meters in the Himalayan mountains (Boxshall and Halsey, 2004), a distance which constitutes about three quarters of the maximum vertical range of the earth's surface. This incredible diversity in habitat use and

diverse body forms make copepods an interesting group of organisms for extensive evolutionary examinations.

The history of the study of copepods has been divided into two phases. The early phase of the 19th century consisted of the gradual recognition of the unity of copepods. The diverse body forms and modes of life caused great confusion in the early classification systems. The second phase started in the mid 19th century and which still continues today focuses on the phylogenetic and evolutionary relationships within the Copepoda (Damkaer, 2002). Historically, many different classification systems have been proposed with numerous revisions in the past ten years. Currently, two infraclasses and nine orders are recognized (Dussart and Defaye, 2001). The description of these is beyond the scope of this thesis but a detailed account can be found in the *Evolution of Copepods* by Huys and Boxshall (1991) and in the *Copepodologist's Cabinet* by Damkaer (2002).

Since the 1960's more than 90 new families of copepods have been proposed (Boxshall and Halsey, 2004). Half of these have been classified as new discoveries and the remaining families were recognised on the basis of revisionary systematic studies. These findings demonstrate the dynamic nature of copepod systematics and the continuing need for strengthening taxonomic reference collections and field work. In spite of the increasing knowledge of this group the evolutionary relationships are still under discussion (Lee, 2000). The difficulties in reconstructing phylogenies are due to a poor fossil record, problems in the selection of suitable characters for phylogenetic analyses and difficulty in species identification.

Three orders of freshwater copepods commonly occur in North America: the Harpacticoida, Cyclopoida and Calanoida. Freshwater Calanoida in North America are

mainly represented by a single family, the Diaptomidae, whose species are examined in this thesis. The Diaptomidae comprises in excess of 500 valid species on most continents and is composed of four subfamilies; Paradiaptominae, Microdiaptominae, Diaptominae, and Speodiaptominae (Boxshall and Halsey, 2004). Although, members of the Diaptomidae are primarily small particle feeders, some are predatory and cannibalistic. Most are planktonic, but some are benthic and a few inhabit subterranean waters (Boxshall and Jaume, 2000). It is not known if the Diaptomidae radiated from a single colonization event or multiple colonization events of the freshwater environment, casting doubts on the monophyletic origin of this group. This uncertainty creates problems in identifying appropriate outgroup taxa for phylogenetic analysis.

In the Americas, diaptomid copepods tend to have restricted distribution patterns and contain many endemic forms (Dussart and Defaye, 2001). There are 182 species, of which 79 (43%) are known to occur in Canada and the U.S.A., 24 (13%) in Mexico and Central America, 6 (3%) in the Antilles, and 84 (47%) in South America (Suárez-Morales *et. al.*, 2005). No species are shared between USA/Canada and South America. In North America, most of the species consist of endemic genera; *Aglaodiaptomus*, *Onychodiaptomus*, *Skistodiaptomus* and two almost endemic genera; *Hesperodiaptomus* and *Leptodiaptomus*. There are four other genera shared with Eurasia; *Acanthodiaptomus*, *Diaptomus*, *Mixodiaptomus*, *Sinodiaptomus* and *Eudiaptomus*, one with Northeastern Asia; *Nordodiaptomus* and one holarctic genus with populations in Africa: *Arctodiaptomus*. The genus *Microdiaptomus* is represented by a single cave dwelling species, *Microdiaptomus cokeri* (Osorio-Tafall, 1942) in Mexico (Elias-Gutierrez and Suárez-Morales, 1998).

Surprisingly few studies have investigated the phylogeny of Diaptomidae, considering how important this family is in freshwaters. No studies have been done to support the subfamilial structure and virtually nothing within the largest subfamily, the Diaptominae. The lack of phylogenetic studies might be a consequence of the inaccessibility and scattered nature of the primary taxonomic literature and the ambiguous morphological relationships that have hindered progress in copepod systematics to date, especially at the family or genus level (Huys and Boxshall, 1991). In recent years, molecular systematics has provided complementary and informative data for inferring the evolutionary relationships among morphologically similar species (Braga *et. al,* 1999). However, very few studies exist at the population, species or genus level for North American diaptomids. Thum (2004) provided the first molecular phylogenetic reconstruction of the North American genera but did not have enough sampled species and phylogenetic resolution to address intrageneric relationships. Boileau and Hebert (1988) examined the population genetic structure of *Hesperodiaptomus arcticus* (Marsh, 1920) using both morphological characters and allozyme markers at Canadian Arctic sites. The study showed high molecular divergence and morphological variability and therefore suggested subdividing *H. arcticus* into three distinct species casting doubt on the cohesiveness of other species in the genus *Hesperodiaptomus*.

The genus *Hesperodiaptomus* has long been considered a taxonomically problematic group, the uncertainty in species delimitation and the presence of morphologically similar forms suggests the possibility of cryptic speciation and warrants a thorough revision using a comprehensive set of morphological and molecular data (Scanlin and Reid, 1996; Granados-Ramírez and Suárez-Morales, 2003). Thus, in this thesis I first examined the phylogenetic

and evolutionary relationships within the genus *Hesperodiaptomus*, and then focused on the evolutionary relationships and population genetic structure of one species, *H. shoshone.*

Hesperodiaptomus shoshone is a high altitude copepod restricted to lentic habitats of the Rocky Mountains and Sierra Nevada. It has been misidentified and confused with *Hesperodiaptomus kenai* (M.S. Wilson, 1953) on multiple occasions and is suspected to consist of cryptic species. The occurrence of *H. shoshone* in isolated high altitude areas provides also a unique opportunity to examine the genetic structure and processes that drive speciation and biodiversity in these communities.

Molecular ecology and copepods

Molecular ecology is a recent discipline that emerged in the mid 80's and includes aspects of molecular biology, ecology, evolution, behavioural ecology and genetics (Beebee and Rowe, 2004). The basis of molecular ecology is understanding how genetic variability is partitioned at different geographical and temporal scales and has important implications in evolutionary biology, ecology and conservation biology.

Populations of most species show some level of genetic structuring which may be due to a number of non-mutually exclusive processes. Genetic drift, gene flow, natural selection, historical processes, predation pressure and life history traits may all shape the genetic structure of populations that has important consequences for both contemporary and long term ecological processes and potential speciation. Genetic structuring also reflects the number of genes exchanged between populations. Low gene flow in populations will affect their evolutionary potential by either promoting local adaptation and ultimately speciation or reducing the fitness by fixation of deleterious mutations which can lead to extinction (Higgins

and Lynch, 2001). On the other hand, high gene flow can prevent local adaptation or fixation of alleles and will slow down the process of speciation (Barton and Hewitt, 1985).

Lakes and ponds represent isolated fragments of aquatic habitat highly appropriate for investigations in evolutionary ecology and genetic variation (Bohonak, 1999). Freshwater invertebrates such as zooplankton provide an ideal system to study population subdivision and speciation. The habitats they occupy are discrete, they occur in large numbers and are easy to sample for genetic analysis. Scientists have speculated as to how these organisms disperse over small and large scale distances, but due to the problems associated with studying such small, short lived organisms, little empirical studies have been completed. Most of our understanding of zooplankton and copepod dispersal is based on anecdotal evidence and broad generalizations (Bohonak and Jenkins, 2003). Zooplankton are believed to rely mostly on passive dispersal by producing resting stages that can survive unsuitable habitat conditions and can be dispersed in a limited fashion using water (Havel *et al.,* 2000; Michels *et al.,* 2001), wind (Jenkins and Underwood, 1998), or other animals as vectors (Bohonak and Whiteman, 1999; Green *et al.,* 2002). Although these studies report means by which zooplankton may disperse, quantifying dispersal is very difficult due to the small size of and large number of dispersing propagules. However, recent developments in molecular ecology provide us with an unprecedented power to detect current and historical dispersal patterns of many aquatic organisms including copepods by examining the genetic structure and amount of gene flow that occurs among populations.

Molecular Markers

Molecular markers have the potential to address questions in ecology and evolution that are difficult to pursue in any other way and are the foundation of molecular ecology. Molecular markers are small sections of the genome (polymorphic proteins or DNA sequences) that can be used as indicators of genome wide variation (Beebee and Rowe, 2004). Not all sections of the genome are equally useful as molecular markers since the level of polymorphism varies depending on the section of the genome. The isolation and characterisation of a wide variety of molecular markers with differing levels of polymorphism appropriate to specific questions of research are of central importance in molecular ecology. Highly polymorphic markers such as microsatellites are preferred in behavioural studies where individual organisms need to be recognised. Population genetic analyses may be more powerful with moderately polymorphic markers while evolutionary relationships among distantly related taxa may be best resolved with more conserved markers. Pros and cons of various markers are discussed by Sunnucks (2000). Three types of markers were used in this thesis to address questions at the genus and species level. They include: nuclear ribosomal DNA sequences, cytochrome oxidase I mitochondrial sequences and microsatellites.

Nuclear rDNA

Nuclear rDNA is a multigene family with copies of RNA genes repeated tandemly multiple times. Ribosomal RNAs are fundamental constituents of ribosomes, which translate mRNA sequences into proteins. Each rRNA gene array, found in the nuclear genome of eukaryotes, is composed of genes coding for three types of rRNAs: 18S, known as the small subunit; 28S, the large subunit and the 5.8S rRNA (Page and Holmes, 1998). These coding

regions are separated from each other by spacers; an external transcribed spacer (ETS) and two internal transcribed spacers (ITS1 and ITS2). Each array of ribosomal genes is transcribed as a single unit which is processed to generate functional rRNAs. These functional rRNA sequences are highly conserved among taxonomic groups. The nuclear rDNA spacer regions evolve more rapidly than the coding regions due to a lack of constraints on mutation of these areas. For this reason, nuclear rRNA gene arrays have been used extensively in molecular systematics to reconstruct relationships between both very distant rDNA sequences and closely related species (Page and Holmes, 1998).

Mitochondrial DNA

Mitochondrial DNA (mtDNA) sequences have been used extensively as genetic markers for molecular systematic and population genetic studies (Avise, 1994; Moritz, 1994). They have several advantageous properties over nuclear genetic markers (Avise, 1994). The mtDNA genome of animals is usually inherited in a uniparental fashion and it has a relatively rapid rate of molecular evolution (Beebee and Rowe, 2004). These features usually lead to non-reticulating (bifurcating) gene trees, rapid geographical sorting and genetic divergence of populations in the absence of gene flow. As a result, phylogeographic studies using mtDNA usually provide more resolution of intraspecific patterns of geographical variation than other classes of molecular markers (Avise, 2000).

Different regions of the mitochondrial genome evolve at different rates, thus certain regions of the mtDNA have been targeted for specific types of studies. MtDNA encodes some of the proteins necessary for oxidative phosphorylation which reside in the

inner mitochondrial membrane. MtDNA also encodes rRNA and tRNA molecules so it can carry out its own protein synthesis. For example, the cytochrome b and control region (CR) have been used for species level or population level phylogeny (Harris *et al.*, 2000; Caudill and Bucklin, 2004) and more conserved regions; 12S, 16S, and COI have been used for molecular systematic studies at species, genus, and family level (Guarnieri, 1996; Lee, 2000).

Microsatellites

Microsatellites, Simple Sequence Repeats (SSRs) (Tautz, 1989) or Short Tandem Repeats (STR's) (Edwards *et al.*, 1991) are sections of DNA composed of repeats of 2-7 nucleotides arranged in a tandem fashion. They have been found in every organism studied so far and are generally assumed to be evenly distributed over genomes (Dietrich *et al.*, 1996) but rarely occur within coding regions (Hancock, 1995). The sequences surrounding the repeat region are usually conserved, allowing PCR primers to be designed so that the repeat region and a short flanking sequence can be amplified. Individuals may differ in the number of repeats present, meaning that the length of the PCR product varies. The size of the PCR product can be scored using electrophoresis and various forms of detection methods.

The high level of polymorphism in microsatellites is attributed to relatively high rates of error during DNA replication (slippage) (Harr *et. al.*, 2000) and recombination (unequal crossover). During replication of a repetitive region, DNA strands may dissociate, and reassociate incorrectly. The replication of the misaligned regions can lead to insertion or deletion of repeat units, thus altering the size of the microsatellite region. The number of repeats in a given microsatellite is known to affect the mutation rate such that longer repeats

are more polymorphic than shorter ones (Primmer *et al.*, 1998). This is possibly due to the increased possibility for a stable misaligned configuration in long stretches of repeats. The second factor that influences microsatellite stability is the interruptions of the repeat. Interrupted microsatellites with insertions or point mutations within the repeat region are known to have lower mutation rates than perfect repeats. This is due to hindrance to formation of slipped intermediates in the interrupted repeats (Petes *et al.*, 1997).

Thesis outline and brief chapter overview

My thesis is composed of four chapters and has a general goal of investigating the molecular ecology and evolution of North American hesperodiaptomid copepods to obtain an improved understanding of processes affecting population dynamics, dispersal and speciation events.

Chapter 1: Phylogeny of the genus *Hesperodiaptomus* (Copepoda) based on nucleotide sequence data of the nuclear ribosomal gene.

In chapter 1, I present the results of a study based upon nuclear ribosomal DNA gene sequence data to address systematic relationships among the genus *Hesperodiaptomus*. As mentioned above, few studies have investigated the Diaptomidae phylogeny, and no studies have examined species relationships within North American diaptomid genera. The present study provides an initial phylogenetic reconstruction of this genus.

Chapter 2: Phylogeography of *Hesperodiaptomus shoshone.*

In chapter 2, I used both nuclear ribosomal DNA sequences and mitochondrial Cytochrome Oxidase I sequences to determine the patterns of genetic differentiation in *H. shoshone* to assess the influence of historical events on dispersal, vicariance and diversification throughout the distribution range of this species in western North America. An understanding of the historical processes on the current distribution patterns provides a means to assess the evolutionary potential of the species and to understand the speciation processes prevalent in this group of organisms.

Chapter 3: Isolation and characterization of microsatellite markers for *Hesperodiaptomus shoshone.*

Highly polymorphic genetic markers such as microsatellites are invaluable for studying fine-scale genetic differentiation and recent dispersal events (Sunnucks, 2000). To date, no microsatellites have been isolated and characterized for any of the North American diaptomid copepods. In this chapter, I present the results of five microsatellite loci that I isolated and characterised for the high altitude diaptomid copepod *H. shoshone*. These loci are useful for investigating genetic variation within and among natural populations of *H. shoshone*. I used these markers in the study described in chapter 4 of this thesis.

Chapter 4: Genetic structure of *Hesperodiaptomus shoshone* (Copepoda) in aestival ponds in the Rocky Mountains of Colorado based on nuclear microsatellite data.

At fine spatial scales, the study of genetic population differentiation provides a unique way of determining how microevolutionary forces such as random drift, gene flow, natural

selection and mutation shape speciation. The overall objective of chapter 4 was to examine the genetic diversity at a local scale of a high altitude freshwater diaptomid copepod *H. shoshone* within and among neighbouring watersheds of the Elk Mountains of Colorado. Understanding the genetic diversity between neighbouring habitats has important consequences for understanding how they will recover from disturbances and provides clues on the evolution of biological diversity in high altitude ecosystems.

References

Avise, J.C. (1994) Molecular markers, natural history and evolution. Chapman and Hall, New York.

Avise, J.C. (2000) Phylogeography : The history and formation of species. Harvard University Press, Cambridge, Massachusetts.

Barton, N.H. and G.M. Hewitt (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16**:113–148.

Beebee, T.J.C., and G. Rowe (2004) An introduction to Molecular Ecology. Oxford University Press, Oxford. 436pp.

Bohonak, A.J. (1999) Dispersal, gene flow, and population structure. *The Quaterly Review of Biology* **74**(1):21-45.

Bohonak, A. and D. Jenkins (2003) Ecological and evolutionary significance of dispersal by freshwater invertebrates. *Ecology Letters* **6**:783-796.

Bohonak, A.J. and H.H. Whiteman (1999) Dispersal of the fairy shrimp *Branchinecta coloradensis* (Anostraca): Effects of hydroperiod and salamanders. *Limnology and Oceanography* **44**:487-493.

Boileau, M.G. and P.D.N. Hebert (1988) Genetic differentiation of freshwater pond copepods at arctic sites. *Hydrobiologia* **167/168**:393-400.

Boxshall, G.A. and D. Jaume (2000) Making waves: The repeated colonisation of fresh water by copepod crustaceans. *Advances in Ecological Research* **31**:61-79.

Boxshall, G.A. and S.H. Halsey (2004) An introduction to copepod diversity. The Ray Society, Andover.

Braga, E., R. Zardoya, A. Meyer and J. Yen (1999) Mitochrondrial and nuclear rRNA based copepod phylogeny with emphasis on the *Euchaetidae* (calanoida). *Marine Biology* **133**:79-90.

Caudill, C.C., A. Bucklin (2004) Molecular phylogeography and evolutionary history of the estuarine copepod, *Acartia tonsa*, on the Nortwest Atlantic. *Hydrobiologia* **511**:91-102.

Damkaer, D.M. (2002) The copepodologist's cabinet: a biographical and bibliographical history. American Philosophical Society, Philadelphia 300pp.

Dietrich, W.F., J. Miller, R. Steen, M.A. Mechant, D. Damron-Boles, Z. Husain and *et. al.* (1996) A comprehensive genetic map of the mouse genome. *Nature* **380**:152-154.

Dussart, B.H. and D. Defaye (2001) Introduction to the Copepoda. 2nd Ed. Guides to the Identification of the Microinvertebrates of the Continental Waters of the World. Backhuys Publishers, Leiden. 344 pp.

Edwards, A., A. Civitello, H.A. Hammond and C.T. Caskey (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American Journal of Human Genetics* **49**:749-756.

Elías-Gutiérrez, E. and E. Suárez-Morales (1998) Redescription of *Microdiaptomus cokei* (crustacea: Copepoda: Diaptomidae) from caves in central Mexico, with the description of a new diaptomid subfamily. *Proceedings of the Biological Society of Washington* **111**(1):199-208.

Green, A.J., J. Figuerola and M.I. Sánchez (2002) Implications of waterbird ecology for the dispersal of aquatic organisms. *Acta Oecologica* **23**:177-189.

Granados-Ramírez, J.G. and E. Suárez-Morales (2003) A new *Hesperodiaptomus* light (Copepoda, Calanoida, Diaptomidae) from Mexico with comments on the distribution of the genus. *Journal of Plankton Research* **25**(11):1383-1395.

Guarnieri, M. (1996) A preliminary study of genetic variation within and among populations of *Diaptomus leptopus* (Copepoda:Calanoida): A DNA sequence analysis of the mitochondrial Cytochrome Oxidase I gene. Concordia University, Montreal 55pp.

Hancock, J.M. (1995) The contribution of slippage-like processes to genome evolution. *Journal of Molecular Evolution* **41**:1038-1047.

Harr, B., B. Zangerl and C. Schlotterer (2000) Removal of microsatellite interruptions by DNA replication slippage: Phylogenetic evidence from *Drosophila*. *Molecular Biology and Evolution* **17**(7):1001-1009.

Harris, D., D.S. Rogers, J. Sullivan (2000) Phylogeography of *Peromyscus furvus* (Rodentia; Muridae) based on cytochrome b sequence data. *Molecular Ecology* **9**(12):2129-2135.

Havel, J.E., E.M. Eisenbacher and A.A. Black (2000) Diversity of crustacean zooplankton in riparian wetlands: Colonization and egg banks. *Aquatic Ecology* **34**:63-76.

Higgins, K. and M. Lynch (2001) Metapopulation extinction caused by mutation accumulation. *Proceedings of the National Academy of Sciences of the USA* **98**:2928–2933.

Humes, A.G. (1994) How many copepods? *Hydrobiologia* **292/293**:1-7.

Huys, R. and G. A. Boxshall (1991) Copepod Evolution. The Ray Society, London.

Jenkins, D.G. and M.O. Underwood (1998) Zooplankton may not disperse readily in wind, rain, or waterfowl. *Hydrobiologia* **387/388**:15-21.

Lee, C.E. (2000) Global phylogeography of a cryptic copepod species complex and reproductive isolation between genetically proximate "populations". *Evolution* **54**(6):2014-2027.

Michels, E., K. Cottenie, L. Neys and L.D. Meester (2001) Zooplankton on the move: First results on the quantification of dispersal of zooplankton in a set of interconnected ponds. *Hydrobiologia* **442**:117-128.

Moritz, C. (1994) Applications of mitochondrial DNA analysis in conservation: A critical review. *Molecular Ecology* **3**:401-411.

Page, R.D.M. and E.C. Holmes (1998) Molecular evolution : a phylogenetic approach. Oxford ; Malden, MA, Blackwell Science.

Petes, T.D. and M. Dominska (1997) Stabilization of microsatellite sequences by variant repeats in the yeast *saccharomyces cerevisiae. Genetics* **146**:491-498.

Primmer, C.R., N. Saino, A.P. Moller and H. Ellegren (1998) Unraveling the processes of microsatellite evolution through analysis of germline mutations in barn swallows *hirundo rustica. Molecular Biology and Evolution* **15**:1047-1054.

Scanlin, M. and J.W. Reid (1996) A new copepod species from California, USA: *Hesperodiaptomus californiensis* (crustacea: Copepoda: Calanoida: Diaptomidae). *Proceedings of the Biological Society of Washington* **109**(1):103-111.

Suárez-Morales, E., J.W. Reid and M. Elias-Gutierrez (2005) Diversity and distributional patterns of neotropical freshwater copepods (Calanoida: Diaptomidae). *International Journal of Hydrobiology* **90**(1):71-83.

Sunnucks, P. (2000) Efficient genetic markers for population biology. *Trends in Ecology and Evolution* **15**(5):199-203.

Tautz, D. (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* **17**:6463-6471.

Thum, R.A. (2004) Using 18S rDNA to resolve diaptomid copepod (Copepoda: Calanoida: Diaptomidae) phylogeny: an example with the North American genera. *Hydrobiologia* **519**: 135–141.

Chapter 1

Phylogeny of the genus *Hesperodiaptomus* (Copepoda) based on nucleotide sequence data of the nuclear ribosomal gene.

Introduction

The members of the genus *Hesperodiaptomus* Light, 1938 include predatory and herbivorous species of copepods that occupy temporary and permanent freshwater habitats. Some species constitute an important component of alpine freshwater communities and raptorial *Hesperodiaptomus* are often considered to be keystone species in these environments (Anderson, 1980; Stoddard, 1987; Paul *et al.,* 1995). Most species of *Hesperodiaptomus* occur in north-western North America but some extend into eastern Canada, and two species have disjunct distributions in the USA and Mexico (Granados-Ramírez and Suárez-Morales, 2003). Although the distribution ranges of four species: *H. arcticus* (Marsh, 1920), *H. eiseni* (Lilljeborg, 1889), *H. shoshone* (S.A. Forbes, 1882) and *H. wardi* (Pearse, 1905) have been reported to extend into Siberia, only two species (*H. arcticus* and *H. eiseni*) have actually been recorded there (Borutskii *et al.,* 1991).

Hesperodiaptomus is the most diverse genus of the sub-family Diaptominae in North America. Although 18 species are generally recognised (Williamson and Reid, 2001; Granados-Ramírez and Suárez-Morales, 2003), it has long been considered as a taxonomically problematic group due to intraspecific morphological variability in males of some species (Schacht, 1897; Wilson, 1953; Boileau, 1991, Kincaid, 1953). On the other hand, females of several species such as *H. arcticus*, *H. breweri* (M.S. Wilson, 1958) and *H. kiseri* (Kincaid, 1953) are similar morphologically and cannot be reliably assigned to a particular species

using the existing data (Scanlin and Reid, 1996). Based upon morphological and molecular data Boileau (1989, 1991) suggested subdividing *H. arcticus* into three species, which would expand the genus into 20 species. Efforts to split the Siberian populations of *H. arcticus* into three additional species based upon minor differences in morphological characters by Streletskaya (1983) were rejected by Borutskii *et al.* (1991). The uncertainty in species delimitation and the presence of morphologically similar forms suggesting the possibility of cryptic species further warrants a thorough revision of the genus using a comprehensive set of morphological and molecular data (Scanlin and Reid, 1996; Granados-Ramírez and Suárez-Morales, 2003).

Species of the genus *Hesperodiaptomus* are characterized by features of the 5th legs, setation patterns of antennules, and features of somites. Detailed descriptions of these features can be found in Wilson and Yeatman (1959), Dussart and Defaye (2001), and Granados-Ramírez and Suárez-Morales (2003). Unfortunately, only a few of these morphological characters can be easily scored for phylogenetic analyses (Huys and Boxshall, 1991; Braga *et al.,* 1999; Bucklin *et al.,* 1999; Thum, 2004). However, in recent years, molecular systematics has provided complimentary and informative data for inferring evolutionary relationships among morphologically similar species (Braga *et al.,* 1999; Lee and Frost, 2002).

The phylogenetic position of *Hesperodiaptomus* within the sub-family Diaptominae remains poorly understood. To date, only two studies have examined the phylogenetic relationships of the sub-family Diaptominae (Light, 1939; Thum, 2004). Light (1939), recognized three distinct groups based on limited morphological data. He considered two groups endemic to North America, namely the *Hesperodiaptomus* group, which includes the

genera *Hesperodiaptomus, Aglaodiaptomus* Light, 1938, *Skistodiaptomus* Light, 1939 and *Onychodiaptomus* Light, 1939 and the monotypic *Leptodiaptomus* group, which included all species of the genus *Leptodiaptomus* Light, 1938. The remaining other group consists of the genera *Arctodiaptomus* Kiefer, 1932; *Mastigodiaptomus* Light, 1939; *Argyrodiaptomus* Brehm, 1933; *Notodiaptomus* Kiefer, 1936 and *Prionodiaptomus* Light, 1939. None of these genera are endemic to North America. Light (1939) speculated that the two endemic groups evolved from a Eurasian ancestor closely related to *Hesperodiaptomus*. A recent molecular phylogenetic study based upon the nuclear 18S ribosomal RNA gene sequence data (Thum, 2004) is generally congruent with Light's (1939) speculations regarding relationships among the North American diaptomid copepods, except for the phylogenetic position of the genus *Leptodiaptomus*. Contrary to Light (1939), who suggested a reciprocal monophyly of *Leptodiaptomus* and *Hesperodiaptomus*, Thum (2004) considers *Leptodiaptomus* to be a lineage derived from the *Hesperodiaptomus* group. However, the 18S rDNA sequences failed to resolve relationships within the genera of the North American Diaptominae because of the overall slow rate of molecular evolution in this gene which were reflected by low bootstrap support values. Faster evolving rDNA regions such as the two internally transcribed spacers (ITS1 and ITS2) may be preferred to resolve such relationships.

To date phylogenetic relationships among species of *Hesperodiaptomus* remain poorly known. Wilson (1953) organised these species into two groups based on the morphological features of the 5th leg of males; one group included species morphologically similar to *H. eiseni* and the other included species similar to *H. shoshone* (Table 1). The eiseni group was distinguished by a modification of the right basis, of regular arrangement of spinules on the distal pad of the second exopod, and the presence of a prominent inner lamellar expansion of

the right coxa. The characteristic features of the shoshone group were a lack of modification of the right basis, a small or absent inner lamellar expansion of the right coxa, a grouped arrangement of spinules of the distal pad of the left exopod and the absence of spinules on the 15^{th} segment of the right antennule. The spinule is not always present in the members of the eiseni group.

The objective of the present study was to reconstruct the phylogeny of the genus *Hesperodiaptomus* based on partial nuclear ribosomal DNA sequences including the ITS1, 5.8S and ITS2 regions (more information on this marker is presented in the general introduction of this thesis). The resulting information will be invaluable for the understanding of interspecific relationships and the evolutionary history with the genus *Hesperodiaptomus.*

Table 1

The classification of the genus *Hesperodiaptomus* into eiseni and shoshone groups based on the morphological data (Wilson, 1953). *Hesperodiaptomus franciscanus* is morphologically distinct from these two groups and therefore placed for the purpose of this thesis in a separate group franciscanus.

Group	Species
eiseni group	*H. augustaensis*
	H. arcticus
	H. breweri
	H. californiensis
	H. eiseni
	H. kiseri
	H. morelensis
	H. nevadensis
	H. schefferi
	H. victoriaensis
	H. wardi
	H. wilsonae
shoshone group	*H. caducus*
	H. hirsutus
	H. kenai
	H. novemdecimus
	H. shoshone
franciscanus group	*H. franciscanus*

Methods

Collection and identification of organisms

Samples used in the present study were either collected during the summers of 2003 and 2004 using a Wisconsin plankton tow net (120 μm mesh size) or acquired from existing collections (Table 2). The newly collected samples were immediately preserved in 95% non-denatured ethanol following the procedure of Bucklin (2000). Species level identifications were completed using keys by Wilson and Yeatman (1959) and Sandercock and Scudder (1996).

Table 2

List of taxa, collection location and collector of species examined in this study.

Species	Populations used in analyses	Collector, year
H. arcticus	Middle Devon Lake, Alberta	Parker, 2004
H. californiensis	Pond #41, Lassen County, California	McFarland, 2005
H. eiseni	Caribou Wilderness lake, California	Pfender, 2002
H. franciscanus	Old Wolf Lake, British Columbia	Matthews, 2004
H. hirsutus	TNC- ADP #9, Agate Lake, Oregon	Dodson, 2005
H. kenai	Butchart Lake, British Columbia	Matthews, 2004
	Waldo Lake, Oregon	Salinas, 2003
	Shirley Lake, British Columbia	Schindler, 2005
H. nevadensis	Peninsula Lake, Alberta	Dayanandan, 2005
H. shoshone	Niwot Ridge, Colorado	Marszalek, 2003
	Middle Smiley pond, Idaho	Marszalek, 2004
	Wahoo Lake 2, California	Kramer, 2004
H. victoriaensis	Plateau Pond, Alberta	Marszalek, 2004

DNA extraction and amplification

DNA from individual copepods was extracted following a modified protocol of Lee and Frost (2002) and Bucklin (2000) for small tissue samples. The copepods, preserved in ethanol or formalin, were soaked at 4°C overnight in 1.5 ml centrifuge tubes containing 400 µl of TE (pH 8.5). On the following day, the TE was removed and 100 µl of lysis buffer (10 mm Tris pH 8.3, 50 mm KCl, 0.5% Tween 20, and 200 μg/ml Proteinase K) was added then the copepods were homogenized with a heat sealed pipet tip. The homogenates were incubated at 65°C for 1 hr. After incubation, 100 µl of Chloroform / Isoamyl alcohol (24:1) was added, the tube was gently inverted for 5 min and centrifuged at 7000 rpm for 5 min. The upper phase was transferred to a clean 1.5 ml centrifuge tube containing 100 µl of Isopropanol. The tubes were gently inverted and the DNA was allowed to precipitate at room temperature for 45 min. The tubes were centrifuged at 14 000 rpm for 10 min and the supernatant was decanted. This procedure was followed by an additional 70% ethanol wash and the pellet was dried at room temperature overnight. The dried pellet was dissolved in 150 µl of TE with 1 µl of RNase and left for 4 hrs at 4°C before PCR reactions were performed.

A 770 bp fragment of the nuclear ribosomal gene family was amplified using universal primers 5F and 4R as given by Beccerra and Venable (1999).

5F 5' – GCA AGT AAA AGT CGT AAC AAG G –3'

4R 5' – TCC TCC GCT TAT TGA TAT GC –3'

The amplified product included the variable ITS1 and ITS2 regions along with the more conserved 5.8S gene. A DNA sequence from one individual from each population was used in the present study. The DNA amplification was performed in 25 µl reactions consisting of 2.5 µl of buffer (0.2 M Tris pH 9.5; 0.25 M KCl; 1mg/ml BSA, 5µl/ml Tween

20), 2.5 mM $MgCl_2$, 250 µM dNTP, 0.4 pmol of each primer, 1.5 units of *Taq* DNA Polymerase, and 5 µl template DNA. The PCR amplifications were performed in an Applied Biosystems thermal cycler with the following temperature profile: 2 min at 94°C (1 cycle), 1 min at 94°C, 30 sec at 50°C, 2 min at 72°C (35 cycles) with a final extension step of 4 min at 72°C. A negative control was included in each PCR run to rule out any contaminations. The PCR products were separated by electrophoresis on a 1% agarose gel containing 0.33 µg/mL ethidium bromide in 1x TBE electrode buffer at 90V for 35 min. The DNA fragments on the gel were visualized with an ultraviolet (302 nm) transilluminator (UVP Inc). A GeneRuler 100-bp DNA ladder (Fermentas) was used for estimating the size of amplified DNA fragments. The PCR products were sequenced in both the forward and reverse directions at the Genome Québec Innovation Centre (Montreal, QC).

Sequence alignment and phylogenetic analysis.

Chromatograms were first imported and processed for quality with the PreGap4 program of the Staden Package (Staden, 1996). The consensus sequence was obtained using the shotgun assembly option of the Gap4 program of the Staden Package.

The DNA sequences were aligned using Clustal X version 1.8.3 software (Thompson *et al.*, 1997), followed by manual refinement with the MacClade software (Maddison and Maddison, 1992). To estimate the phylogenetic information content of the aligned rDNA sequence matrix, the g_1-test (Huelsenbeck, 1991; Hillis and Huelsenbeck, 1992) was computed to assess the skewness of the length distribution of 100 000 random trees generated using PAUP* (Swofford, 2001). The statistical significance of the phylogenetic structure was assessed using the values given in Hillis and Huelsenbeck (1992).

Phylogenetic trees were reconstructed using three different methods, namely i) maximum parsimony (MP), ii) maximum likelihood (ML), and iii) the distance based neighbour joining (NJ). The maximum likelihood (ML) analysis was conducted based on the probability model (GTI +G+I) obtained from MODELTEST 3.0 (Posada and Crandall, 1998) employing a heuristic search, the TBR branch swapping algorithm, and stepwise addition starting tree. Neighbour joining analysis (NJ) was constructed using the evolutionary model obtained from MODELTEST 3.0 (Posada and Crandall, 1998). In the maximum parsimony analysis, gaps were treated as missing data, all characters were weighted equally, and the exhaustive search algorithm was used. All phylogenetic analyses were performed using PAUP* 4.0 software (Swofford, 2001) on a Macintosh computer. Robustness of the branches was assessed using bootstrap analyses with TBR branch swapping and 1 000 bootstrap replicates. The rDNA sequences of *Aglaodiaptomus leptopus* (S.A. Forbes, 1882) (AY275451) and *Mastigodiaptomus nesus* Bowman, 1986 (AY275452) obtained from Genbank were used as outgroups in the analyses.

Results

The aligned nuclear rDNA sequence data matrix resulted in 114 variable sites and 69 parsimony informative sites (Appendix 1). The length of ITS 1 of the species of *Hesperodiaptomus* included in this study ranged between 314 and 316 bp. The length of the ITS 1 region of the outgroups, *A. leptopus* and *M. nesus* was 318 and 314 bp respectively. Insertions and deletions of 1 to 5bp were found in the ITS 1 region. The length of ITS 2 was 195 bp for all species of *Hesperodiaptomus* except for *H. franciscanus* (Lilljeborg, 1889) (196 bp). The length of ITS 2 of the outgroups was 191bp for *A. leptopus* and 195 for *M. nesus*. Insertions and deletions up to 5bp were found in the ITS 2 region of the two outgroup taxa. The polymorphism of nuclear ribosomal DNA sequences varied among the ITS1 and ITS2 regions and the 5.8S gene. Percent sequence divergence of the rDNA sequences of *Hesperodiaptomus* species was highest in ITS1 region and ranged from 1% to 11.8% followed by the ITS 2 region with divergences of 0.6% to 6.6%. No variable sites were found within sequences of the 5.8S gene.

The frequency distribution of the lengths of 10 000 random trees was significantly skewed (mean =643.09; std. dev.= 19.74; g_1=-2.567; $p<0.01$), indicating the prevalence of a phylogenetic signal over noise. The highest likelihood (-lnL = 2740.9412) model selected for the nuclear rDNA dataset based on the MODELTEST version 3.0 (Posada and Crandall, 1998) was a General Time Reversible model (GTR) of a nucleotide substitution with estimated base frequencies (A= 0.2118; C= 0.2564; G= 0.2862; T= 0.2455), a gamma distribution correction for variable sites (G = 0.7154) and the proportion of invariable sites (I = 0.4441).

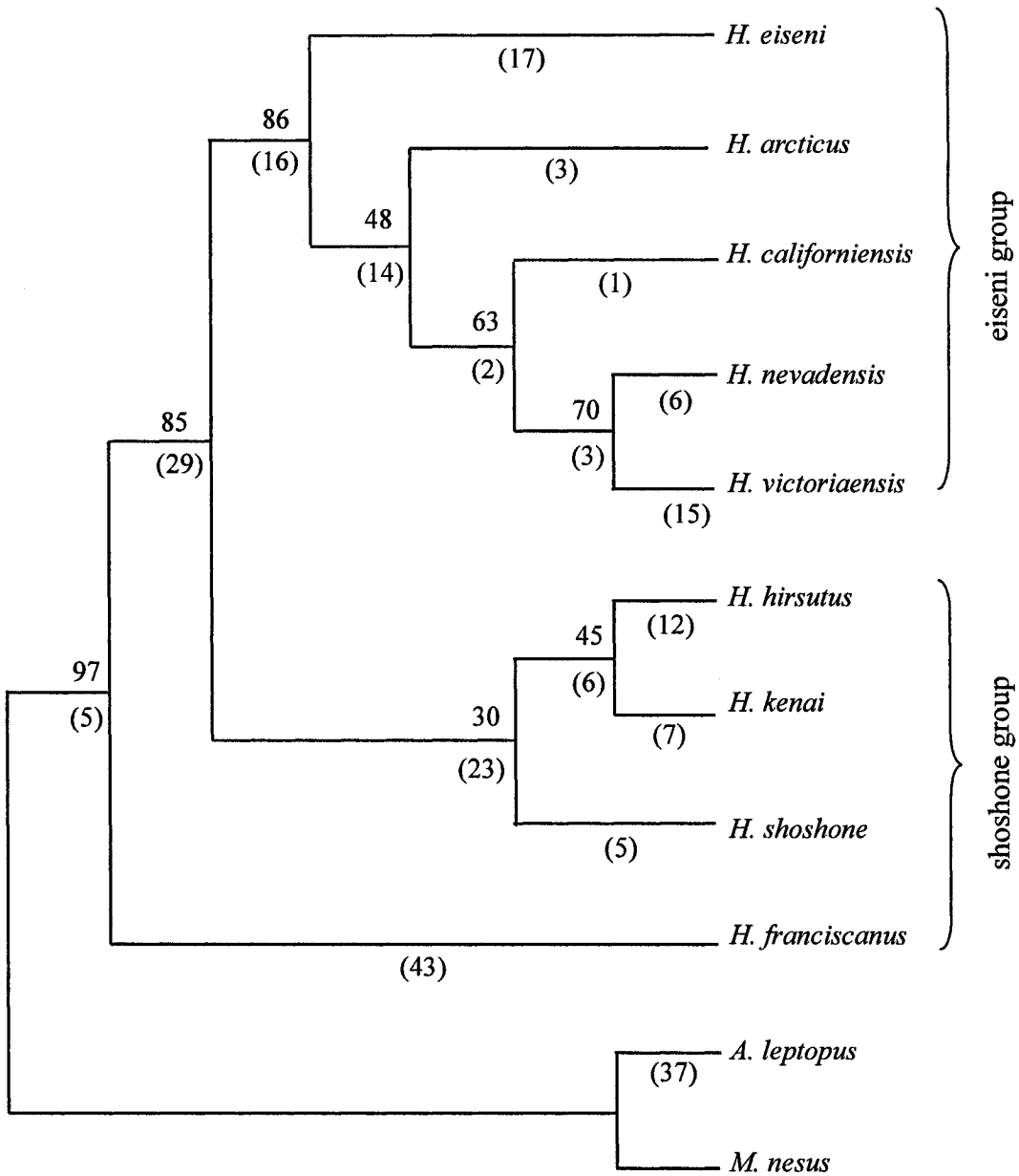
Phylogenetic analysis

The maximum parsimony analysis resulted in a single most parsimonious tree (length = 279; CI =0.774; RI= 0.580) given in Figure 1a. All species of *Hesperodiaptomus* formed a monophyletic group with bootstrap support of 97 and two main lineages. *Hesperodiaptomus franciscanus* occupied a basal position. The remaining species formed two groups recognized as eiseni and shoshone clades. The eiseni clade was composed of *H. eiseni* followed by *H. arcticus*, and *H. californiensis* (Scanlin and Reid, 1996) and highly supported by bootstrap analysis (Figure 1 a). The taxa *H. nevadensis* (Light, 1938) and *H. victoriaensis* (Reed, 1958) were sister to each other and occupied a terminal position within the eiseni clade. Although not supported by high bootstrap values, the shoshone clade was composed of three species *H. shoshone*, *H. kenai* (M.S. Wilson, 1953) and *H. hirsutus* (M.S. Wilson, 1953).

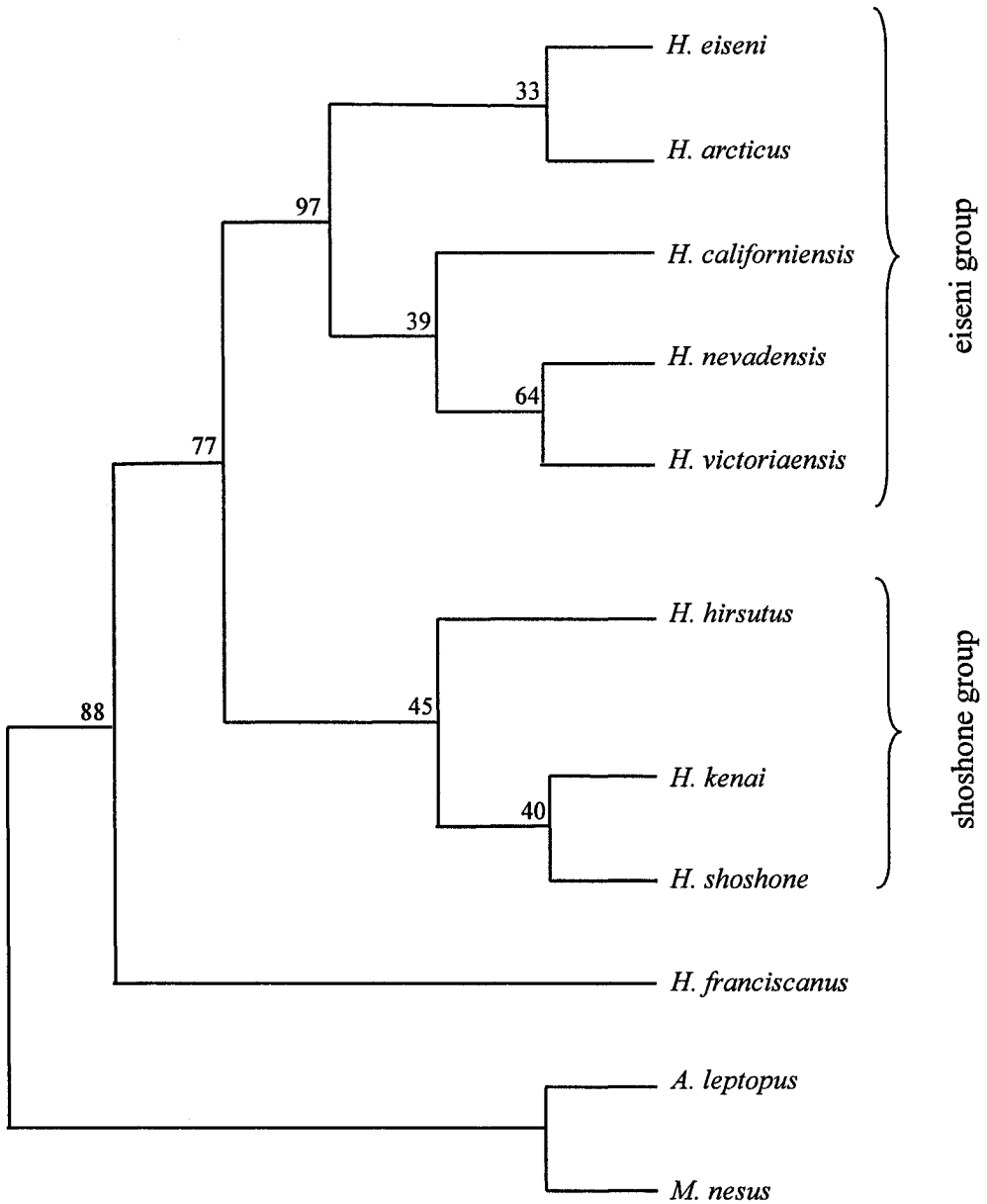
The phylogenetic tree based on neighbor joining (NJ) analysis (Figure 1b) was similar to the phylogenetic tree based on maximum parsimony method except for where *H. arcticus* and *H. eiseni* clustered together as well as the *H. kenai* and *H. shoshone* clustering. In the phylogenetic tree based on maximum likelihood (ML) analysis (Figure 1c), *H. franciscanus* occupied the basal position. However, species of the shoshone group showed a paraphyletic relationship, while the members of the eiseni group were monophyletic.

Figure 1
Phylogenetic reconstruction of the genus *Hesperodiaptomus* based on ITS1, 5.8S and ITS2 sequence data using three phylogenetic methods: a) neighbour joining, b) maximum parsimony, and c) maximum likelihood. Numbers next to nodes are bootstrap values based on 1000 bootstrap replicates. Numbers in the MP trees in parentheses represent the number of parsimonious changes for that branch.

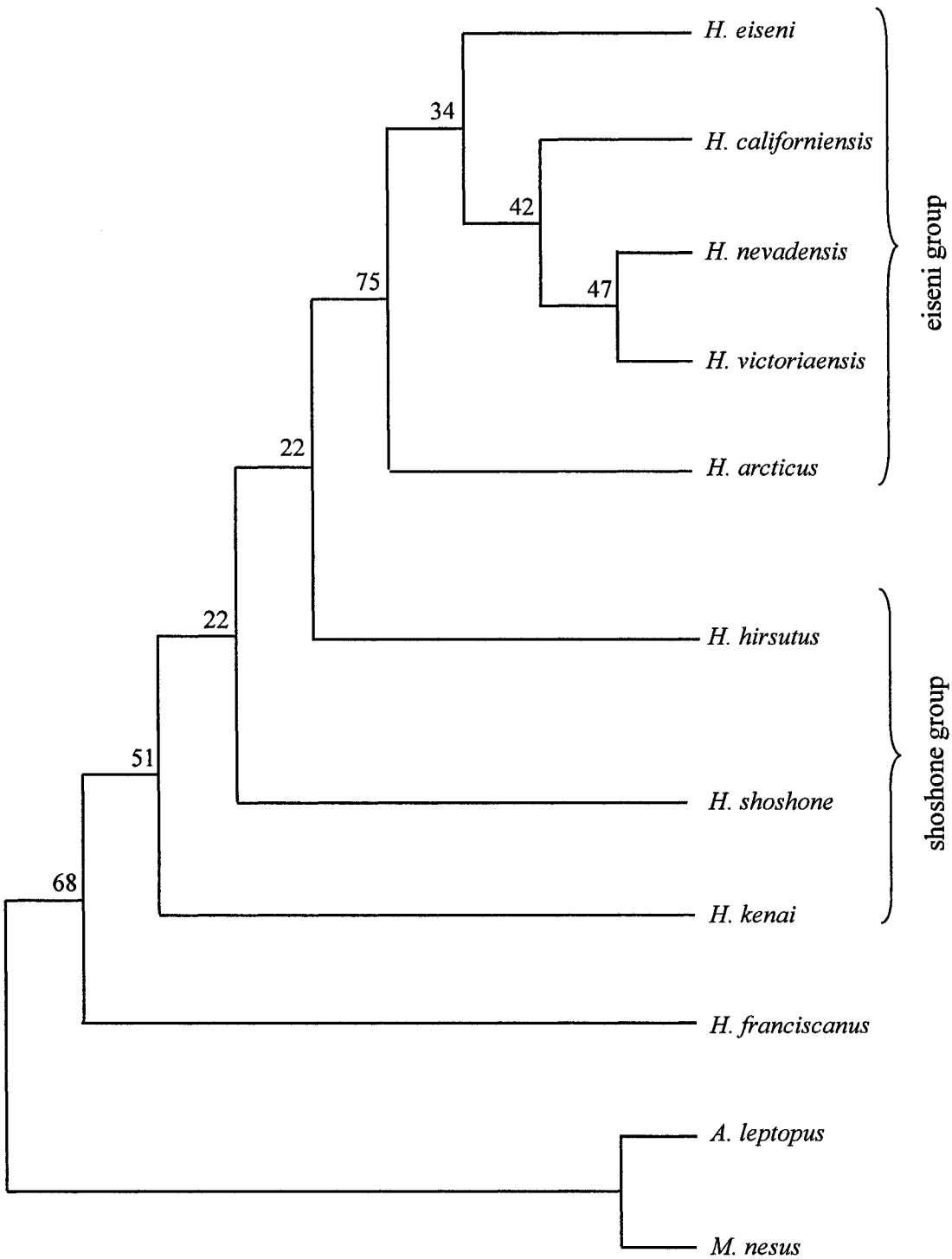
a) Maximum Parsimony



b) Neighbour Joining



c) Maximum Likelihood



Discussion

The nuclear rDNA phylogeny presented here contributes to the understanding of the phylogenetic relationships among the North American hesperodiaptomids. Contrary to earlier speculation, based on differences in maxilliped structures and feeding habits (Anderson, 1967) the monophyly of the genus *Hesperodiaptomus* was supported by all three phylogenetic reconstruction methods.

Phylogenetic trees based on three analytical methods supported the basal position of *H. franciscanus*. This species was originally placed in the genus *Skistodiaptomus* by Marsh (1907) based upon the presence of a small spine near the inner margin of the second exopod of the fifth leg of males and the position of the lateral spine of the same segment. The presence of a process on the ante-penultimate segment of the male geniculate antennule, a characteristic of hesperodiaptomids not present in the *Skistodiaptomus* sp. puzzled Marsh but was apparently not sufficient to exclude *H. franciscanus* from this genus. Due to lack of appropriate equipment, Marsh was unable to examine the characters of the distal pad of the left male exopods which were later recognised as important characters in generic grouping.

The basal position of *H. franciscanus* is in agreement with the current geographical distribution pattern for the genus. *Hesperodiaptomus franciscanus* has a broad geographical distribution along the Pacific Coast from southeastern Alaska to California and east to Alberta (Wilson and Yeatman, 1959; Anderson, 1971). Based on the current diversity of the genus and the fact that at least 11 species are recorded in the British Columbia / Oregon area, Granados-Ramírez and Suárez-Morales (2003) speculated that species of *Hesperodiaptomus* could have radiated from areas of the North Pacific coast of North America 1) to Alaska and

north-eastern Asia, 2) to the south into California and 3) to the east along a high latitude fringe.

The remaining ingroup taxa were grouped into a single monophyletic lineage consisting of two clades: The grouping of the eiseni and shoshone clades is consistent with the morphological character-based classification (Wilson, 1953) of the genus *Hesperodiaptomus*.

The bootstrap support for the taxa within in the *eiseni* clade is low, except for the consistent grouping of *H. victoriaensis* with *H. nevadensis* (Figure 1). From the available morphological characters *H. victoriaensis* is morphologically more similar to *H. californiensis* (Scanlin and Reid, 1996) than to *H. nevadensis*. For example, the two species share many similarities in the morphology of the 5th leg of males. The right coxa is characterised by an inner lamellar expansion, the presence of a longitudinal groove on the posterior face of the right basis, and a protrusion on the inner margin of this segment. The male of *H. californiensis* differs from that of *H. victoriaensis* by the presence a longer endopod in the right 5th leg, and by having a larger inner process on the basis which is placed about midlength rather than at the proximal corner of this segment (Scanlin and Reid, 1996). On the other hand, among other differences, H. *nevadensis* is morphologically more variable between populations and differs from the two previous species by having the inner process of the right basis reduced to a small spinule.

The remaining two species in the clade, *H. eiseni* and *H. arcticus* are both characterised by an expansion in the male 5th leg basis and a lengthwise ridge on the mid-posterior face with a spinous point or denticle. Both species have broad geographic distributions and are morphologically variable, possibly reflecting repeated isolations at multiple refugia during North American glaciations. Such taxa must be considered

cautiously, as it is uncertain whether they exist as biologically and genetically cohesive units (Boileau, 1991). For example, *H. arcticus* is widespread in pond habitats from the Canadian Rockies (Anderson, 1971) as far east as Ungava (Reed, 1959) and from the prairies (Wilson, 1958) to the high arctic (Marsh, 1929) and Siberia (Borutskii *et al.,* 1991). Through allozyme studies, morphological studies and biogeographical analyses Boileau (1989) established that *H. arcticus* is composed of three mostly allopatric species, *H. arcticus* sensu stricto, *H. churchillensis* and *H. nearcticus* (the species used in this study). Boileau (1989) performed a UPGMA cluster analysis Using Nei's unbiased genetic identity at 20 enzyme loci for one *Leptodiaptomus* and four *Hesperodiaptomus* species from 13 Canadian sites. In this analysis *H. arcticus sensu stricto* was found to be basal to the three other hesperodiaptomid species analyzed. Both *H. nearcticus* and *H. churchillensis* shared a greater genetic identity with *H. victoriaensis* than with *H. arcticus sensu stricto,* indicating a polyphyletic origin for the species. Additional studies of this kind are needed to clarify species relationships and only mating experiments among populations and comprehensive molecular surveys that define copepod species more accurately will provide the necessary information to do so.

The topology of the shoshone clade was similar in the maximum parsimony and neighbour joining analyses, but was not supported in the maximum likelihood based phylogenetic tree. The three species included in this clade occur in mountain habitats of western North America and their distribution is mostly allopatric, although *H. kenai* has been also reported at lower elevations. All three species are also included in Wilson's (1953) shoshone group demonstrating congruence between the morphological and molecular characters.

Hesperodiaptomus shoshone has been misidentified and confused with *H. kenai* on multiple occasions based upon the male fifth leg morphology. A re-examination of available specimens by Wilson (1953) found the records of *H. shoshone* by Marsh (1920) from the Pribolof Islands, Alaska, and from Wheat Meadows, California, and those of Carl (1940) from British Columbia to be those of *H. kenai*. Confusion between the two species exists to this day and molecular characters such as the rDNA sequences provided in this study provide additional information for distinguishing the two species. Both species differ by 15 single nucleotide polymorphisms in the rDNA sequences and should prove to be useful in future species identifications.

The phylogeny presented in this study provides an initial step in the reconstruction of the evolutionary relationships in this group of diaptomid copepods. Further studies using more variable regions of the genome along with large number of species having a broader representation of the morphological variation present in the genus are needed to clarify the phylogeny. Accurate phylogenetic reconstructions are also dependent upon accurate and stable taxonomy. This can only be accomplished when all the *Hesperodiaptomus* species are described at which time the biogeographical history of the genus can be elucidated. Once such taxonomic and phylogenetic studies are complete, copepods should prove to be model organisms to evaluate theoretical evolutionary models of speciation, and of morphological and genetic differentiation (Boileau, 1991).

References

Anderson, R.S. (1967) Diaptomid copepods from two mountain ponds in Alberta. *Canadian Journal of Zoology* **45**:1043-1051.

Anderson, R.S. (1971) Crustacean plankton of 146 alpine and subalpine lakes and ponds in Western Canada. *Journal of Fisheries Research Board of Canada* **28**:311-321.

Anderson, R.S. (1980) Relationships between trout and invertebrate species as predators and the structure of the crustacean and rotiferan plankton in mountain lakes. In W. C. Kerfoot (Ed.), Evolution and Ecology of Zooplankton Communities (p. 635-641). University Press of New England, Hanover.

Beccerra, J. and D.L. Venable (1999) Nuclear ribosomal DNA phylogeny and its implications for evolutionary trends in the Mexican *Bursera*. *American Journal of Botany* **86**:1047-1057.

Boileau, M.G. (1989) Biochemical genetic variation in the freshwater copepoda: evolutionary consequences of passive dispersal in freshwater zooplankton. PhD thesis, University of Windsor, Windsor, Ontario. 167pp.

Boileau, M.G. (1991) A genetic determination of cryptic species (Copepoda:Calanoida) and their postglacial biogeography in North America. *Zoological Journal of the Linnean Society* **102**:375-396.

Borutskii, E.V., L.A. Stepanova and M.S. Kos (1991) Revision of the Calanoida of Freshwaters of the USSR. (in Russian; abstract in English). Zoological Institute, USSR Academy of Sciences, Saint Petersburg.

Braga, E., R. Zardoya, A. Meyer and J. Yen (1999) Mitochrondrial and nuclear rRNA based copepod phylogeny with emphasis on the Euchaetidae (Calanoida). *Marine Biology* **133**:79-90.

Bucklin, A. (2000) Methods for Population Genetic Analysis of Zooplankton. Chapter 11. In R. Harris, P. Wiebe, J. Lenz, H. Skjoldal, and M. Huntley (Eds.), The Zooplankton Methodology Manual, International Council for the Exploration of the Sea (p 533 – 570). Academic Press, London.

Bucklin, A., M. Guarnieri, R.S. Hill, A.M. Bentley and S. Kaartvedt (1999) Taxonomic and systematic assessment of planktonic copepods using mitochrondrial COI sequence variation and competitive, species-specific PCR. *Hydrobiologia* **401**:239-254.

Carl, G.C. (1940) The distribution of some Cladocera and free-living Copepoda in British Columbia. *Ecological Monographs* **10**:55-110.

Dussart, B.H. and D. Defaye (2001) Introduction to the Copepoda. 2nd Ed. Guides to the Identification of the Microinvertebrates of the Continental Waters of the World. Backhuys Publishers, Leiden. 344 pp.

Granados-Ramírez, J.G. and E. Suárez-Morales (2003) A new *Hesperodiaptomus* Light (Copepoda, Calanoida, Diaptomidae) from Mexico with comments on the distribution of the genus. *Journal of Plankton Research* **25**(11):1383-1395.

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

Huelsenbeck, J.P. (1991) Tree-length distribution skewness: An indicator of phylogenetic information. *Systematic Zoology* **40**(3):257-270.

Huys, R. and G.A. Boxshall (1991) Copepod Evolution. The Ray Society, London. 468pp.

Kincaid, T. (1953) A contribution to the taxonomy and distribution of the American freshwater calanoid crustacea. The Calliostoma Company, Seattle. 73pp.

Lee, C.E. and B. Frost (2002) Morphological stasis in the *Eurytemora affinis* species complex (Copepoda: Temoridae). *Hydrobiologia* **480**:111-128.

Light, S.F. (1939) New American subgenera of *Diaptomus* Westwood (Copepoda, Calanoida). *Transactions of the American Microscopical Society* **58**(4):437-484.

Maddison, W.P. and D.R. Maddison (1992) MacClade version 3: Analysis of phylogeny and character evolution. Sinauer Associates, Massachusetts.

Marsh, C.D. (1907) A revision of the North American species of *Diaptomus*. *Transactions of the Wisconsin Academy of Sciences, Arts, and Letters* **15**(2):381-516.

Marsh, C.D. (1920) Freshwater Copepoda. *Report of the Canadian Arctic Expedition 1913-1918. Ottawa* 7:1-25.

Marsh, C.D. (1929) Distribution and key to North American copepods of the genus *Diaptomus* with the description of a new species. *Proceedings of the United States National Museum* **75**(14):1-27.

Paul, A.J., P.R. Leavitt, D.W. Schindler and A.K. Hardie (1995) Direct and indirect effects of predation by a calanoid copepod (subgenus: *Hesperodiaptomus*) and nutrients in a fishless alpine lake. *Canadian Journal of Fisheries and Aquatic Science* **52**:2628-2638.

Posada, D. and K.A. Crandall (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**:917-818.

Reed, E.B. (1959) The distribution and ecology of freshwater Entomostraca in arctic and subarctic North America. PhD thesis, University of Saskatchewan, Saskatoon, Saskatchewan. 142pp.

Sandercock, G.A. and G.G.E. Scudder (1996) Key to the Species of Freshwater Calanoid Copepods of British Columbia. British Columbia: Resources Inventory Committee Publications.http://srmwww.gov.bc.ca/risc/pubs/aquatic/calanoid/.

Scanlin, M. and J.W. Reid (1996) A new copepod species from California, USA: *Hesperodiaptomus californiensis* (Crustacea: Copepoda: Calanoida: Diaptomidae). *Proceedings of the Biological Society of Washington* **109**(1):103-111.

Schacht, F.W. (1897) The North American species of *Diaptomus. Illinois Natural History Survey Bulletin* **5**:97-207.

Staden, R. (1996) The Staden Sequence Analysis Package. *Molecular Biotechnology* **5**:233-241.

Stoddard, J.L. (1987) Microcrustacean communities of high-elevation lakes in the Sierra Nevada, California. *Journal of Plankton Research* **9**:631-650.

Streletskaya, E.Y. (1983) The "eiseni" group of the genus *Hesperodiaptomus* (Copepoda:Calanoida) and a new species *H. koolensis* from the Chukot Peninsula. *Zoologicheskii Zhurnal* **62**:1474-1480.

Swofford, D.L. (2001) PAUP* 4.0 - Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Assoc., Sunderland, Massachusetts.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins (1997) The CLUSTAL_X windows interface; Flexible strategies for multiple-sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **27**: 2682-2690.

Thum, R.A. (2004) Using 18S rDNA to resolve diaptomid copepod (Copepoda: Calanoida: Diaptomidae) phylogeny: an example with the North American genera. *Hydrobiologia* **519**: 135–141.

Williamson, C.E. and J.W. Reid (2001) Copepoda. In J.H. Thorp and A.P. Covich (Eds.), Ecology and Classification of North American Freshwater Invertebrates (p 915–954). Academic Press, New York.

Wilson, M.M. and H.C. Yeatman (1959) Free-Living Copepoda. In W.T. Edmonson and H.B. Ward (Eds.) Ward's and Whipple's Freshwater Biology (p.735-794). John Wiley and Sons, New York.

Wilson, M.S. (1953) New and inadequately known North American species of the copepod genus *Diaptomus*. *Smithsonian Miscellaneous Collections* **122**(2):1-30.

Wilson, M.S. (1958) New records and species of calanoid copepods from Saskatchewan and Louisiana. *Canadian Journal of Zoology* **36**:489-497.

Chapter 2

Phylogeography of *Hesperodiaptomus shoshone*

Introduction

The combination of population genetics, molecular systematics and biogeography has led to the creation of a new discipline known as phylogeography (Avise, 1994). Phylogeography is a field of study concerned with the principles and processes governing the geographic distributions of genealogical lineages, especially those within and among closely related species (Avise, 2000). In other words, phylogeography deals with the historical phylogenetic components of the spatial distribution of gene lineages (Avise, 2000).

Freshwater organisms have been the target of a considerable amount of phylogeographic study (e.g. fish: Bermingham and Avise, 1986; Bernatchez and Wilson, 1998; cladocerans: Cox and Hebert, 2001; Hyalella: Väinölä *et al.,* 2001; enallagma: Turgeon and McPeek, 2002; gastropods: Bunje, 2005 among many others). Such analyses have revealed historical patterns of range expansion, vicariance, lineage introgression and long-distance dispersal in different regions of the world. They have also provided invaluable information on the history, evolution and biogeography of different aquatic species.

The goal of the present study is to determine the patterns of genetic differentiation of populations of the freshwater copepod, *Hesperodiaptomus shoshone* (S.A. Forbes 1893), based on mtDNA and nuclear rDNA data, to assess the influence of historical events on dispersal, vicariance and diversification throughout its range in North America. An understanding of the influence of historical processes on the current distribution patterns

provides a means to assess the evolutionary potential of the species and to understand the speciation processes prevalent in this group of organisms.

Hesperodiaptomus shoshone is a sexually reproducing calanoid copepod that occurs in high altitude lentic aquatic systems of the Rocky Mountains and the Sierra Nevada. The insular nature of the distribution of aquatic environments in this system and the complex geological and glacial history of the area provide an unprecedented opportunity to study the phylogeography of passively dispersing aquatic invertebrates.

Pleistocene glaciations

The Pleistocene glaciations were arguably the most important historical events that occurred during the evolutionary life span of extant species. These glaciations may have influenced the present distribution and genetic diversity of all living taxa in the northern hemisphere (Hewitt, 1996). The Pleistocene epoch, which began about 2 million years ago and extended to about 10 000 years ago, was a time of successive warm and cold periods. During this epoch, there have been at least four major glaciations in North America, each named after the state where evidence of its existence was first obtained (Nebraskan, Kansan, Illinoian, and Wisconsinan) (Flint, 1971). Only the events related to the most recent (Wisconsinan) glaciation are known in detail, as each ice sheet obliterated most traces of the previous glaciations (Wilson and Hebert, 1998).

The Wisconsinan glaciation reached maximum ice coverage 23 000-18 000 years ago with a major deglaciation occurring between 15 000 and 8 000 years ago (Dyke and Prest, 1987). During this period, two enormous sheets of glacial ice covered large areas of northern North America. The Cordilleran ice sheet covered the mountain region between the Pacific

Coast and the Rocky Mountains from just south of the 49th parallel to the Yukon and Alaska. The Laurentide ice sheet extended over almost all of the rest of Canada east of the Rockies, and over much of the northern United States, especially in the east. Numerous smaller alpine glaciers also formed in the mountains of the west: the Cascades, Central and Southern Rockies, and the Sierra Nevada.

Glaciations had remarkable effects on freshwater habitats, both through the destruction of old systems and the widespread creation of new lakes and rivers. Advancing glacial fronts caused widespread habitat destruction and displacement or extirpation of local populations (Pielou, 1991). Aquatic species were particularly affected, as opportunities for dispersal were limited to direct water connections leading away from the advancing ice fronts (Bernatchez and Wilson, 1998). Aquatic species displaced by the glaciers would also have had to survive for periods lasting thousands of years in fringe habitats along glacial margins. Such a scenario for survival of copepod populations is highly plausible due to the presence of resting eggs in diaptomids that allow persistence in temporary unpredictable environments. In fact, Hairston et al. (1995) found that the viability of some diapausing copepod eggs can be in the order of a few hundred years. Additionally, many current copepod populations still persist in habitats that are less than a square meter in size and under ten centimetres in depth (pers. obs.). In some regions, meltwater from retreating glaciers during deglaciation events may have formed large proglacial lakes, which facilitated the dispersal of freshwater species over vast geographical ranges.

Use of molecular tools to study phylogeography

Mitochondrial DNA has emerged as a powerful tool for examining biogeographic events, both above and below the species level (Avise *et al.,* 1987). In many species, the maternal inheritance, lack of recombination and high mutation rate relative to nuclear DNA make mtDNA an ideal tool for studying the relatedness of populations (phylogeography) and dispersal events (Avise *et al.,* 1987). Since the genetic structure of freshwater copepods is largely determined by the island like nature of their habitats, phylogeographic patterns of mtDNA diversity could be eadily interpreted in terms of historical effects rather than contemporary dispersal.

The mitochondrial gene targeted for this study is a portion of the cytochrome oxidase I gene (Cox I). Parts of this gene have been shown to vary significantly within species of marine copepods (Bucklin *et al.,* 1996) and in one freshwater copepod, *Aglaodiaptomus leptopus* (Guarnieri, 1996). In addition, this part of the Cox I gene has already been sequenced for multiple populations in several copepod species providing a database to assess the typical level of intraspecific variation (Bucklin *et al.,* 1996).

In addition to the mtDNA, partial nuclear ribosomal DNA gene sequences, including the ITS1, 5.8S and ITS2 regions were also employed in the phylogeographic analyses (more information on rDNA is provided in the general introduction). The two internally transcribed spacers have been sequenced in a few diaptomid copepod species, providing a database to assess the typical level of variation (Thum, unpublished Genbank submissions).

Objective:

To assess the influence of historical events on the evolutionary patterns and the present genetic structuring in the freshwater diaptomid copepod *H. shoshone* to address the following questions:

1) Is the pattern of genetic variation in *Hesperodiaptomus shoshone* geographically structured throughout its distribution range in North America?

2) Is there a correlation between genetic and geographical distances?

3) Is there a relationship between the phylogeographic structure of *H. shoshone* populations and geological history?

Methods

Study species

The species chosen for this study is the diaptomid copepod, *H. shoshone*. This species reproduces sexually, and females must copulate before a clutch of eggs is extruded (Watras and Haney, 1980). The life cycle includes six naupliar stages and five copepodid stages before the final moult into adulthood, after which growth stops (Edmonson and Ward, 1959). *Hesperodiaptomus shoshone* is restricted to high altitude lakes and ponds of the Rocky Mountains from British Columbia to Colorado, and the Sierra Nevada. This species is univoltine and matures early in the summer in Colorado (Maly and Maly, 1974).

Study populations

During the summer of 2003 and 2004, one hundred and fifty ponds and lakes were sampled from a geographic region extending from Alberta to Colorado and California. Approximately 90 of the lakes and ponds sampled contained populations of *H. shoshone.* Fifty-one populations throughout the range of *H. shoshone* were used in the phylogenetic and phylogeographic analyses (Table 1 and Figure 1). The ponds were selected in order to get the widest geographic distribution and replication in each watershed. The sister species, *Hesperodiaptomus kenai* (M.S. Wilson 1953) was used as an outgroup in the phylogenetic analyses.

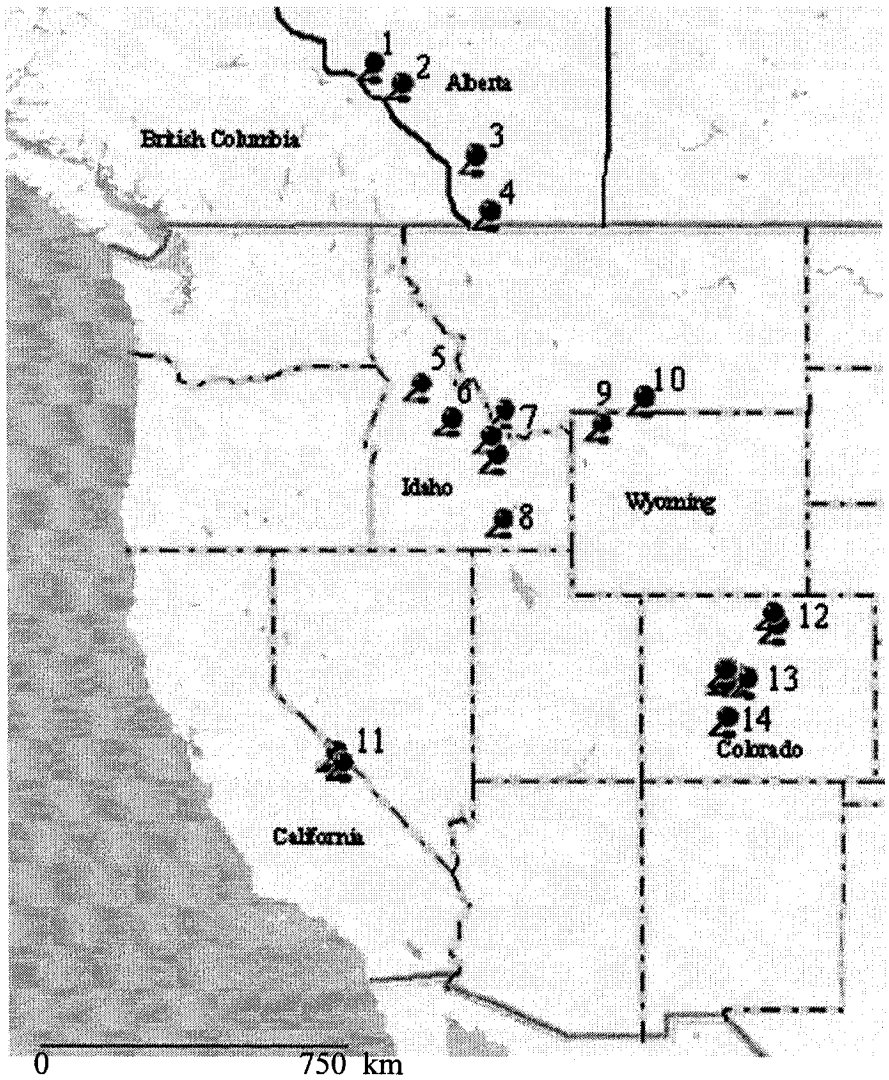
Table 1
Sampling locations of *Hesperodiaptomus shoshone* populations used in the phylogeographic analyses. Province and state abbreviations used: AL: Alberta, CA: California, CO: Colorado, ID: Idaho, WY:Wyoming.

Code	State Prov.	Lake ID	Pond/Lake/ Basin name	Altitude (m)	Coordinates Latitude	Longitude	Yr	Collector
AL1	AL		Plateau Pond	2124	50° 13.90'N	114° 30.50'W	04	M.Marszalek
AL2	AL		Teardrop Pond	2024	50° 12.10'N	114° 26.77'W	04	M. Marszalek
AL3	AL		Wilcox Lake	2358	52° 14.72'N	117° 11.92'W	04	M. Marszalek
AL4	AL		Surprise Lake	1664	52° 47.89'N	117° 37.86'W	04	M. Marszalek
AL5	AL		Mid. Rowe Lake	2140	49° 03.20'N	114° 03.40'W	04	M. Marszalek
AL6	AL		Upp. Rowe Lake	2154	49° 03.1'N	114° 03.63'W	04	M. Marszalek
CA1	CA	10222		3554	37° 05.21'N	118° 32.14'W	04	A. Kramer
CA2	CA	10230		3533	37° 04.43'N	118° 29.99'W	04	A. Kramer
CA3	CA	10223		3495	37° 05.00'N	118° 31.97'W	04	A. Kramer
CA4	CA	10257		3587	37° 03.46'N	118° 26.90'W	04	A. Kramer
CA5	CA	50224	Wahoo 2	3443	37° 13.71'N	118° 42.82'W	04	A. Kramer
CA6	CA	10263		3450	37° 02.09'N	118° 28.31'W	04	A. Kramer
CA7	CA	50176	Lobe, Upper	3291	37° 14.53'N	118° 45.19'W	03	A. Kramer
CA8	CA	41203	"Pavilion, Lower"	3323	37° 14.28'N	118° 47.53'W	03	A. Kramer
CA9	CA	50424	Split Lake	3413	37° 21.95'N	118° 44.58'W	03	A. Kramer
CA11	CA	50193	Petite, Lower	3504	37° 18.29'N	118° 42.54'W	03	A. Kramer
CA12	CA	50154	Goethe, Upper	3514	37° 13.16'N	118° 42.16'W	03	A. Kramer
CA13	CA	50194	Petite, Upper	3504	37° 18.26'N	118° 42.42'W	03	A. Kramer
CA14	CA	40222	Ramona	3290	34° 44.96'N	115° 02.51'W	01	A. Kramer
CA15	CA	52103	Frog Lake	3632	37° 15.27'N	118° 40.76'W	03	A. Kramer
CA18	CA	50423	Spire	3523	37° 22.02'N	118° 45.05'W	03	A. Kramer
CA19	CA	50207	Puppet Lake	3422	37° 17.06'N	118° 44.05'W	03	A. Kramer
PF3	CA	11910	Alpine	3675	37° 12.55N	118° 41.45'W	02	M. Pfrender
CO1	CO		Glacier Gorge 1	3429	40° 15.61'N	105° 38.05'W	04	M. Marszalek
CO3	CO		Niwot Ridge 1	3075	40° 01.87'N	105° 33.29'W	03	M. Marszalek
CO4	CO		Niwot Ridge 2	3079	40° 01.81'N	105° 33.43'W	03	M. Marszalek
CO5	CO		Lake City 1(CR1)	3554	38° 03.18'N	107° 21.26'W	03	M. Marszalek

Table 1 cont.

CO7	CO		Schofield 1	3713	39° 02.75'N	107° 02.75'W	04	M. Marszalek
CO8	CO		West Maroon Pass	3630	39° 01.91'N	107° 00.60'W	04	M. Marszalek
CO9	CO		Hasley Basin 1	3577	39° 03.46'N	107° 02.79'W	04	M. Marszalek
CO10	CO		Hasley Basin 8	3541	39° 03.96'N	107° 02.87'W	04	M. Marszalek
CO11	CO		Peeler BasinU4	3357	38° 53.76'N	107° 04.58'W	03	M. Marszalek
CO12	CO		Redwell Basin	3431	38° 53.59'N	107° 03.62'W	03	M. Marszalek
CO13	CO		Mexican Cut L11	3400	39° 01.79'N	107° 03.98'W	03,04	M. Marszalek
CO14	CO		Mexican Cut U2	3460	39° 01.55'N	107° 03.84'W	03,04	M. Marszalek
CO15	CO		Baldy Basin 1	3476	38° 58.76'N	107° 02.36'W	04	M. Marszalek
CO16	CO		Kettle ponds K4	2860	38° 56.70'N	106° 58.60'W	03	M. Marszalek
CO17	CO		Taylor Park (TP2)	3218	38° 56.29'N	106° 38.87'W	03	M. Marszalek
CO18	CO		Cottonwood Pass	3675	38° 49.72'N	106° 24.65'W	04	M. Marszalek
CO20	CO		Mexican Cut L6	3409	39° 01.68'N	107° 03.79'W	03,04	M. Marszalek
ID1	ID		Rough Pond 2	2925	43° 42.95'N	113° 50.38'W	04	M. Marszalek
ID5	ID		Middle smiley pond	3020	43° 42.38'N	113° 50.24'W	04	M. Marszalek
ID8	ID		Lava Butte SW	2325	45° 16.54'N	116° 07.45'W	04	M. Marszalek
ID2	ID		Tech Sergeant pond	2487	44° 31.64'N	115° 11.45'W	04	M. Marszalek
ID4	ID		Pond 2 Soldier Lakes	2553	44° 31.44'N	115° 11.52'W	04	M. Marszalek
ID9	ID		Mt. Harrison	2616	42° 18.35'N	113° 39.29'W	04	M. Marszalek
ID10	ID		Mill Pond 2	2849	44° 38.79'N	113° 39.56'W	04	M. Marszalek
ID3	ID	WL9389	U. Mill Lake	2849	44° 38.72'N	113° 39.86'W	04	M. Marszalek
WY1	WY		"Beartooth 8"	3051	44° 56.97'N	109° 29.04'W	03	M. Marszalek
WY2	WY		"Beartooth 9"	3053	44° 56.84'N	109° 28.98'W	03	M. Marszalek
WY3	WY	WL7790	Shoshone Lake	2378	44° 22.33'N	110° 42.74'W	03	M. Marszalek

Figure 1
Sampling locations of *Hesperodiaptomus shoshone* populations used in the phylogeographic analyses. The numbers of the locations on the figure correspond to populations listed in Table 1: 1) AL4; 2) AL3; 3)AL1, AL2; 4) AL5, AL6; 5) ID8; 6) ID2, ID4; 7) ID1, ID3, ID5, ID10; 8) ID9; 9) WY3; 10) WY1, WY2; 11) CA1, CA2, CA3, CA4, CA5, CA6, CA7, CA8, CA9, CA11, CA12, CA13, CA14, CA15, CA18, CA19, PF3; 12) CO1, CO3, CO4; 13) CO7, CO8, CO9, CO10, CO11, CO12, CO13, CO14, CO15, CO16, CO17, CO18, CO20.



Sample collection and Identification

Copepods were collected using a Wisconsin plankton tow net (120 μm mesh size) and were preserved immediately in 95% ethanol following the procedure of Bucklin (2000). Individual copepods were identified from each population using the Wilson and Yeatman key (1959) prior to DNA extraction and amplification.

Molecular methods

DNA extraction

DNA from individual copepods was extracted by a technique modified from Lee and Frost (2002) and Bucklin (2000) for small tissue samples. The protocol is described in detail in Chapter 1.

Design of specific mtDNA primers for H. shoshone

A 740 bp fragment of the cytochrome oxidase I mitochondrial gene was amplified from 8 individuals (CO13 population) using universal primers by Folmer *et al.* (1994):

LCO-1490 5' - GGT CAA CAA ATC ATA AAG ATA TTG G – 3'

HCO-2198 5' - TAA ACT TCA GGG TGA CCA AAA AAT CA – 3'

Amplification was performed in 25 μl reactions consisting of 2.5 μl of buffer (0.2 M Tris pH 9.5; 0.25 M KCl; 1 mg/ml BSA, 5 μl/ml Tween 20), 2.5 μM MgCl2, 250 μM dNTP, 0.4 pmol of each primer, 1.5 units of *Taq* DNA Polymerase, and 5μl of template DNA. PCR was performed with the following profile: 1 min at 94°C (1 cycle), 1 min at 94°C, 30 sec at 37°C, 1 min at 72°C (40 cycles) with a final extension step of 2 min at 72°C. A negative control was included with each PCR mix to detect possible contamination.

PCR products were separated via electrophoresis using a 1% TBE agarose gel stained with 0.33 μg/mL ethidium bromide in 1x TBE electrode buffer. Electrophoresis was carried out apparatus at 90V for 35 min and DNA fragments were visualized with an ultraviolet transilluminator (UVP Inc) at 302nm. A GeneRuler 100-bp DNA ladder (Fermentas) was used to estimate DNA fragment sizes. The amplified DNA (4 reactions per individual for a total volume of 100 μl) was purified using a QIAGEN PCR purification kit and checked again on a gel. Sequencing reactions were performed in the forward and reverse directions using ABI Big Dye™ Terminator v 3.0 or 3.1 cycle sequencing ready reaction kits following the manufacturer's protocol. Following the sequencing reaction, the products were purified through NH_4OAc and ethanol precipitation, dried, resuspended in 15 μl of formamide EDTA, denatured at 94°C for 4 minutes and analyzed on a ABI 310 automated genetic analyzer (Applied Biosystems, Foster City CA).

Sequence processing

Chromatograms were first imported and processed for quality with the PreGap4 program of the Staden Package (Staden, 1996). The consensus sequence was then obtained using the shotgun assembly option of the Gap4 program of the Staden Package (Staden, 1996). Ambiguities between sequences were inspected visually with the Trev program of the same software.

Three new primers (LCF1, LCF2, HC546R) were designed based on the consensus sequence obtained above using the software Oligo 3.4 (Rychlik and Rhoads, 1989) as universal primers were not successful in amplifying some of these samples. The sequences of these primers are as follows.

LCF1 5'-TCG GGG ATG ACC AAA TTT ATA – 3'

LCF2 5'- GTT GTC ACA GCT CAT GCG TT -3'

HC0546R 5'- CGC TAG AAC TGG CAG AGA TA -3'

These new primers were used for PCR amplifications of one to four individuals per population. Some primers were more efficient in PCR amplification of samples from some populations than others due to possible mutations at primer binding sites. The same PCR conditions were used as above but with annealing temperatures ranging between 40-55°C depending on the population examined. The PCR products were checked on 1% TBE agarose gels and sequenced using the ABI 3730XL Automated DNA sequencer at the McGill University and Genome Québec Innovation Centre (Montreal, QC). Chromatograms were processed as described above.

Nuclear ribosomal DNA amplification and sequencing

A 770 bp fragment of the nuclear ribosomal gene family was amplified in individuals from 48 populations using universal primers as described by Beccerra and Venable (1999).

5F 5'- GCA AGTAAAAGTCGTAACAAGG -3'

4R 5' – TCCTCCGCTTATTGATATGC -3'

Amplification was performed in 25 µL reactions consisting of 2.5 µl of buffer (0.2M Tris pH 9.5; 0.25 M KCl; 1 mg/ml BSA, 5 µl/ml Tween 20), 2.5 mM $MgCl_2$, 250 µM dNTP, 0.4 pmol of each primer, 1.5 units of *Taq* DNA polymerase and 5 µL template DNA. PCR was performed with the following profile: 2 min at 94°C (1 cycle), 1 min at 94°C, 30 sec at 50°C, 2 min at 72°C (35 cycles) with a final extension step of 4 min at 72°C , a negative control was included in each PCR reaction. The PCR products were checked on 1% agarose

gels and sent for sequencing at the McGill University and Genome Québec Innovation Centre (Montreal, QC). The sequences were processed as described above.

Data Analysis

Sequence alignment

Sequences were aligned using both Clustal X version 1.8.3 (Thompson *et al.*, 1997) and DAMBE version 4.0.3 (Xia and Xie, 2001), followed by manual refinement of the multiple sequence alignment using MacClade (Maddison and Maddison, 1992). DAMBE version 4.0.3 (Xia and Xie, 2001) was used for aligning the mtDNA cytochrome oxidase I protein coding nucleotide sequences against aligned amino acid sequences keeping the within codon sequences intact. Other multiple sequence alignment programs often introduce frame shift indels within codons in the aligned protein coding sequences, which are erroneous as the biologically meaningful aligned sequences of functional genes should have complete codons without interruptions by indels. DAMBE version 4.0.3 (Xia and Xie, 2001) solves the problem of alignment artefacts by aligning the protein coding nucleotide sequences against aligned amino acid sequences (Xia and Xie, 2001). All alignments are included in Appendix 2 and 3.

Analysis of the phylogenetic signal

To estimate the phylogenetic information content of the mtDNA and rDNA character matrices, the g_1-test (Huelsenbeck, 1991; Hillis and Huelsenbeck, 1992) was computed to assess the skewdness of the tree length distribution of 100 000 trees randomly generated with PAUP* (Swofford, 2001). Given a data set, the tree lengths under a maximum parsimony criterion for all possible topologies (or a random sample of them) are computed. If there is

no phylogenetic signal in the data, the distribution of tree lengths tends to be symmetrical. If there is a phylogenetic signal, the distribution tends to be left skewed (Hillis and Huelsenbeck, 1992). The probability of the phylogenetic structure was assessed using the values provided by Hillis and Huelsenbeck (1992).

Substitution Saturation

The presence of substitution saturation in the aligned sequences was tested with DAMBE version 4.0.43 (Xia and Xie, 2001) by plotting the number of transversions and transitions divided by sequence length against genetic distances for 1^{st} / 2^{nd} and 3^{rd} codon positions for mtDNA and for all positions for the rRNA gene, as it is not a protein coding gene. Transitional substitutions occur more frequently than transversional substitutions. Although they should both increase linearly with increasing genetic distance, as the genetic distance becomes larger (i.e. more divergent sequences are compared), the observed numbers of transitions reaches a plateau, due to multiple substitutions at a given site. Beyond that point, it is no longer possible to infer the expected genetic distance from the observed distance and the sequences are said to be saturated. Similarities between them could be the result of chance alone rather than homology. In coding regions, the 3^{rd} codons positions evolve much faster than the 1^{st} and 2^{nd} and are more likely to be saturated at a lower genetic distance. For this reason saturation was tested for both the 1^{st} /2^{nd} and 3^{rd} codon positions. Additionally, substitution saturation was evaluated with the saturation index developed by Xia *et al.* (2003) as implemented in DAMBE version 4.0.43 (Xia and Xie, 2001). Simulation studies show that phylogenetic information is essentially lost when the observed saturation is equal to or greater than half of full substitution saturation (Xia *et al,* 2003). Comparison of

the saturation index expected when assuming half of full saturation (Issc) with the observed saturation index (Iss) was performed with a *t* test in DAMBE version 4.0.3 (Xia *et al.,* 2003). In the absence of saturation, Iss should be significantly smaller than Issc.

Descriptive statistics

Percent nucleotide sequence divergence (uncorrected p) among haplotypes was obtained using the software MEGA 3.1 (Kumar *et al.*, 2004). Haplotype diversity (*h*), nucleotide diversity (π) (Nei 1987), and Tajima's (1989) *D* test for departure from neutrality were calculated for both molecular markers with the program DNASP version 4.10 (Rozas *et al.*, 2003).

Phylogenetic analyses

Phylogeny reconstruction for the mtDNA sequences was based on a 455 bp region corresponding to positions 135-588 in the harpacticoid copepod *Tigriopus japonicus* mtDNA cytochrome oxidase I gene sequence (Machida *et al.,* 2002;GenBank accession number AB060648.1; GI:21320907) and on a 739 bp region of the nuclear rDNA region that included ITS1, 5.8s and ITS2 regions. The data were analyzed using traditional phylogenetic methods and nested clade analyses using the software programs PAUP* 4.0 (Swofford, 2001) and GEODIS version 2.4 (Posada *et al.,* 2000), respectively.

Although traditional methods for estimating species phylogenies are of limited utility to assess within-species phylogeny, they were performed in order to compare the efficiency of the three different methods of analysis (nested clade analysis, phylogeny, isolation by distance and mismatch distribution). The difficulties of reconstructing within-species phylogenies with classical phylogenetic methods include lack of resolution and the difficulty

in identifying a root (Crandall *et al.*, 1994). Within-species data sets usually have fewer characters for phylogenetic analysis, reducing the statistical power of traditional phylogenetic methods. Also, in natural populations most, haplotypes in the gene pool exist as sets of multiple identical copies. The ancestral haplotypes persist in the population and are sampled together with their descendants. A problem related to the persistence of ancestral haplotypes in a population is that a single ancestral haplotype will often give rise to multiple descendant haplotypes, yielding a tree with true multifurcations. Traditional phylogenetic trees are generally based on a bifurcating tree (Posada and Crandall, 2001).

All analyses were performed on unique haplotypes in order to minimize computational time. Data were subjected to three different methods of phylogenetic reconstruction: i) maximum parsimony (MP), ii) maximum likelihood (ML), and iii) neighbour joining (NJ). Maximum likelihood (ML) analysis was conducted on the probability model obtained from MODELTEST 3.0 (Posada and Crandall, 1998) using a heuristic search, TBR branch swapping, 10 random addition replicates and a random addition starting tree. Neighbour joining analysis (NJ) was also constructed with the evolutionary model obtained from MODELTEST 3.0 (Posada and Crandall, 1998). In the maximum parsimony analysis, gaps were treated as missing data and all characters were weighted equally. The heuristic search algorithm was implemented for 1 000 random sequence replicates with tree bisection and reconnection (TBR) branch swapping. The 50% majority rule consensus was calculated using all of the resulting most parsimonious and maximum likelihood trees. Robustness of the MP and NJ tree topologies were tested with the bootstrap analysis with 1000 replicates each. Due to computational constraints, bootstrap with 500

replicates was used to infer the robustness of the ML trees. All phylogenetic analyses were conducted using PAUP* 4.0 (Swofford, 2001).

Nested clade analysis

Nested clade analysis (Templeton, 1998) allows for statistical testing of the null hypothesis of no association between haplotype and geographical location. This method has the power to discriminate between phylogeographical patterns due to restricted gene flow and historical events operating at the population level (Templeton, 1998). It has been applied to scenarios as diverse as the origin and subsequent trajectory of human evolution (Hammer *et al.*, 1998) and to the phylogeographical structure of national and regional faunas (Creer *et al.*, 2001; Schultheis *et al.*, 2002). The first step of such an analysis is the estimation of the haplotype network using the statistical parsimony procedure developed by Templeton *et al.* (1992). The statistical parsimony algorithm begins by estimating the maximum number of differences among haplotypes as a result of single substitutions (i.e. those that are not the result of multiple substitutions at a single site) with a 95% statistical confidence (Posada and Crandall, 2001). This number is called the parsimony connection limit. After this, haplotypes differing by one change are connected, then those by two, by three and so on, until all the haplotypes are included in a single network or the parsimony connection limit is reached (Posada and Crandall, 2001). When parsimonious connections are not justified ($P < 95\%$) non-parsimonious connections between haplotypes can be still established using equations provided by Templeton *et al.* (1992).

Subsequently, the nesting algorithm (Templeton *et al.*, 1987, 1992; Crandall, 1996) assigns a series of nested clades to the set of plausible haplotype networks. The nesting

procedure consists of nesting n-step clades within (n+1) step clades, where n refers to number of transitional steps used to define the clade (Crandall, 1996). By definition each haplotype is a zero-step clade. The (n+1) step clades are formed by the union of all n-step clades that can be joined together by (n+1) mutational steps. The nesting rules start at the tips of the haplotype network and move one mutational step into the interior, uniting all haplotypes connected by this procedure into a 1-step clade. Following the initial 1-step clades from the tips, the procedure is repeated on the more interior portions of the haplotype tree until all haplotypes have been placed into 1-step clades. The next level of nesting rules uses the 1-step clades as units rather than individual haplotypes. The nesting rules are the same, however 2-step clades are now formed. The nesting procedure is repeated until a nesting level is reached such that the next higher nesting level would result in a single category spanning the original haplotype network. Many missing intermediates might be found in the network (they are represented by "o") and might be due to unsampled or extinct haplotypes. These missing intermediates must be considered in the nesting procedure to assure overall consistency.

Once the haplotype network has been converted into a nested design, the simplest test for geographical association treats each sample location as a categorical variable. An exact permutational contingency analysis of categorical variation is performed as described in Templeton and Sing (1993). Although this contingency analysis can detect significant geographical associations and localize them within the nested cladogram it does not incorporate any information about the geographical distances or positions among the sample locations (Templeton *et al.,* 1995). Thus, for all the clades where the nested contingency

analysis is significant a more elaborate analysis that uses information on geographical distance is then employed.

Two measures of geographical distance are then computed: the clade distance D_c,: and the nested clade distance D_n. The clade distance measures the average distance that an individual bearing a haplotype from the clade of interest lies from the geographical center of all individual bearing haplotypes from the same clade (Templeton, 1998). The nested clade distance measures the average distance that an individual bearing a haplotype from the clade of interest lies from the geographical center of all individuals bearing haplotypes from the next higher level nesting clade that contains the clade of interest (Templeton, 1998). Contrasts in distance measures between tip clades (clades that are not interior nodes in the haplotype network) and clades immediately interior to them in the nested design are important in determining geographical structuring of genetic variation (Templeton *et al.*, 1995; Templeton, 1998). The statistical significance of the different distance measures and the interior-tip contrasts are determined by random permutation testing (Templeton, 1998). The two major reasons for the failure to reject the null hypothesis (random geographical distribution for all clades within a nesting category) are:1) the samples are inadequate to detect geographical structuring even though it exists, and 2) the population is panmictic over the sample area. As there is no way of differentiating between the two alternatives, biological inference is confined to clades in which the null hypothesis is rejected by using the inference key presented by Templeton (1998).

A haplotype network was constructed by TCS version 1.21 (Clement *et al.*, 2000), and clades were nested according to rules outlined in Crandall (1996). GEODIS version 2.4 (Posada *et al.*, 2000) was used to test for significant associations between haplotype and

geography and to compute the distance measures and contrasts between tip-interior clades. Where significance is detected, the inference key of Templeton (1998) was used to determine the likely cause of the associations (the revised version can be downloaded from http://darwin.uvigo.es/download/geodisKey_11Nov05.pdf).

Mismatch Distribution

Demographic histories of the main regions were additionally inferred by plotting mismatch distributions (Slatkin and Hudson, 1991). Mismatch distributions are calculated by counting the number of nucleotide site differences between each pair of individuals and using a bar diagram to display the relative frequencies of pairs that differ by zero sites, one site etc. Analysis of the mismatch distribution can provide insight into the historical demography of a population. Empirical data and computer simulations have shown that a stable population will produce a multimodal distribution, as the rise of new mutations is offset by the loss of variation due to random genetic drift, resulting in random lineage sorting (Harpending *et al.*, 1998). In contrast, the pairwise differences in a population that has expanded in the past will fit a Poisson distribution, due to the rate of accumulation of new mutations being greater than the loss of variation through drift.

The expected mismatch distribution was generated using the software DNASP version 4.10 (Rozas *et al.*, 2003). Testing whether the observed distribution of pairwise differences deviated significantly from that expected under the sudden population expansion model was done using a Kologorov- Smirnov test in SPSS 11.5 (1999).

One concern about using Mismatch distribution for a group of subdivided populations, as found in this study, is that the assumption of random mating is violated.

However, Rogers (1995) showed that the theory behind Mismatch distribution analysis is robust and should approximately hold true even when populations are completely isolated.

Isolation by Distance

Mantel's tests (Legendre and Legendre, 1998) were performed to test the correlation between genetic and geographic distances. A simple Pearson's R correlation coefficient cannot be used because the cases are not independent (e.g. the distance between sequence 1 and 3 is not independent of the distance between sequence 1 and 4 because sequence 1 is involved in both). This test indicates whether differentiation among the populations occurred through gradual isolation by distance.

Pairwise geographic distances between populations were calculated while accounting for the curvature of the earth geographic distances in the R package 3.0 (Casgrain and Legendre, 2001). Saturation curves (refer to results section) suggest that there is no mutational saturation within *H. shoshone*, and thus simple percent divergence were used as a measure of genetic distance (Kuchta and Tan, 2005). The computer program IBD (Bohonak, 2002) was used to calculate Mantel tests (10 000 randomizations) and the coefficient of determination.

Results

Descriptive Statistics

Alignment of the 455 bp mtDNA region resulted in 106 variable sites while the alignment of the 739 bp rDNA sequences resulted in 22 variable sites. A few populations were discarded from analyses due to poor sequencing or failed PCR reactions. Levels of haplotype diversity and nucleotide diversity were different between geographic regions (Table 2). In the mtDNA data the Colorado populations showed the highest haplotype (h) and nucleotide diversity (π). As opposed to the mtDNA data, the rDNA Colorado populations have the lowest haplotype (h) diversity. Furthermore, the Colorado populations are represented by only three haplotypes with the majority of the populations restricted to one haplotype.

Uncorrected ("*p*") sequence divergence levels among all *H. shoshone* haplotypes ranged from 0.2 % to 17.3 % in the mtDNA data set (Table 3) and from 0.01% to 1.9 % in the rDNA dataset (Table 4). Average sequence divergence between the "North" and "South" lineages was 1.9 % (std. dev. 0.11) in the rDNA dataset and 16.9 % (std. dev. 1.3) in the mtDNA sequences while the divergence between *H. shoshone* and the outgroup *H. kenai* was 2.4 % and 19 % for rDNA and mtDNA respectively.

The result of the Tajima's test applied to unique haplotypes of the mtDNA and rDNA sequences and *H. kenai* outgroup showed no significant departure from neutrality (mtDNA: $D = 0.139$, $p > 0.10$; rDNA; $D = 0.291$, $p > 0.10$).

Table 2
Descriptive statistics of *Hesperodiaptomus shoshone* populations for mtDNA and rRNA gene sequence diversity. The statistics are reported by geographical region; North (Idaho, Wyoming, and Alberta), CA (California), and CO (Colorado).

	mtDNA				**rRNA gene**			
	North	***CA***	***CO***	**Total**	***North***	***CA***	***CO***	**Total**
Number of populations	15	16	17	48	16	17	15	48
Number of individuals sequenced	24	32	23	79	18	33	26	77
Number of polymorphic sites	25	8	72	106	7	4	7	22
Number of haplotypes	11	12	15	38	6	5	3	14
Haplotype Diversity (h)	0.963	0.958	0.978	0.998	0.783	0.794	0.257	0.881
Nucleotide Diversity (π)	0.018	0.001	0.047	0.099	0.002	0.002	0.002	0.011

Table 3

Uncorrected ("*p*") sequence divergence levels among the 48 *Hesperodiaptomus shoshone* haplotypes and the sister taxa *Hesperodiaptomus kenai* mtDNA cytochrome oxidase I sequence data. Refer to Table 1 and Table 6 for more information on the populations and haplotypes.

	1	2	3	4	5	6	7	8	9	10	11	12
1 HAL2												
2 HAL3	0.004											
3 HCA1	0.162	0.164										
4 HCA2	0.164	0.167	0.002									
5 HCA4	0.162	0.164	0.004	0.002								
6 HCA6	0.162	0.164	0.007	0.004	0.002							
7 HCA7	0.164	0.162	0.013	0.016	0.013	0.011						
8 HCA8	0.16	0.162	0.007	0.009	0.007	0.009	0.007					
9 HCA9	0.164	0.162	0.011	0.013	0.011	0.013	0.002	0.004				
10 HCA11	0.162	0.16	0.009	0.011	0.009	0.011	0.004	0.002	0.002			
11 HCA13	0.162	0.164	0.004	0.007	0.004	0.007	0.009	0.002	0.007	0.004		
12 HCA14	0.16	0.162	0.007	0.009	0.007	0.009	0.011	0.004	0.009	0.007	0.002	
13 HCA18	0.164	0.162	0.007	0.009	0.007	0.009	0.007	0.004	0.004	0.002	0.002	0.004
14 HPF3	0.162	0.164	0.004	0.007	0.004	0.007	0.009	0.007	0.007	0.009	0.004	0.007
15 HCO1	0.16	0.156	0.098	0.096	0.098	0.096	0.093	0.1	0.096	0.098	0.102	0.104
16 HCO3	0.162	0.158	0.1	0.098	0.1	0.098	0.096	0.102	0.098	0.1	0.104	0.107
17 HCO4	0.16	0.156	0.1	0.098	0.1	0.098	0.096	0.102	0.098	0.1	0.104	0.107
18 HCO5	0.167	0.162	0.104	0.102	0.104	0.107	0.104	0.107	0.102	0.104	0.109	0.111
19 HCO8	0.158	0.153	0.1	0.098	0.096	0.093	0.091	0.098	0.093	0.096	0.1	0.102
20 HCO7	0.151	0.147	0.098	0.096	0.093	0.091	0.089	0.096	0.091	0.093	0.098	0.1
21 HCO11	0.158	0.158	0.073	0.076	0.078	0.076	0.073	0.076	0.076	0.078	0.078	0.08
22 HCO12	0.158	0.158	0.078	0.08	0.082	0.08	0.078	0.08	0.08	0.082	0.082	0.084
23 HCO13	0.153	0.149	0.084	0.082	0.08	0.078	0.076	0.082	0.078	0.08	0.084	0.087
24 HCO14	0.153	0.149	0.084	0.082	0.08	0.078	0.076	0.082	0.078	0.08	0.084	0.087
25 HCO15	0.151	0.151	0.078	0.08	0.078	0.076	0.073	0.076	0.076	0.078	0.078	0.08
26 HCO16	0.142	0.138	0.093	0.091	0.089	0.087	0.087	0.091	0.089	0.089	0.093	0.096
27 HCO17	0.149	0.144	0.096	0.098	0.1	0.098	0.091	0.098	0.093	0.096	0.1	0.102
28 HCO18	0.158	0.153	0.1	0.098	0.1	0.098	0.096	0.102	0.098	0.1	0.104	0.107
29 HCO20	0.156	0.151	0.091	0.089	0.087	0.084	0.082	0.089	0.084	0.087	0.091	0.093
30 HWY3	0.031	0.031	0.164	0.162	0.16	0.16	0.162	0.162	0.162	0.16	0.164	0.162
31 HID1	0.02	0.016	0.162	0.16	0.158	0.158	0.16	0.16	0.16	0.158	0.162	0.16
32 HID2	0.024	0.02	0.158	0.156	0.153	0.153	0.156	0.156	0.156	0.153	0.158	0.156
33 HID3	0.022	0.018	0.164	0.162	0.16	0.16	0.162	0.162	0.162	0.16	0.164	0.162
34 HID8	0.024	0.02	0.162	0.16	0.158	0.158	0.16	0.16	0.16	0.158	0.162	0.16
35 HID9	0.024	0.02	0.164	0.162	0.16	0.16	0.162	0.162	0.162	0.16	0.164	0.162
36 HID10	0.024	0.02	0.164	0.162	0.16	0.158	0.16	0.162	0.162	0.16	0.164	0.162
37 HWY1	0.027	0.027	0.167	0.164	0.162	0.162	0.164	0.164	0.164	0.162	0.167	0.164
38 HWY2	0.027	0.027	0.162	0.16	0.158	0.158	0.16	0.16	0.16	0.158	0.162	0.16
39 *Hkenai*	0.211	0.209	0.198	0.196	0.198	0.196	0.193	0.2	0.196	0.198	0.202	0.204

Table 3 cont.

	13	14	15	16	17	18	19	20	21	22	23	24	25
1 HAL2													
2 HAL3													
3 HCA1													
4 HCA2													
5 HCA4													
6 HCA6													
7 HCA7													
8 HCA8													
9 HCA9													
10 HCA11													
11 HCA13													
12 HCA14													
13 HCA18													
14 HPF3	0.007												
15 HCO1	0.1	0.098											
16 HCO3	0.102	0.1	0.007										
17 HCO4	0.102	0.1	0.002	0.004									
18 HCO5	0.107	0.104	0.098	0.1	0.1								
19 HCO8	0.098	0.096	0.047	0.044	0.044	0.084	0.007						
20 HCO7	0.096	0.093	0.049	0.047	0.047	0.084	0.047	0.049					
21 HCO11	0.08	0.073	0.049	0.051	0.051	0.089	0.047	0.036	0.024				
22 HCO12	0.084	0.078	0.047	0.044	0.044	0.084	0.038	0.031	0.031	0.031			
23 HCO13	0.082	0.08	0.036	0.033	0.033	0.082	0.029	0.031	0.027	0.027	0.004		
24 HCO14	0.082	0.08	0.036	0.033	0.033	0.078	0.024	0.027	0.027	0.022	0.031	0.031	
25 HCO15	0.08	0.073	0.051	0.053	0.053	0.093	0.047	0.044	0.029	0.022	0.031	0.016	0.038
26 HCO16	0.091	0.091	0.04	0.038	0.038	0.082	0.036	0.038	0.038	0.038	0.016	0.016	0.038
27 HCO17	0.098	0.096	0.087	0.089	0.089	0.049	0.071	0.064	0.064	0.064	0.069	0.064	0.073
28 HCO18	0.102	0.1	0.093	0.091	0.091	0.022	0.076	0.076	0.08	0.071	0.069	0.064	0.084
29 HCO20	0.089	0.087	0.038	0.036	0.036	0.084	0.027	0.029	0.033	0.033	0.007	0.007	0.033
30 HWY3	0.162	0.164	0.169	0.171	0.169	0.171	0.164	0.158	0.169	0.173	0.162	0.162	0.167
31 HID1	0.16	0.162	0.158	0.16	0.158	0.16	0.153	0.147	0.164	0.164	0.151	0.151	0.158
32 HID2	0.156	0.158	0.153	0.156	0.153	0.16	0.149	0.142	0.16	0.16	0.147	0.147	0.153
33 HID3	0.162	0.164	0.16	0.162	0.16	0.162	0.151	0.144	0.167	0.167	0.153	0.153	0.16
34 HID8	0.16	0.162	0.158	0.16	0.158	0.162	0.151	0.147	0.164	0.164	0.151	0.151	0.158
35 HID9	0.162	0.164	0.162	0.164	0.162	0.164	0.153	0.147	0.169	0.169	0.156	0.156	0.162
36 HID10	0.162	0.164	0.158	0.16	0.158	0.162	0.149	0.142	0.164	0.164	0.151	0.151	0.158
37 HWY1	0.164	0.167	0.162	0.164	0.162	0.169	0.156	0.149	0.158	0.162	0.151	0.151	0.156
38 HWY2	0.16	0.162	0.162	0.164	0.162	0.169	0.158	0.151	0.162	0.167	0.156	0.156	0.16
39 *Hkenai*	0.2	0.198	0.191	0.193	0.191	0.213	0.198	0.198	0.196	0.202	0.191	0.191	0.202

Table 3 cont.

	26	27	28	29	30	31	32	33	34	35	36	37	38
1 HAL2													
2 HAL3													
3 HCA1													
4 HCA2													
5 HCA4													
6 HCA6													
7 HCA7													
8 HCA8													
9 HCA9													
10 HCA11													
11 HCA13													
12 HCA14													
13 HCA18													
14 HPF3													
15 HCO1													
16 HCO3													
17 HCO4													
18 HCO5													
19 HCO8													
20 HCO7													
21 HCO11													
22 HCO12													
23 HCO13													
24 HCO14													
25 HCO15													
26 HCO16	0.067												
27 HCO17	0.069	0.04											
28 HCO18	0.018	0.071	0.071										
29 HCO20	0.153	0.149	0.164	0.164									
30 HWY3	0.14	0.144	0.149	0.153	0.024								
31 HID1	0.136	0.14	0.149	0.149	0.029	0.004							
32 HID2	0.142	0.147	0.151	0.156	0.027	0.002	0.007						
33 HID3	0.14	0.144	0.151	0.153	0.029	0.004	0.004	0.007					
34 HID8	0.144	0.149	0.153	0.158	0.024	0.009	0.013	0.007	0.013				
35 HID9	0.14	0.144	0.149	0.153	0.029	0.004	0.009	0.002	0.009	0.009			
36 HID10	0.142	0.147	0.162	0.153	0.013	0.029	0.033	0.031	0.033	0.029	0.033		
37 HWY1	0.147	0.147	0.162	0.158	0.009	0.024	0.029	0.027	0.029	0.024	0.029	0.004	
38 HWY2	0.191	0.204	0.213	0.193	0.204	0.204	0.207	0.207	0.209	0.202	0.204	0.196	0.198
39 *Hkenai*													

Table 4
Uncorrected ("*p*") sequence divergence levels among the 48 *Hesperodiaptomus shoshone* haplotypes and the sister taxa *Hesperodiaptomus kenai* for nuclear ribosomal DNA sequence data. Refer to Table 1 and Table 7 for more information on the populations and haplotypes.

	1	2	3	4	5	6	7
Hapl 1							
Hapl 2	0.001						
Hapl 3	0.004	0.003					
Hapl 4	0.003	0.001	0.001				
Hapl 5	0.004	0.003	0.003	0.001			
Hapl 6	0.007	0.005	0.005	0.004	0.005		
Hapl 7	0.02	0.022	0.022	0.02	0.022	0.016	
Hapl 8	0.016	0.018	0.02	0.019	0.02	0.015	0.012
Hapl 9	0.015	0.016	0.019	0.018	0.019	0.014	0.011
Hapl 10	0.014	0.015	0.018	0.016	0.018	0.012	0.009
Hapl 11	0.016	0.018	0.02	0.019	0.02	0.015	0.012
Hapl 12	0.018	0.019	0.022	0.02	0.022	0.016	0.014
Hapl 13	0.016	0.018	0.018	0.016	0.018	0.012	0.009
Hapl 14	0.018	0.019	0.019	0.018	0.019	0.014	0.008
H.kenai	0.028	0.027	0.027	0.026	0.027	0.022	0.024

	8	9	10	11	12	13	14
Hapl 1							
Hapl 2							
Hapl 3							
Hapl 4							
Hapl 5							
Hapl 6							
Hapl 7							
Hapl 8							
Hapl 9	0.001						
Hapl 10	0.003	0.001					
Hapl 11	0.003	0.001	0.003				
Hapl 12	0.001	0.003	0.004	0.001			
Hapl 13	0.003	0.001	0.003	0.003	0.004		
Hapl 14	0.004	0.003	0.004	0.004	0.005	0.001	
H.kenai	0.026	0.024	0.023	0.026	0.027	0.023	0.024

Analysis of the phylogenetic signal and substitution saturation

The length-frequency distribution from 10,000 random trees was significantly skewed for both the mtDNA (mean =1231.19; std. dev. = 66.98; g_1=-0.54; $p<0.01$) and rDNA (mean =100.38; std. dev. = 8.5; g_1=-0.59; $p<0.01$) sequences, indicating the prevalence of phylogenetic signal over noise in both data sets.

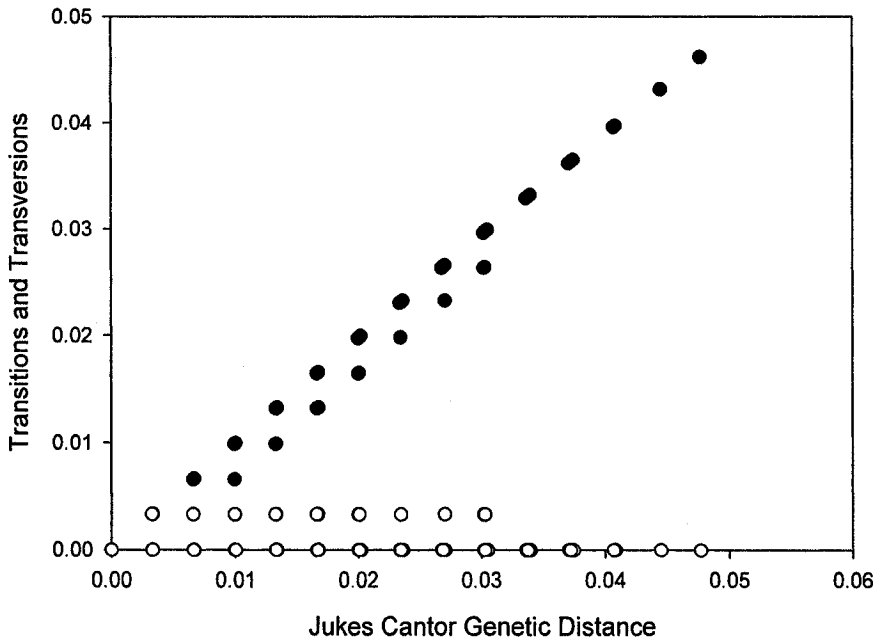
There was no saturation in the mtDNA sequences when both 1st / 2nd and 3rd codon positions (Figure 2) were analysed. Although, 3rd codon positions no longer increase linearly and are gradually reaching a plateau, transitions still outnumber transversions, indicating that enough signal was still retained for phylogenetic analyses. Expected and observed saturation indices were also compared for both 1st / 2nd and 3rd codon positions for the mtDNA and all positions for the rDNA dataset. The significant difference between the observed (Iss) and expected saturation (Iss_c) indices further confirmed the lack of substitution saturation in the mtDNA sequences (Table 5). There was no evidence of substitution saturation in the rDNA sequences from the transition and transversion plots (Figure 3) and from the substitution saturation indices (Table 5).

Table 5
Observed (Iss) and expected (Iss_c) saturation indices for mtDNA and rDNA data.

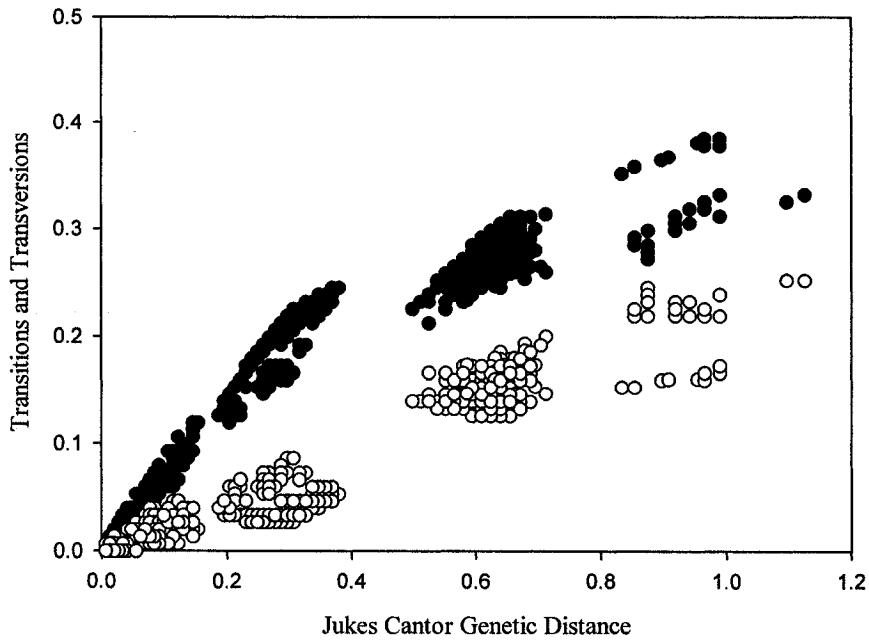
Marker	***Positions***	***Iss***	***Iss $_c$***	***p***
mt DNA	1st / 2nd	0.160	0.784	< 0.01
	3rd	0.363	0.773	< 0.01
rDNA	all	0.142	0.808	<0.01

Figure 2
Number of transitions and transversions versus divergence plots for the *Hesperodiaptomus shoshone* mitochondrial DNA sequence phylogeography dataset. The estimated number of transitions and transversions for each pairwise comparison is plotted against Jukes Cantor evolutionary distance. Transitions are represented by filled circles (●), and transversions are shown by open circles (○).

a) 1^{st} / 2^{nd} codon positions



b) 3rd codon position only



c) All positions

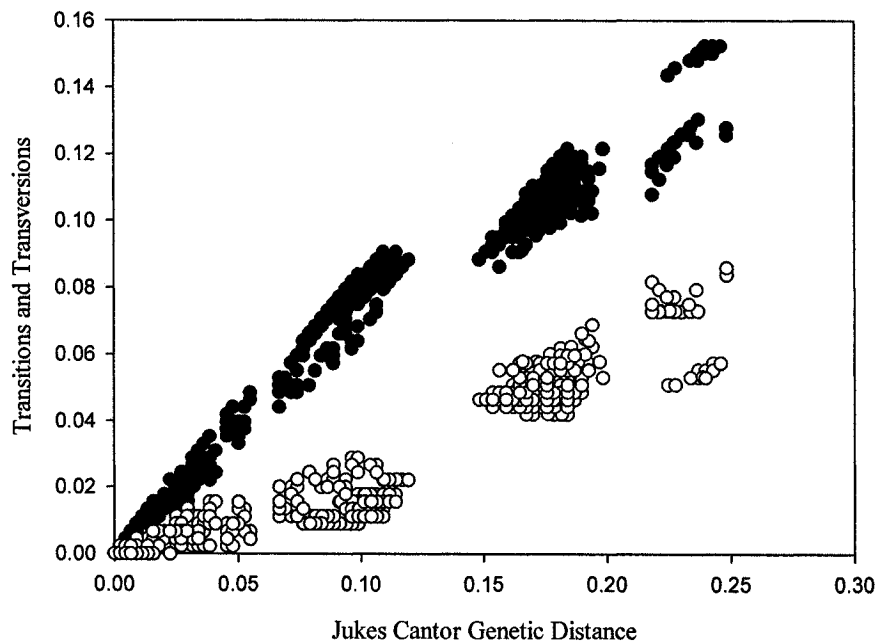
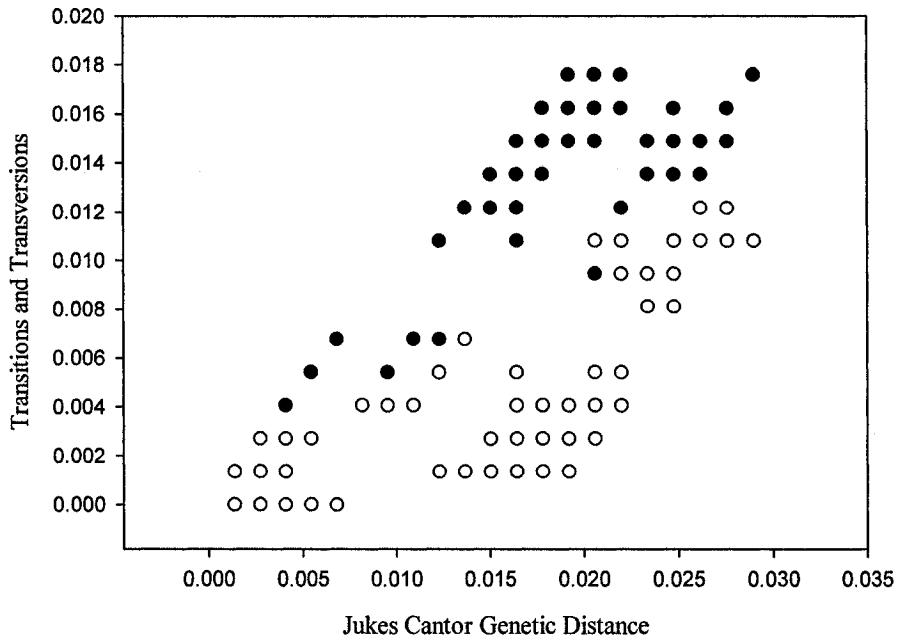


Figure 3
Number of transitions and transversions versus divergence plots for the *Hesperodiaptomus shoshone* rDNA sequence phylogeography dataset. The estimated number of transitions and transversions for each pairwise comparison is plotted against Jukes Cantor evolutionary distance. Transitions are represented by filled circles (●), and transversions are shown by open circles (○).



Phylogenetic Analysis

Alignment of DNA sequences was unambiguous, with no insertions or deletions among the taxa for both mtDNA and rDNA sequences. A total of 38 unique haplotypes was identified in the 48 *H. shoshone* mtDNA sequences amplified (Table 6). Only 48 of the 51 populations were successfully amplified with mtDNA primers. Table 6 shows the number of individuals sequenced per population for the mtDNA. The following populations were not included in the mtDNA analysis due to lack of amplification or poor sequence quality: AL1, AL4 and CA19. Most of the haplotypes were unique to each sampled population and only 7 haplotypes were found in more than one population. When translated into amino acids, no stop codons were observed in the mtDNA sequences and only five amino acid substitutions were detected among all haplotypes. The nuclear rDNA sequences were less variable and a total of 14 unique haplotypes (labelled H1-H14) were identified in the 739 bp dataset (Table 7). For the rDNA dataset 48 out of the 51 populations were successfully amplified except for: CO15, CO20 and ID10 populations.

Table 6
Haplotype identifications with corresponding sampled populations of the mtDNA sequences used in the nested clade analysis and phylogenetic analyses. Refer to Table 1 for further information on the sampled populations. Square brackets represent the number of individuals sequenced in each population.

Hapl.	Populations
HAL2	Alberta: AL2 [2]
HAL3	Alberta: AL3 [2], AL5 [2], AL6 [2]
HCA1	California: CA1 [2], CA3 [2]
HCA2	California: CA2 [2]
HCA4	California:CA4 [2], CA5 [2] , CA12 [2]
HCA6	California: CA6 [2]
HCA7	California: CA7 [2], CA15 [2]
HCA8	California: CA8 [2]
HCA9	California: CA9 [2]
HCA11	California: CA11 [2]
HCA13	California: CA13 [2]
HCA14	California: CA14 [2]
HCA18	California: CA18 [2]
HPF3	California: PF3 [2]
HCO1	Colorado: CO1 [1]
HCO3	Colorado: CO3 [1]
HCO4	Colorado: CO4 [1]
HCO5	Colorado: CO5 [1]
HCO7	Colorado: CO7 [1], CO9 [1], CO10 [1]
HCO8	Colorado: CO8 [1]
HCO11	Colorado: CO11 [1]
HCO12	Colorado: CO12 [1]
HCO13	Colorado: CO13 [4]
HCO14	Colorado: CO14 [1]
HCO15	Colorado: CO15 [1]
HCO16	Colorado: CO16 [1]
HCO17	Colorado: CO17 [1]
HCO18	Colorado: CO18 [1]
HCO20	Colorado: CO20 [4]
HID1	Idaho: ID1 [2]
HID2	Idaho: ID2 [2], ID4 [2]
HID3	Idaho: ID3 [1], ID5 [2]
HID8	Idaho: ID8 [2]
HID9	Idaho: ID9 [1]
HID10	Idaho: ID10 [1]
HWY1	Wyoming: WY1 [1]
HWY2	Wyoming: WY2 [1]
HWY3	Wyoming: WY3 [1]

Table 7
Haplotype identifications with corresponding sampled populations for the nuclear rDNA sequences used in the nested clade analysis and phylogenetic reconstructions. Refer to Table 1 for further information on the sampled populations.

Haplotype	Populations
H1	Alberta: AL4 [2], AL5 [1]
H2	Alberta: AL1 [1], AL3 [1]
H3	Alberta: AL2 [1]
H4	Wyoming: WY1 [1], WY2 [2], WY3 [1]; Alberta: AL6 [1]; Idaho: ID5 [1 ID8 [1], ID9[1]
H5	Idaho: ID3 [1]
H6	Idaho: ID1 [1], ID2 [1], ID4 [1]
H7	Colorado: CO3 [2], CO4 [1], CO5 [2], CO7 [1], CO9 [1], CO10 [1], CO11 [3], CO12 [2], CO13 [1], CO14 [1], CO16 [1], CO17 [2], CO18 [2]
H8	California: CA6 [2], CA13 [2], CA14 [2], CA15 [2]
H9	California: CA6 [2], CA7 [2], CA11 [2], CA13 [2], CA15 [2], CA18 [2], CA19 [2]
H10	California: CA19 [2]
H11	California: CA1 [2], CA2 [2], CA3 [2], CA4 [2], CA7 [2], CA8 [2], CA9 [2], CA18 [2], PF3 [1]
H12	California: CA2 [2]
H13	Colorado: CO8 [3]; California: CA12 [2], CA5 [2]
H14	Colorado: CO1 [3]

Independent likelihood ratio tests as implemented in MODELTEST version 3.7 (Posada and Crandall, 1998) selected a General Time Reversible model (GTR) of nucleotide substitution with estimated base frequencies (A= 0.2457; C= 0.2870; G= 0.2548; T= 0.2125) and proportion of invariable sites (I = 0.9019) for the rDNA dataset. The Kimura three-parameter (K81) (Kimura, 1981) model of nucleotide substitution was selected for the mtDNA cytochrome oxidase I partial sequence, with estimated base frequencies (A= 0.2466; C= 0.1913; G= 0.2208; T= 0.3412), proportion of invariable sites (I = 0.5377) and rates for variable sites following a gamma distribution (G`= 0.6911).

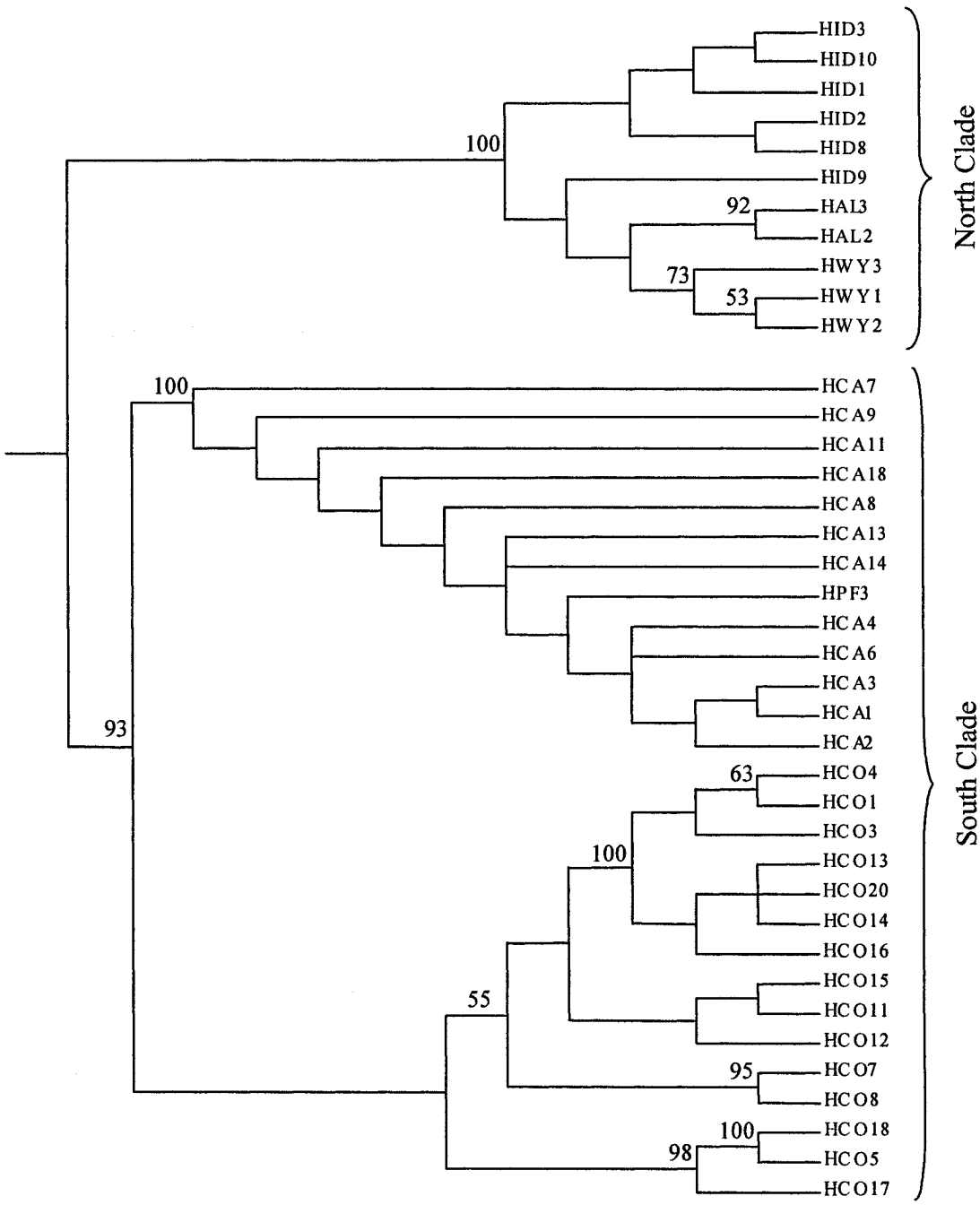
All three (parsimony, maximum likelihood, and neighbour joining) analyses identified similar tree topologies. An equally weighted maximum parsimony analysis of unique haplotypes of *H. shoshone* populations resulted in 100 most parsimonious trees (length 300; CI=0.647; RI=0.918) in the mtDNA sequences and 88 most parsimonious trees (length 41; CI=0.717; RI=0.819) in the rDNA dataset. Figures 4 and 5 show a 50% majority rule consensus of the most parsimonious trees and bootstrap confidence values for its nodes for the mtDNA and rDNA. Two main lineages with high bootstrap support (77-100 %) can be recognized in *H. shoshone* according to the MP phylogeny: (1) the north lineage which includes populations from Idaho, Alberta and Wyoming; (2) the south lineage which includes populations from Colorado and California. Both maximum parsimony, and neighbour joining analyses based upon mtDNA support a subdivision of the south lineage into Colorado and California clades. No such divisions were observed in the mtDNA maximum likelihood and in the rDNA phylogenetic reconstructions.

The maximum likelihood analysis of mtDNA placed the root of the *H. shoshone* phylogeny between HWY1 and HWY2 haplotypes while the neighbour joining analysis

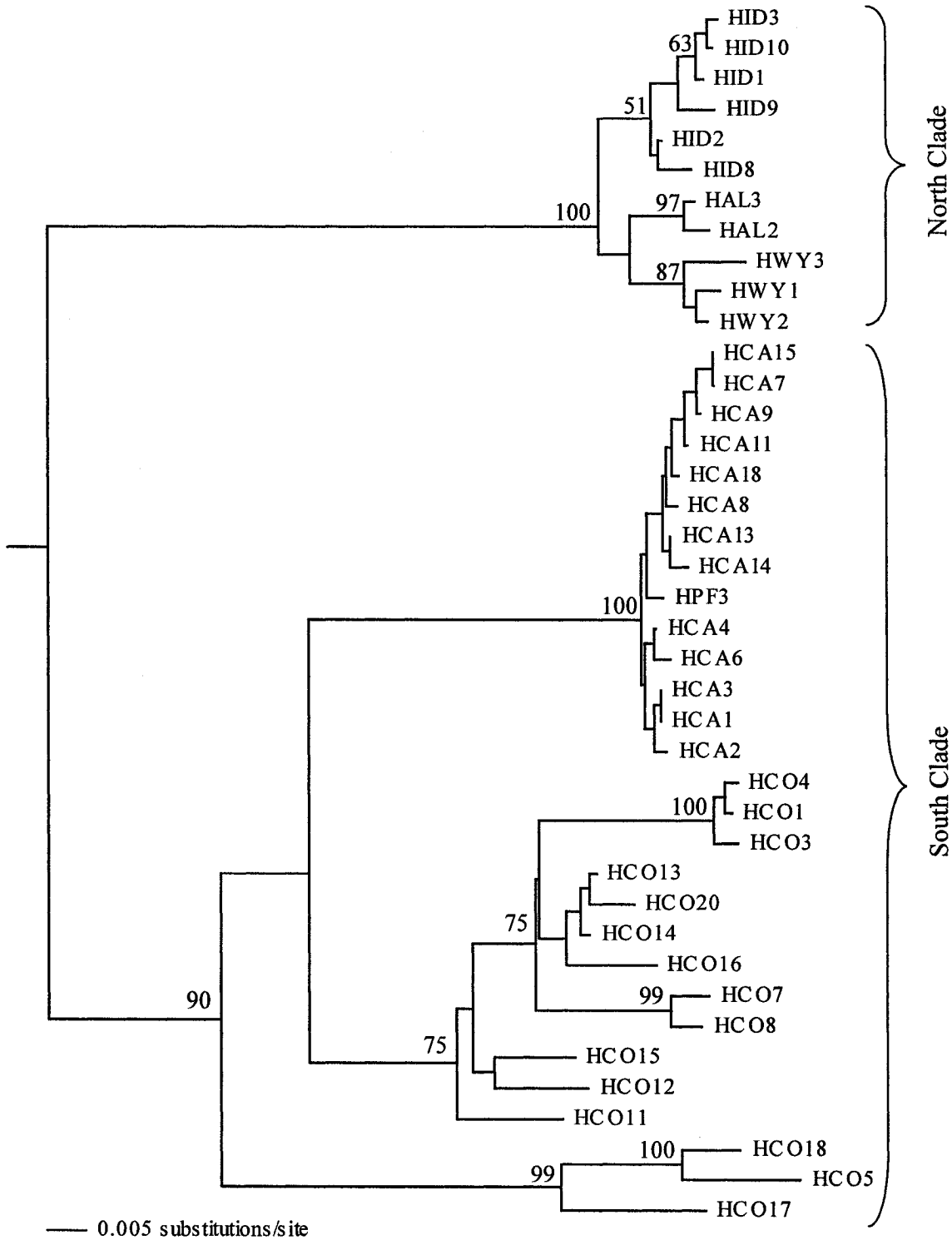
placed it between HCO4 and HCO1 haplotypes (not shown). This placement is highly unlikely as explained in the discussion and therefore midpoint rooting was used instead.

Figure 4
Phylogeny reconstruction based upon mtDNA cytochrome oxidase I of *Hesperodiaptomus shoshone* haplotypes using three phylogenetic methods: a) maximum parsimony, b) neighbour joining, and c) maximum likelihood. Haplotypes are shown at branch tips, further information on haplotypes and populations are given in Table 1 and Table 6. Numbers next to nodes are bootstrap values based on 1000 bootstrap replicates for maximum parsimony and neighbour joining and on 500 bootstrap replicates for maximum likelihood.

a) Maximum Parsimony



b) Neighbour Joining



c) Maximum Likelihood

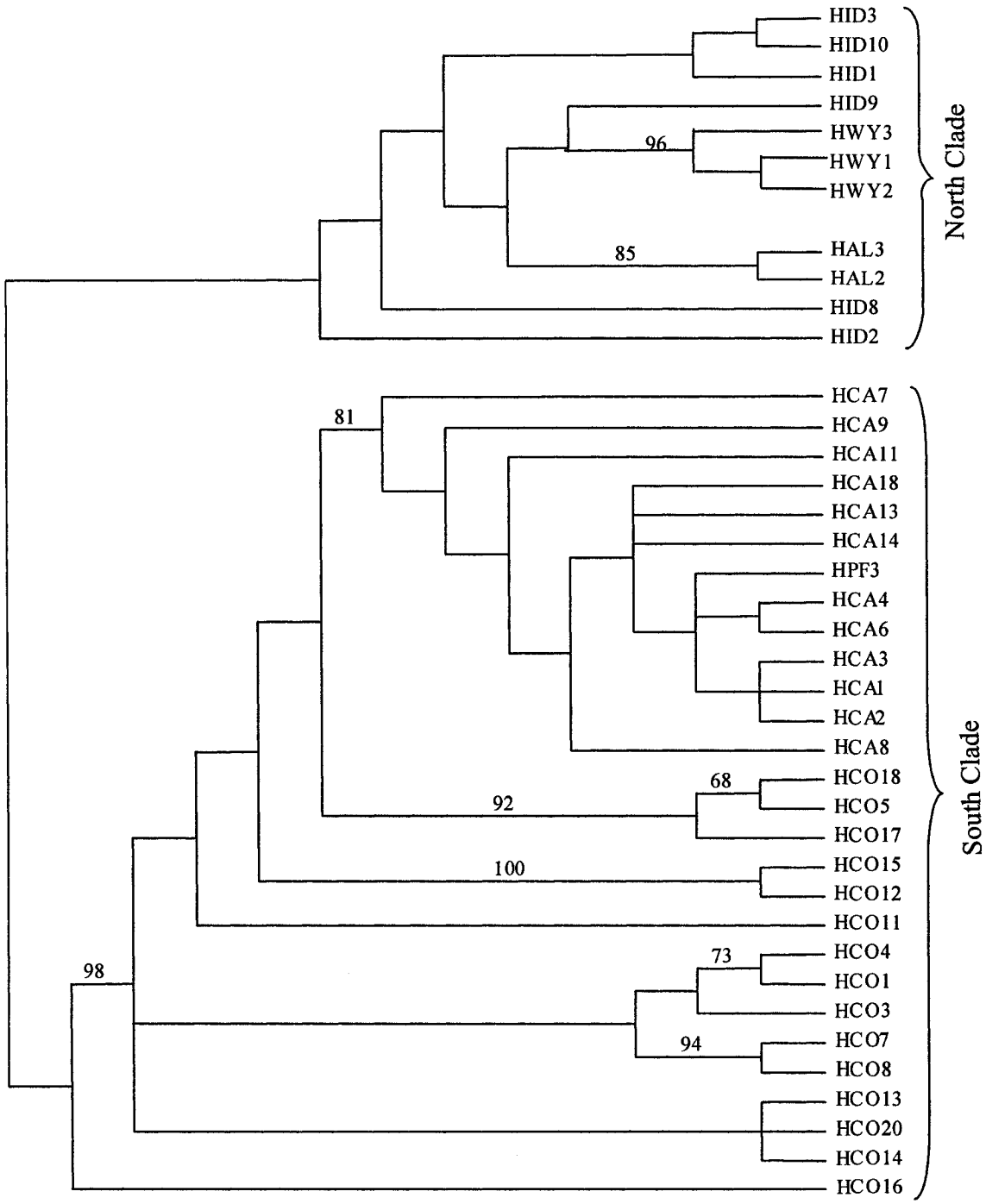
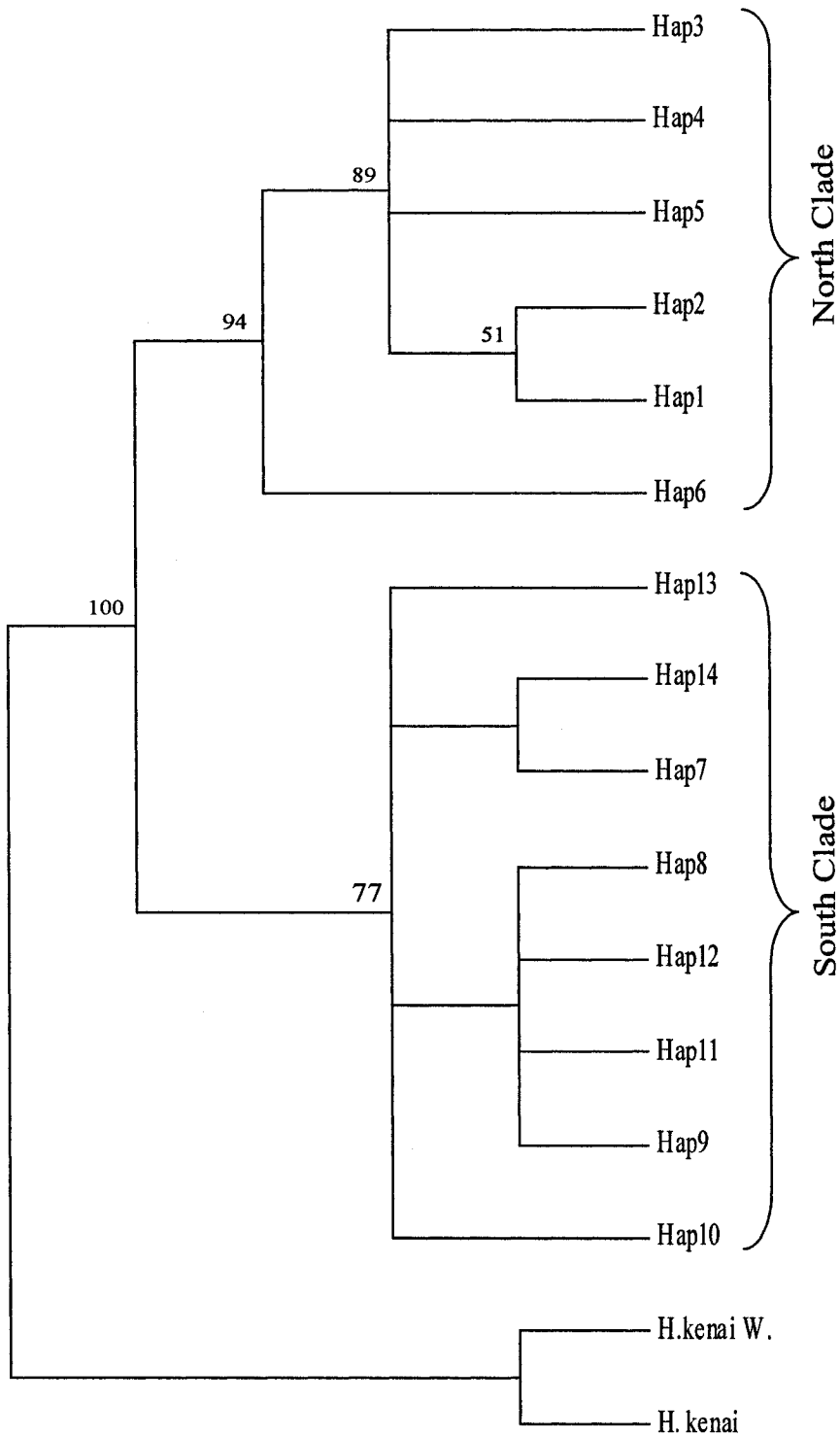
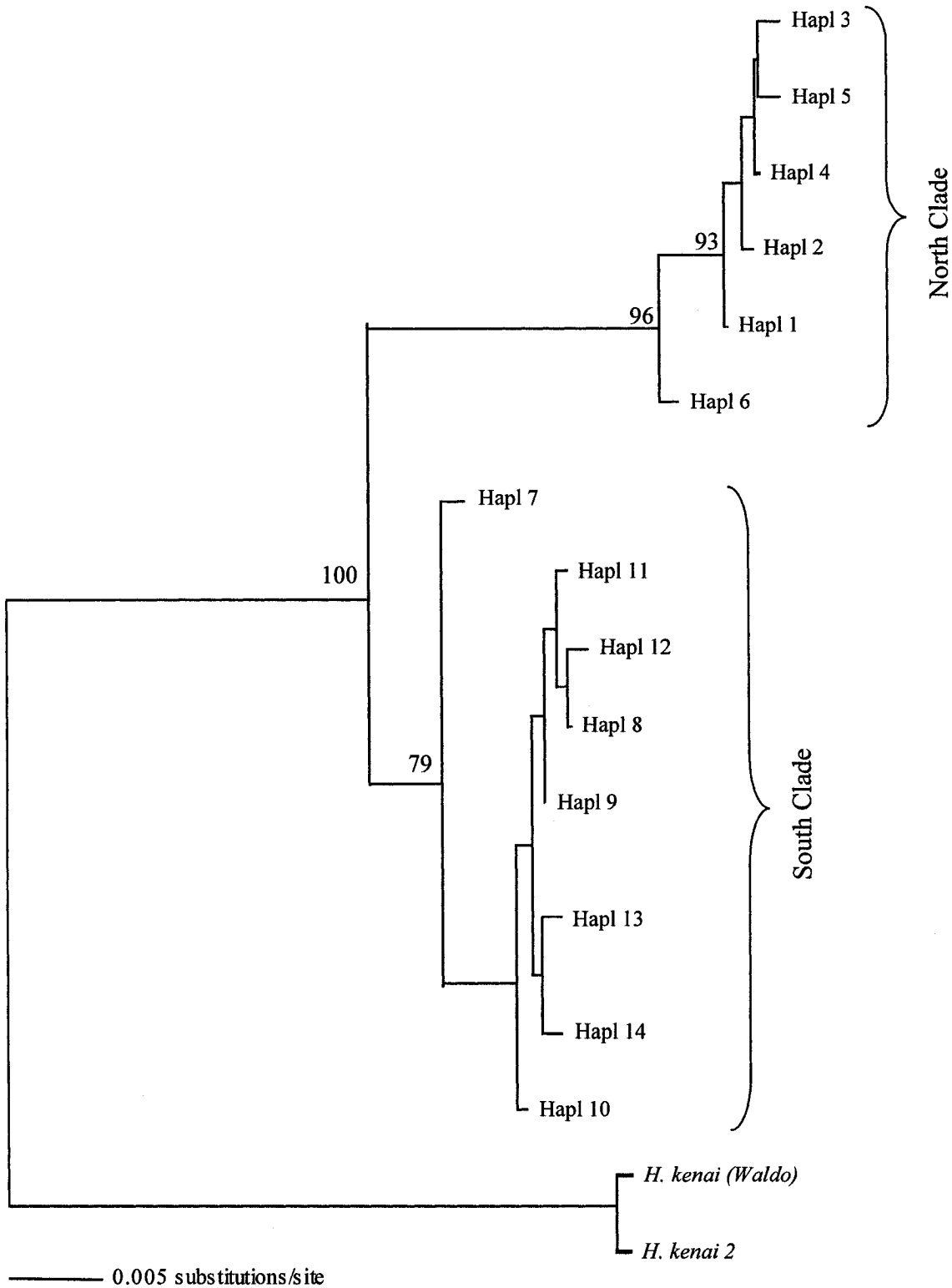


Figure 5
Phylogeny reconstruction based upon nuclear rDNA of *Hesperodiaptomus shoshone* haplotypes using three phylogenetic methods: a) maximum parsimony, b) neighbour joining, and c) maximum likelihood. Haplotypes are shown at branch tips, further information on haplotypes and populations are given in Table 1 and Table 7. Numbers next to nodes are bootstrap values based on 1000 bootstrap replicates for maximum parsimony and neighbour joining and on 500 bootstrap replicates for maximum likelihood.

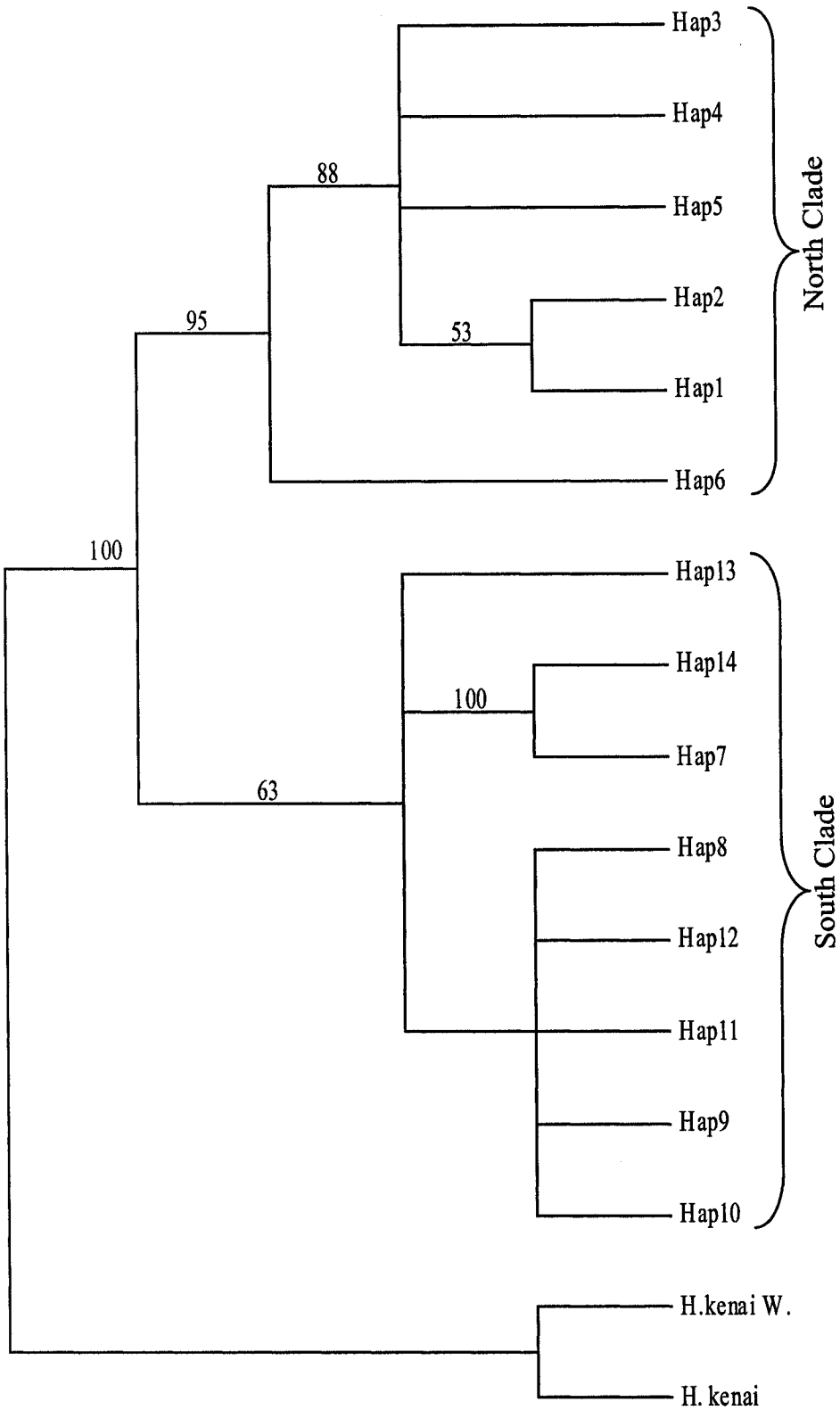
a) Maximum Parsimony



b) Neighbour Joining



c) Maximum Likelihood



Nested clade analysis

Population history inferred from mitochondrial COI gene sequence data

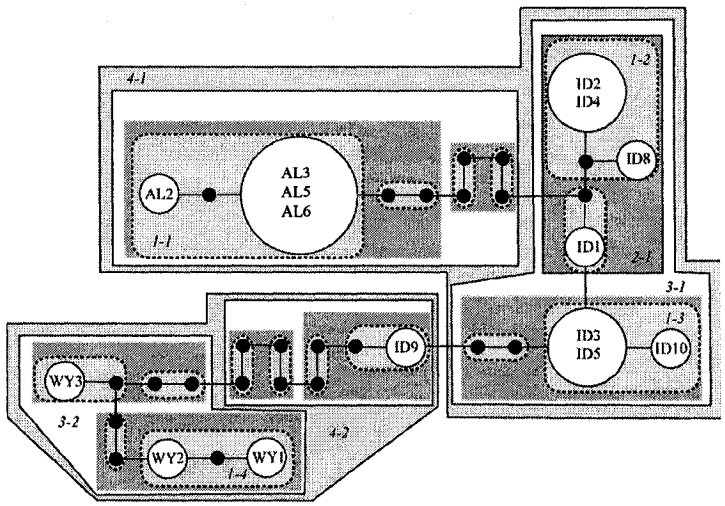
According to the 95% confidence limit of parsimonious connections (p = 0.953) as revealed by the estimation procedure (Templeton *et al.*, 1992) implemented in TCS version 1.21 (Clement *et al.*, 2000), the network-building algorithm joined haplotypes up to four mutational steps apart, resulting in multiple networks as well as isolated haplotypes. The mtDNA based networks and the isolated haplotypes could not be reliably joined due to the large number of mutations separating the networks and single haplotypes (1 to 102 mutations). The resolution obtained through NCA deteriorates with increasing divergence (Templeton *et al.*, 1992). Therefore, the highly divergent mtDNA sequences were analysed as three groups, based upon geographical and phylogenetic grouping (North Clade, Colorado and California).

The large divergence based upon the mtDNA sequences in the Colorado populations prevented the reconstruction of a single network and resulted in 3 small networks and 6 unique haplotypes (HCO5, HCO11, HCO12, HCO15, HCO17, HCO18). The uncorrected sequence divergence between these haplotypes ranged from 2 to 9 % with a mean of 6 % (std. dev. 2.6) (Table 3). Therefore the mtDNA sequences of Colorado populations were removed from the nested clade analysis. The California populations were connected with 95% parsimony probability up to 10 mutational steps (p = 0.952). The North Clade sequences were joined into a single network with 10 mutational steps (p = 0.951). Figure 6b shows the intraspecific networks and nesting designs obtained using the nesting algorithm given in Templeton *et al.* (1987) for the mtDNA North Lineage and California clades.

The null hypothesis of no association between geographic location and haplotype position within a clade was rejected for 6 clades (Table 8). The most likely biological processes that might have caused the observed significant associations for all of these clades are described in Table 9 and 10. These include allopatric fragmentation, and contiguous range expansion.

Figure 6
The nested haplotype networks based upon mitochondrial COI gene sequence data of *Hesperodiaptomus shoshone* clades. Haplotype connections are based on a P > 0.95 probability of being parsimonious. Only the north lineage and California haplotype are shown due to the large divergence among Colorado sequences which prevented a reconstruction of a haplotype network. Branches connecting haplotypes represent single substitutions. Filled circles represent unsampled haplotypes. The size of the circle is proportional to the number of populations contained in a given haplotype. Hierarchical nesting levels are denoted by boxes and numbered clades.

a) North Lineage



b) California populations

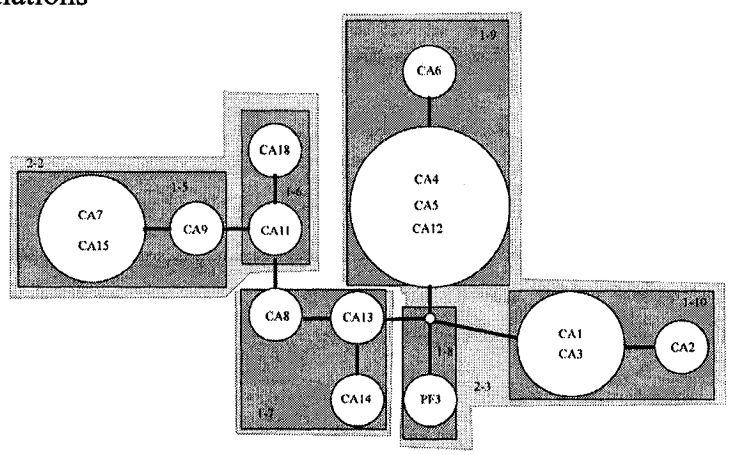


Table 8
Nested contingency analysis of geographical associations of mtDNA data from *Hesperodiaptomus shoshone*. Clades not showing genetic or geographical variation are excluded (no test is possible within such nested categories).

Clade	Permutational χ^2 statistic	Probability
North Clade		
Clade 3-1	12	0.02
Clade 4-1	20	0.001
Total Cladogram	24	0.01
California		
Clade 2-2	10	0.04
Clade 2-3	66	< 0.01
Total Cladogram	32	< 0.01

Table 9

Results of GEODIS analysis showing: clade (D_c), nested (D_n), and interior-tip clade (I-T) distances for the North Lineage based upon cytochrome oxidase I gene sequences. Only clades with a significant geographical/genetic association are shown. Superscripts represent associations that are significant at the 5% level: L significantly large, s significantly small). Chain of inference refers to the key provided by Templeton (2004), numbers indicate steps followed in the inference key to reach the conclusion. D_c, D_n and I-T represent the geographical dispersion of a clade, the distance between all individuals with haplotype x from all those with haplotype y, and average distance between tip and interior clades, respectively.

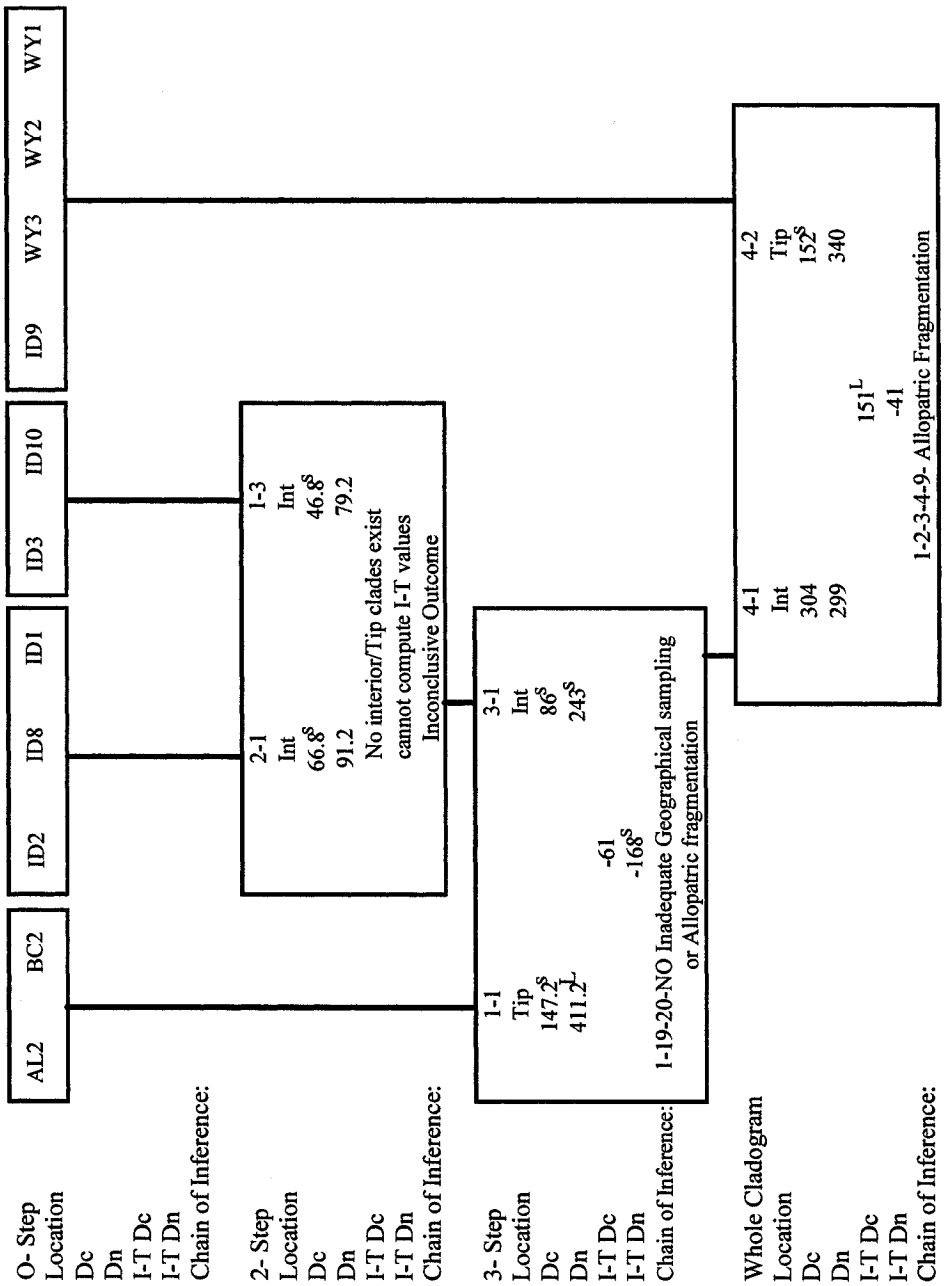
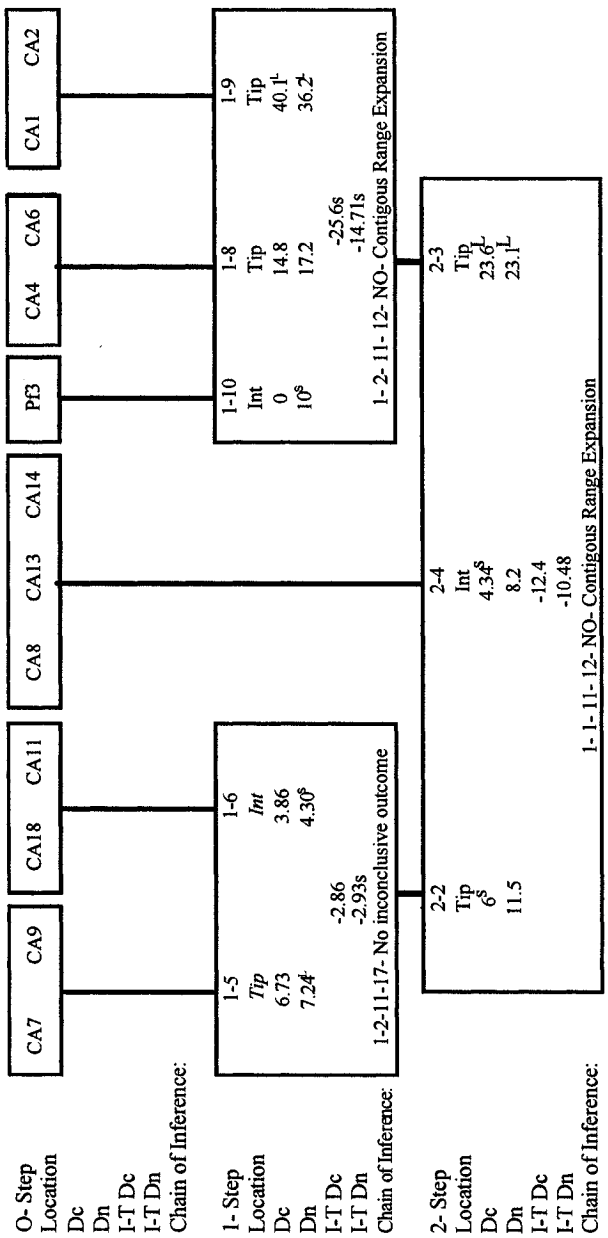


Table 10

Results of GEODIS analysis showing clade (D_c), nested (D_n), and interior-tip clade (I-T) distances based upon mitochondrial cytochrome oxidase I sequences of California populations. Only clades with a significant geographical/genetic association are shown. Superscripts represent associations that are significant at the 5% level: L significantly large, s significantly small). Chain of inference refers to the key provided by Templeton (2004), numbers indicate steps followed in the inference key to reach the conclusion. D_c, D_n and I-T represent the geographical dispersion of a clade, the distance between all individuals with haplotype x from all those with haplotype y, and average distance between tip and interior clades, respectively.



Population history inferred from rDNA

According to the 95% confidence limit of parsimonious connections (P = 0.956), the network-building algorithm joined haplotypes up to 11 mutational steps apart, resulting in one network for the rDNA data. The nested design resulted in four nesting levels (Figure 7). In the NCA, the null hypothesis of no geographical association among clades was rejected at different nesting levels (Table 11) showing that both population structure and population history played important roles in determining the current distribution of rDNA haplotypes. The results obtained by the inference key (Templeton, 2004) are given in Table 12. At the lower nesting levels insufficient genetic resolution prevented discrimination between range expansion, colonization and restricted dispersal/gene flow. At the highest nesting level allopatric fragmentation was inferred.

Figure 7
The nested haplotype network for rDNA *Hesperodiaptomus shoshone* clades. Haplotype connections are based on a P > 0.95 probability of being parsimonious. Branches connecting haplotypes represent single substitutions. Filled circles represent unsampled haplotypes. The size of the circle is proportional to the number of populations contained in a given haplotype. Hierarchical nesting levels are denoted by boxes and numbered clades. Descriptions of populations included in each haplotype are given in Table 7.

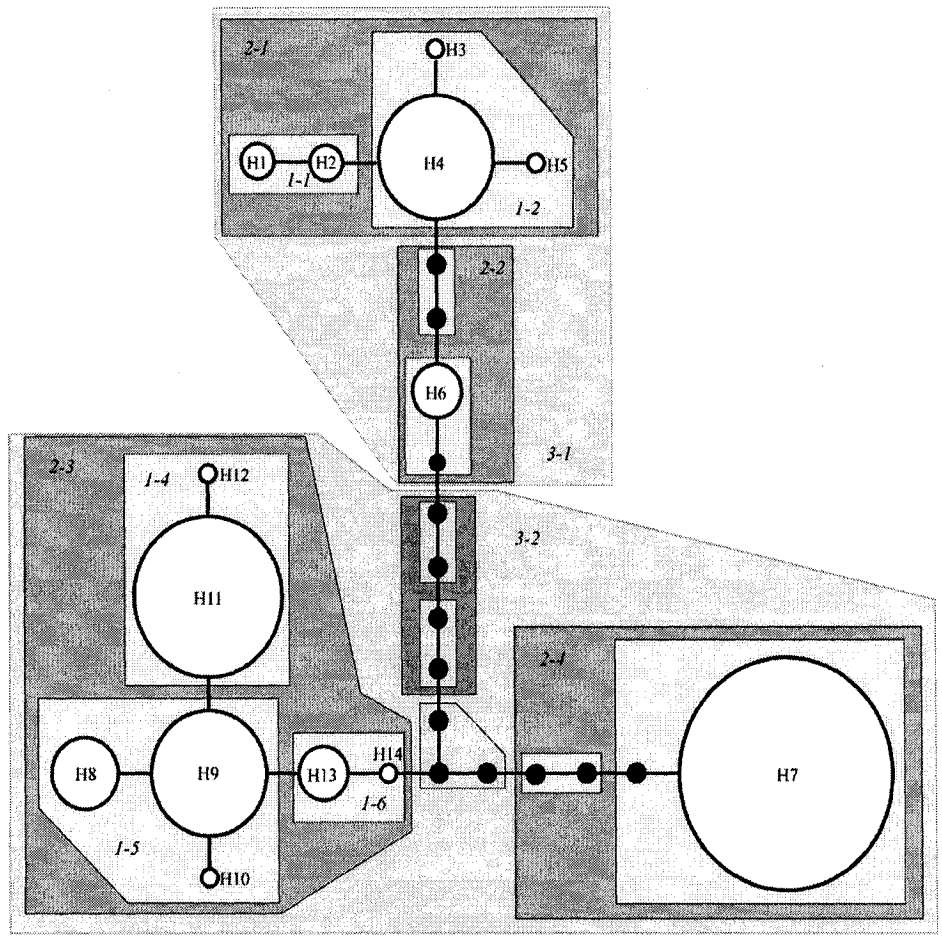
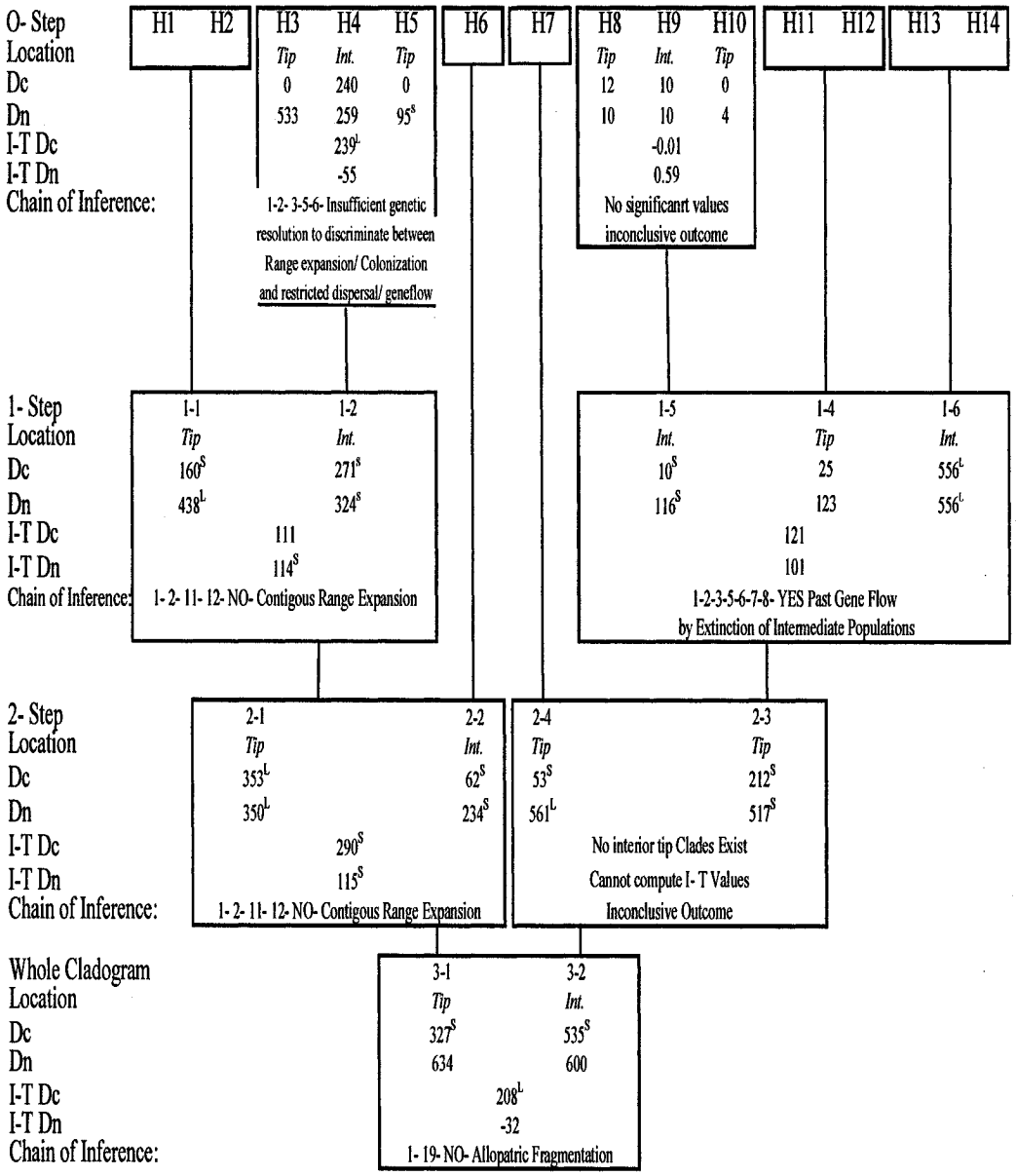


Table 11
Nested contingency analysis of geographical associations for nuclear rDNA data from *Hesperodiaptomus shoshone*. Clades not showing genetic or geographical variation are excluded (no test is possible within such nested categories).

Clade	Permutational χ^2 statistic	Probability
1-2	1.00	0.004
1-5	33.85	0.004
1-6	24.00	< 0.001
2-3	89.70	< 0.001
3-1	30.00	0.001
3-2	78.00	< 0.001
Total Cladogram	108.00	< 0.001

Table 12

Results of GeoDis analysis showing clade (D_c), nested (D_n), and interior-tip clade (I-T) distances for rDNA sequences. Only clades with a significant geographical/genetic association are shown. Superscripts represent associations that are significant at the 5% level: L significantly large, s significantly small). Chain of inference refers to the key provided by Templeton (2004), numbers indicate steps followed in the inference key to reach the conclusion. D_c, D_n and I-T represent the geographical dispersion of a clade, the distance between all individuals with haplotype x from all those with haplotype y, and average distance between tip and interior clades, respectively.



Mismatch Distribution

The mismatch distribution for both all mtDNA and rDNA populations is multimodal and significantly different from a hypothesis of sudden population expansion (mtDNA: $z = 2.174$, $p < 0.001$; rDNA: $z = 1.250$, $p < 0.08$) indicating stable population size over the range of the species. The results for the mismatch distributions for the full dataset and different regions are reported in Figures 8 and 9 for the mtDNA and rDNA respectively.

The hypothesis of a sudden population expansion is supported by the mtDNA and rDNA mismatch distributions for the California populations with no significant difference between the observed and expected distribution under the expansion model (mtDNA: $z = 0.788$, $p = 0.564$; rDNA: $z = 0.690$, $p = 0.722$) (Figure 8d and Figure 9d). The north clade mismatch distribution suggests a signal of allopatric fragmentation for the mtDNA sequences ($z = 1.96$, $p < 0.001$, Figure 8b) and a signal of population expansion for the rDNA sequences ($z = 1.069$, $p = 0.187$, Figure 9b). The multimodal Colorado mismatch distribution is indicative of a stable population in the mtDNA data ($z = 1.795$, $p < 0.003$, Figure 8c). Since the rDNA sequences from Colorado are represented by very few haplotypes (see Table 7) the mismatch distribution cannot be reliably interpreted for this region (Figure 9c).

A significant large negative F_S value, indicative of recent population expansion, was only detected in the rDNA California population. The more conservative Tajima's test indicated no significant deviation from neutrality in the mtDNA and rDNA analyses except for the rDNA Colorado populations.

Table 13
Statistical tests of neutrality and demographic parameter estimates for *Hesperodiaptomus shoshone* COI mitochondrial and rDNA data sets. Significant values of $p < 0.05$ are indicated by asterisks and of $p<0.01$ by double asterisks.

	All populations	**North Clade**	**Colorado**	**California**
mtDNA				
Tajima *D*	1.146	0.2407	-0.461	1
Fu *Fs*	-3.516	-2.713	-1.711	-6.809**
rDNA				
Tajima *D*	2.046	-0.419	-1.958*	0.182
Fu *Fs*	1.894	-0.888	2.207	-1.508

Figure 8
Mismatch distribution of the mitochondrial COI DNA sequences from all the data and the three main geographically defined *Hesperodiaptomus shoshone* groups. The observed distribution is represented by vertical bars. The expected distribution as predicted by the "sudden expansion" model of Rogers (1995) is depicted by the continuous line. The Kolmogorov-Smirnov test statistic and significance values are shown in the top right corner of each graph.

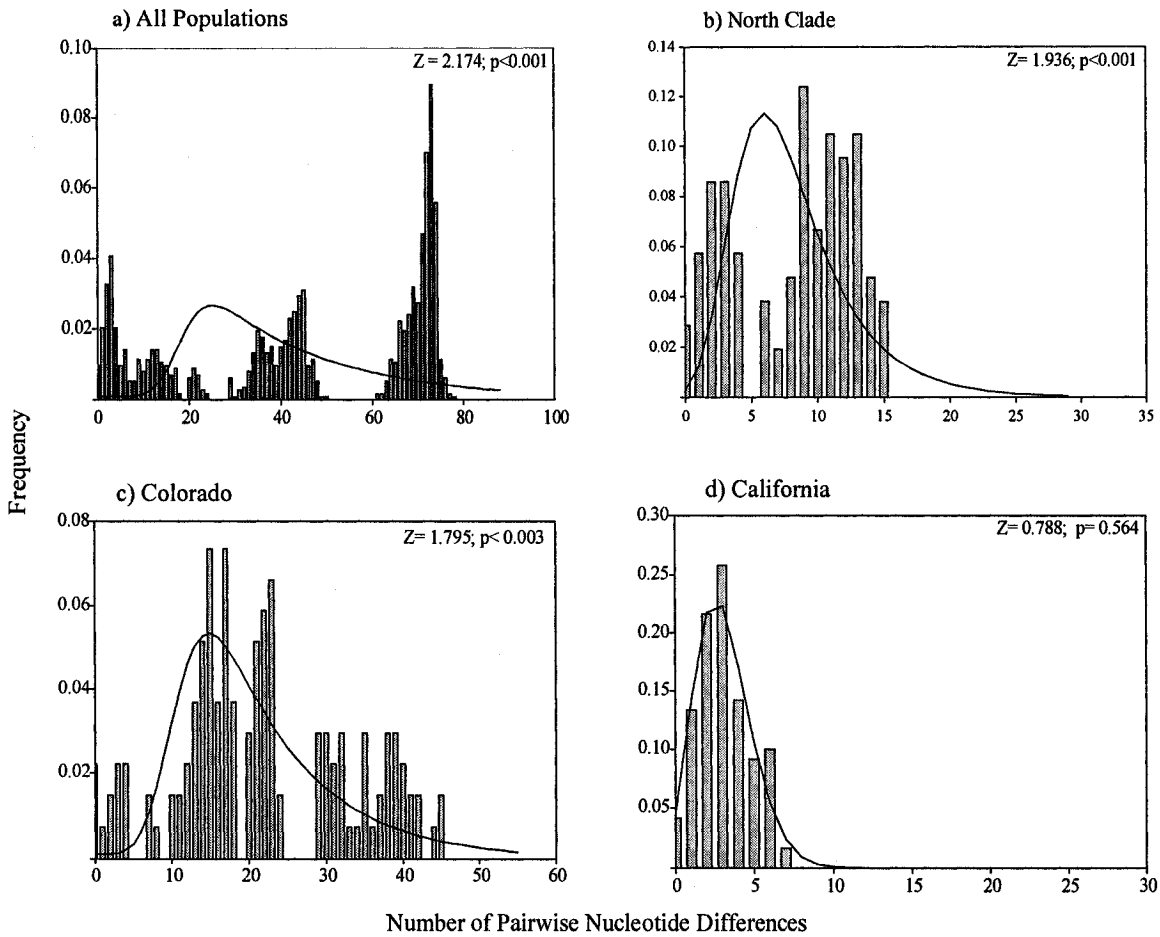
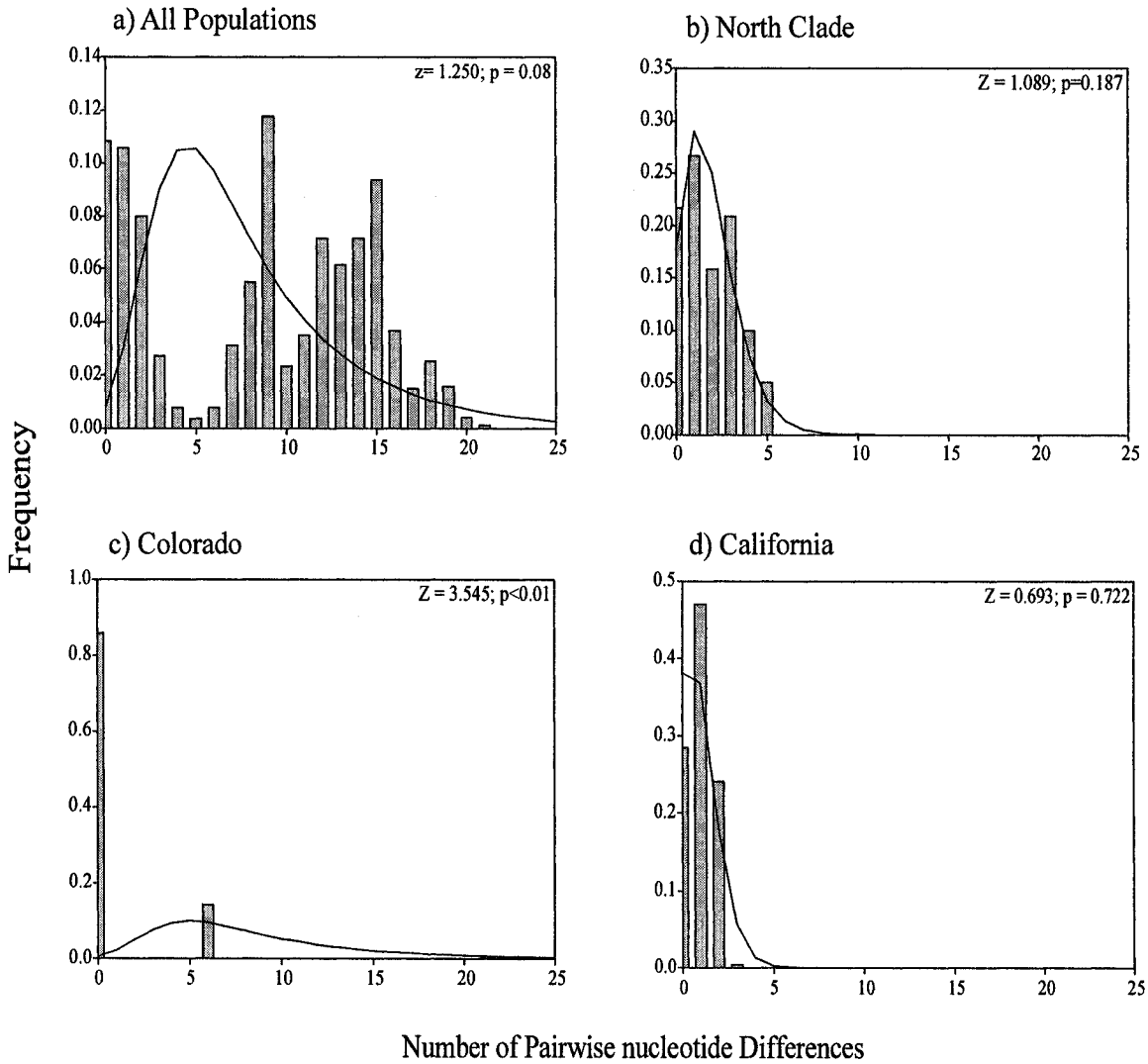


Figure 9
Mismatch distribution for the rDNA sequences from all the data and the three main geographically defined *Hesperodiaptomus shoshone* groups. The observed distribution is represented by vertical bars. The expected distribution as predicted by the "sudden expansion" model of Rogers (1995) is depicted by the continuous line. The Kolmogorov-Smirnov test statistic and significance values are shown in the top right corner of each graph.



Isolation by Distance

Mantel tests revealed a significant association between geographic and genetic distances in mtDNA (R^2=0.529, $p < 0.01$) and rDNA (R^2= 0.482, $p < 0.01$) datasets at the spatial scale of this study (Table 14 and Figures 10 and 11). However, diverse signals of isolation by distance were revealed by focusing on smaller geographic scales. In the less divergent rDNA no significant isolation by distance correlations were observed in the north clade or the Colorado and California populations. In the mtDNA data a significant correlation between geographic and genetic distance was found in the Colorado populations ($R^2 = 0.405$, $p < 0.01$). When the north eastern Colorado populations (CO1, CO3, and CO4) were excluded from the analysis, the association in Colorado populations became significantly stronger ($R^2 = 0.766$). No a priori rationale exists for determining the limits of isolation by distance analysis. Thus various methods of population grouping are usually explored (Kuchta and Tan, 2005).

Table 14
Mantel Test results for matrix correlation between genetic distance and geographic distance for mtDNA and rDNA sequences, with 10 000 randomizations. In the mtDNA Colorado clade followed by an asterisk, the north eastern Colorado populations (CO1, CO3, and CO4) were excluded from the analysis

	Z	R^2	p
rDNA			
All Data	35.79	0.482	< 0.01
North Clade	0.625	0.008	0.374
Colorado (all)	0.392	0.035	0.330
California	0.281	0.003	0.508
mtDNA			
All Data	323.11	0.529	<0.01
North Clade	4.994	0.366	<0.01
Colorado (all)	11.278	0.405	<0.01
Colorado *	6.177	**0.766**	**<0.01**
California	1.012	0.001	0.598

Figure 10
Isolation by Distance plots for the mitochondrial COI gene sequences from all groups and the three main geographically defined *Hesperodiaptomus shoshone* groups. a) all populations, b) North Clade, c) Colorado, d) California.

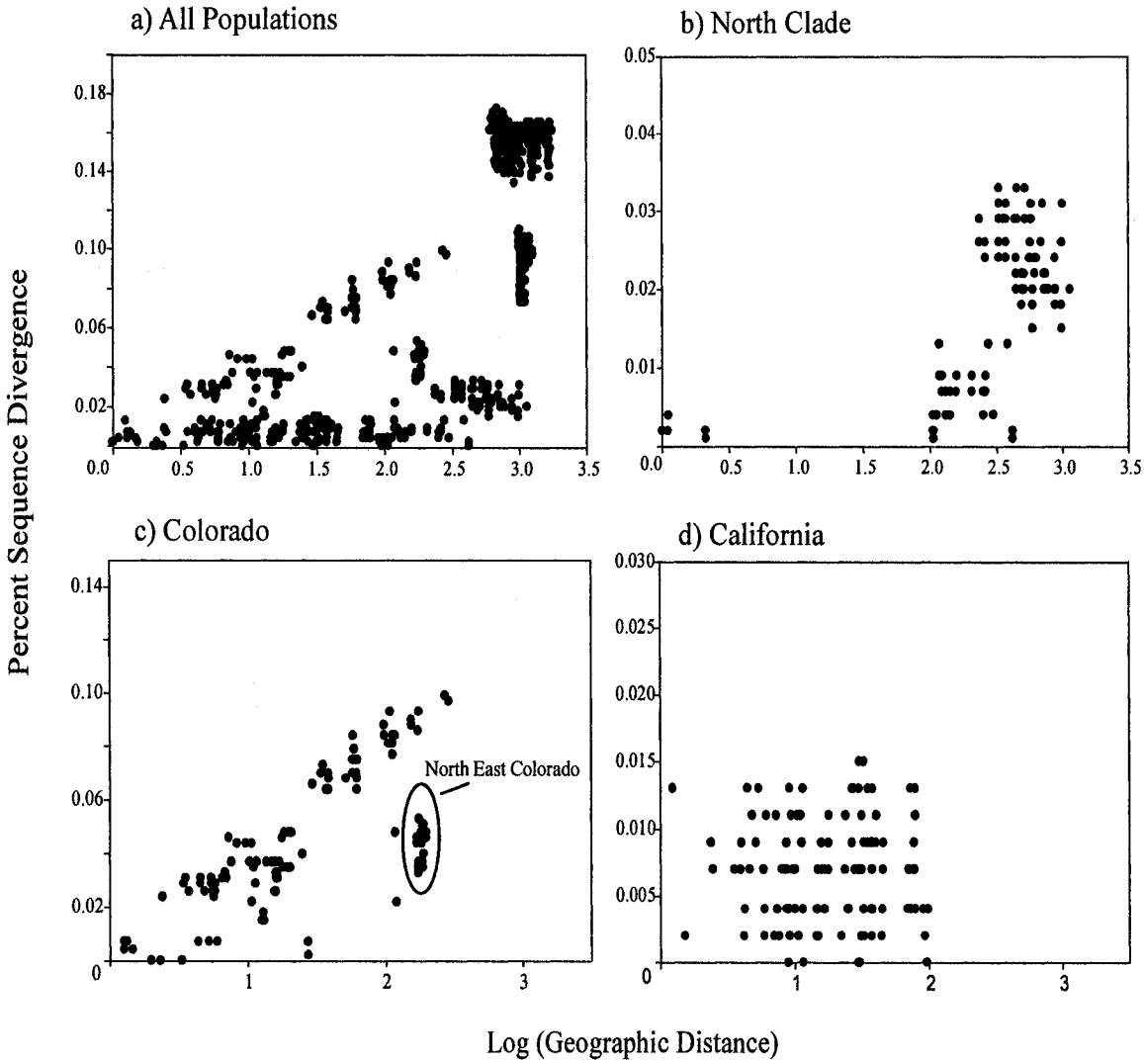
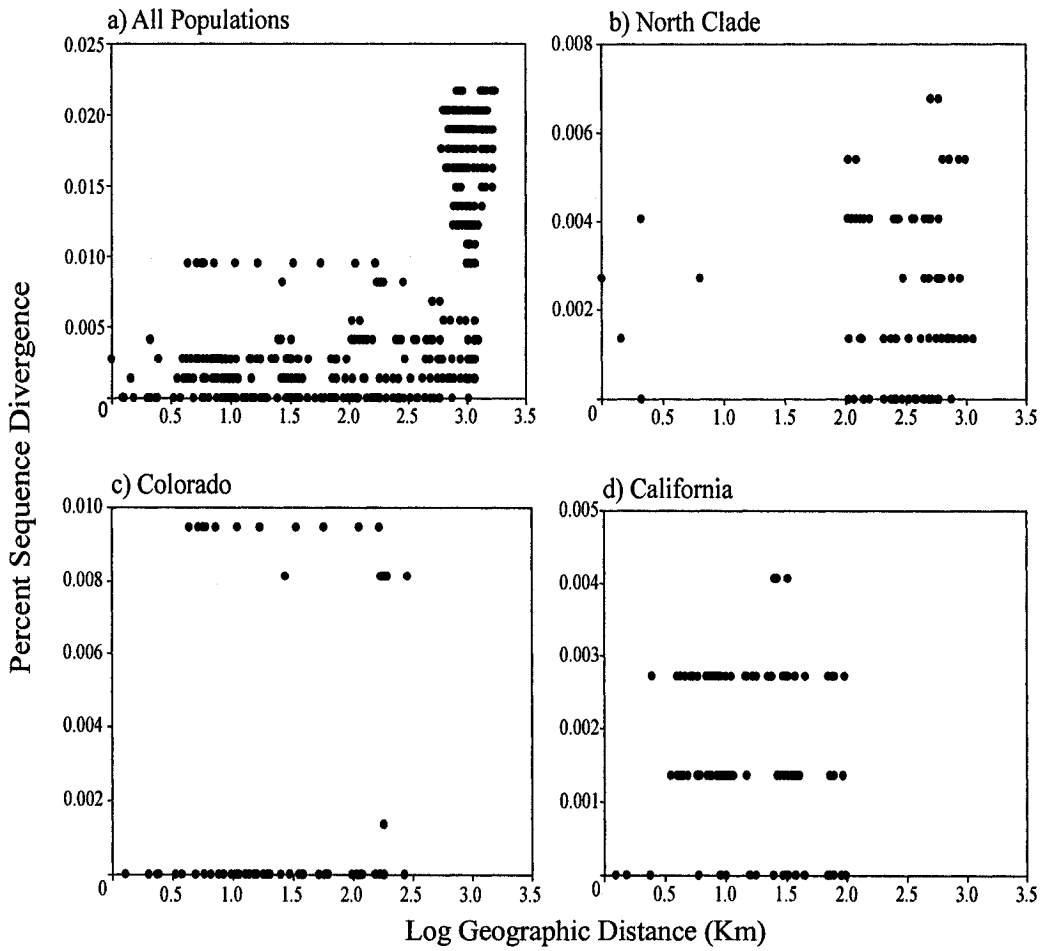


Figure 11
Isolation by Distance plots for the rDNA sequences from all populations and the three main geographically defined *Hesperodiaptomus shoshone* groups. a) all populations, b) North Clade, c) Colorado, d) California.



Discussion

A complex phylogeographic and phylogenetic pattern was observed within the North American H. *shoshone* populations. The results of this study reveal a genetically structured distribution and identify two reciprocally monophyletic groups. Judging from their levels of sequence divergence, the clades most likely represent long separated lineages. The main lineages show strong concordance with major geographic areas of Northwestern North America: The North Lineage is restricted to Alberta, Idaho and Wyoming, while the South Lineage is only found in California and Colorado.

An accumulation of phylogeographic animal and plant data in recent years has allowed comparison of the phylogeographic structure among codistributed species. Three broad distributional patterns characterize Northwestern North American endemics; widespread in the northwest; restricted to the Rocky Mountains; and disjunct, with populations in the coastal ranges and separate populations in the Rockies. Most of the studies to date have concentrated on the Northern Rockies and coastal ranges and have not included the Southern Rockies distribution (see review by Brunsfeld *et al.*, 2001). The distribution of *H. shoshone* is unusual since it occurs only at high altitudes, no populations occur in the coastal ranges of the Cascades and the distribution extends to the Southern Rocky Mountains. Furthermore, very few studies have examined the historical biogeography of invertebrates in this region and such a study provides a model of passive dispersal for montane freshwater invertebrates.

Differences between molecular markers

Molecular phylogenies and phylogeographic analyses are commonly used to infer evolutionary history, often in order to understand the effects of historical events on the evolution of organisms. Mitochondrial markers have been predominately used to examine these questions. However, the advantages of using multiple, independent markers have become increasingly evident (Arbogast *et al.*, 2002; Knowles and Maddison, 2002), given that single markers may have histories nonconcordant with the organismal history due to stochasticity or selection on a marker. The two markers used in this study provide important complementary information and demonstrate the limitations of each marker for this type of analysis. While the mitochondrial markers were too divergent for some analyses, the rDNA sequences did not provide enough resolution.

The extreme mitochondrial sequence divergence (up to 17.3 %, Table 3) found in this study among conspecific populations of *H. shoshone* is noteworthy, but not unusual among some copepod species. Interpopulation allozyme divergence was found to be about an order of magnitude higher in harpacticoid copepods than in many other species studied (Burton and Lee, 1994) and mitochondrial differentiation has been shown to exceed 18% (Burton, 1998; Edmans, 2001) in *Tigriopus californiensis*. Interestingly the sequence divergence between the north and south lineages is comparable to the divergence of its sister species, *H. kenai*, casting doubts on its integrity as a single species and suggesting the possibility of occurrence of two cryptic species. Additionally, the large sequence divergence in the mtDNA gene sequences resulted in problems with reconstruction of the maximum parsimony network and with the consistency of accurate placement of the *H. kenai* outgroup. Furthermore, such large sequence divergence prevents an accurate calibration of the molecular clock, especially

since no fossil records exist for *H. shoshone*. The best means of calibrating the molecular clock for *H. shoshone* COI mtDNA is from data for geminate species of shrimp on either side of the Isthmus of Panama (Knowlton *et al.*, 1993; Edmans, 2001). This calibration gives an estimated COI mitochondrial substitution rate of 2.2– 2.6% per million years for species spanning a similar range of divergence (6–20%). Using this estimation, it would seem that the North Lineage may have separated from the South Lineage during the mid-Miocene to early Pliocene (6–8 Mya). However, since no fossil records are available, such divergence time estimation should be interpreted cautiously. High divergence in mtDNA in *H. shoshone* might be due to high mutation rates, perhaps caused by the high UV alpine environments occupied by the species or, perhaps, by frequent bottlenecks which may accelerate genetic drift.

On the other hand, the limited variation in the rDNA sequences of *H. shoshone* resulted in a poor resolution in the intraspecific evolutionary analysis and phylogeography of this species. Although nuclear rDNA is expected to be less variable, the comparatively low rDNA genetic divergence of the Colorado populations (3 haplotypes, Table 7) is particularly contradictory to the highest amount of genetic variability found in the mitochondrial DNA in Colorado populations. The rDNA sequences have been used frequently in molecular systematics because they include both highly conserved and highly variable sequences. Despite their popularity, they do not always evolve in a simple manner. Often, when intra- and interspecific sequences are compared, interesting patterns emerge, with a high degree of sequence similarity within species, yet a great deal of divergence between them (Page and Holmes, 1998). This pattern is often explained by concerted evolution and arises because genetic mechanisms, particularly those of unequal crossing over and gene conversion,

transfer DNA sequences between genes so they evolve together. Concerted evolution is an important process in multigene families since it implies that mutations can spread to all members of the species. Therefore, the point of sequence divergence does not necessarily correspond to a speciation event but could correspond to the time of the last gene conversion or unequal crossing over. Such an event could explain the observed inconsistency of sequence variability between the mitochondrial and ribosomal gene sequences in the Colorado populations.

Variation may also occur when concerted evolution is not fast enough to homogenize different copies in the face of high rates of mutation and / or gene flow. Both the Northern lineage and California populations experienced recent expansions and higher gene flow compared to the more isolated Colorado populations, resulting in the observed pattern.

Phylogenetic reconstruction

The phylogenetic reconstructions reveal two highly supported reciprocally monophyletic clades that show concordance with major geographic areas of Northwestern North America: The North Lineage is restricted to Alberta, Idaho and Wyoming, while the South Lineage to California and Colorado. Except for the two major lineages, all mtDNA and rDNA trees generated exhibited largely unresolved topologies with low bootstrap support, irrespective of the method (parsimony, neighbour joining, or maximum likelihood). In the South Lineage, the monophyly of the California clade is highly supported by bootstrap replicates in the mtDNA dataset while no such pattern is found in the rDNA sequences. Lack of support for the monophyletic origin of the Colorado populations in both mtDNA and

rRNA sequences points to the possibility of long distance dispersal between Colorado and California populations.

The large divergence of the mtDNA gene sequences resulted in alternate root placement depending on the phylogenetic reconstruction method employed. The mtDNA maximum likelihood analysis placed the root of the *H. shoshone* phylogeny between WY1 and WY2 populations while the neighbour joining analysis placed it between CO4 and CO1 populations (not shown). This placement was considered incorrect and midpoint rooting (Figure 4) was used instead for the following reasons. First, the monophyly of the north and south clades is well supported by mtDNA maximum parsimony and rDNA analyses. Second, NJ and ML bootstrap replicates do not support such root placement. Finally, the length of the branch connecting the ingroup to the outgroup is excessively large compared to any branches connecting the ingroup taxa. Under such circumstances, random rooting artefacts may arise due to long branch attraction (Huelsenbeck *et al.*, 2002). These lines of evidence suggest that the *H. kenai* might not be an appropriate outgroup despite being the closest taxon to *H. shoshone*. Therefore, due to lack of alternative outgroups, midpoint rooting was employed for the mtDNA maximum likelihood and neighbour joining analyses (Figure 4). Unlike the outgroup rooting, the midpoint rooting is consistent with available biogeographic and phylogenetic evidence. The topology of the midpoint-rooted tree preserves the monophyly of both lineages and places the root between north and south clades although the bootstrap support for both lineages is not as strong as in the maximum likelihood analysis.

Phylogeography

The nested clade analysis supports the phylogenetic reconstruction methods mentioned above. In addition, it identifies historical processes that might be responsible for the current observed genetic structure of *H. shoshone* populations. These analyses establish the presence of two main, divergent lineages in *H. shoshone* with strong geographical segregation forming southern and northern clades, which indicates a deep phylogeographical structure. The division into two lineages was probably produced through a past event of allopatric fragmentation (Avise, 2000) as inferred from the rDNA dataset and the large number of mutations separating the mtDNA sequence data.

Nested clade analyses performed on both markers are mostly concordant with a few differences that could be explained by the variability of the markers. Since mutation rates are lower in the nuclear rDNA than the mtDNA sequences, the analysis provides complementary inferences at different temporal scales. The historical processes inferred from the rDNA should represent much older events while those inferred from mtDNA more recent ones. Without accurate molecular dating, however it is impossible to infer where the inferences of the two markers should overlap.

Habitat contractions caused by Pleistocene glaciations have been proposed as a major process in reducing genetic diversity in northern latitudes (Hewitt, 1996). According to this hypothesis, loss of genetic diversity due to genetic drift in bottlenecked refugial populations would have been followed by rapid range expansions northward as ice sheets receded. If such expansions occurred too recently to accumulate point mutations, low genetic diversity and weak phylogeographical patterns would be observed today (Mila *et al.*, 2000). The mtDNA sequences from the North Lineage have a lower genetic diversity than the southern

Colorado populations indicative of recent post glacial expansions. Both mismatch distributions and nested clade analyses suggest that fragmentation and population expansions are responsible for the current genetic structuring of *H. shoshone* in the North Lineage.

Some contradicting *H. shoshone* distribution data might suggest a different scenario. The *H. shoshone* distribution is disjunct in Alberta. A dozen populations are known to exist in Jasper National park (including the two sampled populations AL3, and AL4), but interestingly no populations occur in Banff just south of Jasper, two populations in the southern Alberta foothills (AL1, AL2) and a few populations in Waterton Lakes National Park (AL5, AL6). Anderson (1967, 1971, 1974) sampled these areas meticulously for zooplankton thus the distribution of *H. shoshone* is very well known for this region. There is some evidence that all these sampling locations served as refugia for at least some aquatic organisms during the last glaciation. Crossman and McAllister (1986) referred to a possible Banff-Jasper refugium, and at least two other high valley areas in Jasper National Park are suspected to have been unglaciated during the Wisconsian (Hughes, 1955). Donald (1985) also suggested the presence of ice free areas in the Waterton Lakes Park area, possibly since the early Wisconsian glacial maximum.

One population from the North Lineage of particular interest occurs at Plateau Pond, Alberta (AL1). Numerous nunataks (mountains that were not covered by glaciers) are believed to have existed just west of the ice free corridor that separated the Cordilleran and Wisconsin ice sheets during the last glaciation. Valley glaciers descending eastward from the Cordilleran ice sheet would often flow around the lower mountains and hills east of the continental divide (Pielou, 1991). One of the refugia formed in this way is Plateau Mountain. There is substantial geological evidence that this mountain was not glaciated (see Beaty,

1975; Burns, 1980). Along with *H. shoshone,* a rare diaptomid, *Hesperodiaptomus victoriaensis* (Reed, 1958), also occurs in Plateau Pond (Anderson, 1967). When first found, this was one of just four known populations of *H. victorianensis* and the only one known outside the Arctic (Anderson, 1967). Further extensive sampling efforts by Anderson in the Southern Canadian Rocky Mountains, foothills and adjacent prairies failed to locate any other populations (Anderson, 1971, 1974). Thus, the population of *H. victoriaensis* in Plateau Pond is considered to be a true glacial relict. Since all the Alberta populations of *H. shoshone* probably survived the last glaciation in the multiple refugia mentioned above we should expect extensive genetic divergence in these isolated populations. This is not the case however, in the markers examined. Only two mtDNA haplotypes (HAL2 and HAL3; Table 6) separated by one mutation were found in the Alberta populations, although DNA from the Plateau pond (AL1) and Surprise Lake (AL4) individuals could not be successfully amplified and sequenced and they were not included in the mtDNA analysis. The rDNA sequences surprisingly proved to be more variable with four haplotypes (H1, H2, H3 and H4; Table 7) in the Alberta populations separated by 3 mutations, but the Plateau pond population did not represent a unique haplotype. Thus, while the geological data suggests the presence of refugia the low genetic diversity between those populations indicates recent expansions.

Another interesting pattern inferred from the northern lineage is the allopatric fragmentation between clade 4-1 and 4-2 of the mtDNA data separating Alberta and central Idaho from southern Idaho, Wyoming and the Beartooths. Allopatric fragmentation is also consistent with the bimodal signal observed in the mismatch distribution analysis. Such a pattern has been inferred for other taxa from the region, such as the chipmunk *Tamias*

amoenus (Demboski and Sullivan, 2003) and represents a break between the mountain ranges by the Snake River plane.

Populations in the South Lineage were less affected by Pleistocene habitat contractions and should show higher levels of nucleotide diversity, stronger phylogeographical structure, and a multimodal distribution of pairwise nucleotide differences (mismatch distribution) indicating long-term demographic stability. Such results were observed in the mtDNA data but not the rDNA dataset.

Major discrepancies exist between the Colorado mtDNA dataset and the nuclear rDNA sequences as discussed above. In the mtDNA data, the large sequence divergence resulted in three networks that could not be reliably joined. The interpretation for nested clade analysis is fragmentation which is supported by a multimodal mismatch distribution.

The California populations of *H. shoshone* occur only in the high Sierra Nevada. The Sierra Nevada was far from the two ice sheets during the last glaciation but the mountain range had a discontinuous ice cap that extended for 450 km and was 50 km wide in some areas (Norris and Webb, 1990). The current Sierra Nevada glacier population consists of approximately 498 ice features, including perennial ice patches and mountain glaciers (Norris and Webb, 1990). These features range in size from small ice patches with areas of less than 0.1 km^2, to Palisade Glacier, the largest in the Sierra Nevada, with an area of 1.58 km^2 (Norris and Webb, 1990). All the sampled and sequenced *H. shoshone* populations occur in close proximity to existing ice features (i.e. Matthers Glaciers, Goethe Glacier, and Palisade Glacier), in cirque or paternoster lakes formed by receding glaciers. Mitochondrial nested clade and mismatch distribution analyses suggest recent range expansions for these

populations consistent with the geology of the region. The rDNA nested clade analysis did not provide enough resolution to infer any historical interpretation for the region.

Nested clade analysis has limitations inherent to the sampling scheme, which depends on comprehensive sampling of the distribution range of a species. In this case, the distribution of *H. shoshone* in some regions is poorly known. The sampling sites for this study were obtained by anecdotal evidence, reports dating from the early 1900's, publications and personal accounts. Many of the sites where the species was present historically currently did not contain *H. shoshone*, possibly due to extinctions, temporal variability, fish stocking, misidentifications, or habitat alteration. For example, an account placing *H. shoshone* in the front range of eastern Colorado in a series of ponds and lakes in the valley of Beaver Creek, called the Seven Lakes, described by Ward (1904) proved to be unsuccessful. The valley was actually flooded and converted into reservoirs which currently provide water for Colorado Springs. These sampling limitations resulted in few cases of significant geographical associations of haplotypes for which inconclusive outcomes were obtained.

Isolation by distance and Demographic analyses

The accumulation of genetic differentiation with increased geographical separation, as a result of restricted dispersal relative to geographical range, was first explored by Wright (1943), who termed the phenomenon isolation by distance (IBD). IBD has been expanded upon and was used to develop methods for inferring equilibrium conditions between gene flow and genetic drift among natural populations (Hutchinson and Templeton, 1999). Under equilibrium conditions, the values and the variability of genetic distances should both increase with geographical distance, as gene flow counteracts the effect of drift (Turgeon and Bernatchez, 2001). Isolation by distance is expected between populations if gene flow is

prevalent but dependent on geographical distance, because the exchange of individuals should be more frequent among nearby populations than among populations located further away (e.g., Wright, 1946; Slatkin, 1993; Hutchison and Templeton, 1999). An absence of IBD is commonly interpreted to mean that populations have not reached equilibrium between gene flow and drift. Following a recent range expansion, populations established from the same source will be genetically similar reflecting a recent shared evolutionary history. In this scenario, little or no relationship should be detected between genetic and geographic distances coupled with low variance in divergence (Hutchinson and Templeton, 1999). If gene flow is prevalent in a region with low genetic drift, no patterns between genetic and geographic distances should also be observed. If the region becomes fragmented, or gene flow is very low, drift would become more influential than gene flow (Hutchinson and Templeton, 1999, McLean *et. al.,* 1999). In this scenario, the resulting pattern should reflect extreme isolation were populations experience drift without relation to geographic distances at which point a high degree of scatter in genetic distances should be observed.

A significant isolation by distance correlation was found in both data sets analysed. This relationship explains 48.2–54.5 % of the variation (Figures 10 and 11) and is consistent with the gene flow and drift equilibrium hypothesis proposed by Hutchinson and Templeton (1999). The lack of perfect fit can be explained by a number of factors operating within smaller geographical areas such as lack of equilibrium between gene flow and drift, natural selection, limits to the power of genetic markers or long distance dispersal (Hellberg, 1994)

When isolation by distance was performed for smaller geographical regions non equilibrium conditions indicative of past range expansions were observed in the recently glaciated areas of California in both mtDNA and rDNA analyses and for the north clade

rDNA (Figures 10 and 11). In Colorado a significant IBD pattern was observed for the mtDNA marker

The mismatch distribution of the observed number of differences between pairs of haplotypes was unimodal for the clades for which nested clade analysis and IBD indicated range expansions (Rogers and Harpending, 1992). The mismatch distribution for the North Lineage based upon mtDNA was bimodal, consistent with nested clade analysis and indicative of allopatric fragmentation. By contrast, the rDNA mismatch patterns in the North lineage were consistent with a range expansion further supported by nested clade and IBD analyses. The mtDNA mismatch patterns in the Colorado populations are consistent with expectations of a population unaffected by sudden demographic change supported by a significant isolation by distance pattern.

The significance of neutrality tests varied depending on the statistic and the region. These tests are not very powerful and were computed to verify the inferences of population expansion independently. Tajima D was significant and negative only for rDNA Colorado populations. Tajima's *D* statistic is based on the difference between two estimates of the amount of genetic variation. The estimate can be calculated using either the number of segregating sites or the average number of pairwise differences (Page and Holmes, 1998). In a constant-size neutral equilibrium population, the expectation of Tajima's *D* is nearly zero because the expectations of both estimates are the same. However, natural selection changes the structure of genetic variation within populations because alleles persist for different times under selection compared with drift (Page and Holmes, 1998). When some kind of balancing selection is acting, Tajima's *D* tends to be positive. On the other hand, selective sweeps and purifying selection can generate negative values of Tajima's *D*. However, changes in

population size can also affect Tajima's *D*. In a population with decreasing size, the expectation of Tajima's *D* is positive, while a negative Tajima's *D* is predicted for a population with increasing size. Therefore, when the observed value of Tajima's *D* deviates significantly from 0, it is very difficult to know what the cause of such a deviation is. In the case of the Colorado populations the negative value of Tajima D for rDNA are most likely due to selection rather than population expansion. The values of Fu and Fs were only significantly negative for California populations, indicating rapid population expansions.

This study reveals the advantage of using multiple analyses to explain a species' evolutionary history from molecular data. In this study, four main approaches were employed: phylogeny reconstructions, phylogeography (nested clade analysis), isolation by distance and demographic analyses (mismatch distributions, neutrality tests). No single analysis can provide a history as complete as did the combination of these techniques. For example, the phylogenetic reconstructions were helpful regarding broad genetic patterns, especially with the identification of the two main lineages. However, phylogenetic reconstructions provide little information about relationships among populations within each region, or about the processes that generated the genetic structure. Nested clade analysis was useful in identifying the likely processes causing the current genetic structure. Mismatch distribution and the other neutrality tests provide an independent test of the data to determine whether inferences from the nested clade analysis (such as range expansions) are supported. Yet, demographic tests alone can prove to be misleading because various micro-evolutionary forces leave similar genetic fingerprints (e.g., expansion and selection).

Conclusion

This study has assessed the phylogeographic relationships of *H. shoshone* populations inhabiting mountain ranges of western North America. It reveals a complex history of fragmentation, isolation by distance and population expansions. Three major patterns emerge in the *H. shoshone* phylogenetic and phylogeographic analyses. First, large genetic divergences in mtDNA and rDNA separate populations in the Northern Rocky Mountains from those in the Sierra Nevada and Southern Rockies. Each of these lineages represents a unique evolutionary unit. Although the molecular sequences indicate the presence of two cryptic species, further morphological and mating experiments are needed to make such conclusions. Second, Pleistocene glaciations had an important impact on the genetic structuring of *H. shoshone* populations. Third, a weak isolation by distance effect was observed suggesting limited dispersal between *H. shoshone* populations. Together, these results illustrate the complicated biogeographical history of freshwater crustaceans. Further studies on this and other taxa are necessary to test the generality of these results with respect to other freshwater crustaceans and aquatic invertebrates.

References

Anderson, R.S. (1967) Diaptomid copepods from two mountain ponds in Alberta. *Canadian Journal of Zoology* **45**:1043-1051.

Anderson, R.S. (1971) Crustacean plankton of 146 alpine and subalpine lakes and ponds in Western Canada. *Journal of Fisheries Research Board of Canada* **28**:311-321.

Anderson, R. S. (1974) Crustacean plankton communities of 340 lakes and ponds in and near the National Parks of the Canadian Rocky Mountains. *Journal of Fisheries Research Board of Canada* **31**:855-869.

Arbogast, B.S., S.V. Edwards, J. Wakeley, P. Beerli and J.B. Slowinski (2002) Estimating Divergence Time from molecular data on phylogenetic and population genetic time scales. *Annual Review of Ecology and Systematics* **33**(1):707-740.

Avise, J.C. (1994) Molecular markers, natural history and evolution. Chapman and Hall, New York.

Avise, J.C. (2000) Phylogeography : the history and formation of species. Harvard University Press, Cambridge, Mass.

Avise, J.C., C.A. Reeb, and N.C. Saunders (1987) Geographic population structure and species differences in mitochrondrial DNA of mouthbrooding marine catfishes (*Aridae*) and demersal spawning toadfishes (*Batrachoididae*). *Evolution* **41**(5):991-1002.

Beaty, C.B. (1975) The landscapes of southern Alberta: a regional geomorphology. PhD Thesis. University of Lethbridge, Alberta 95pp.

Beccerra, J. and D.L. Venable (1999) Nuclear ribosomal DNA phylogeny and its implications for evolutionary trends in the Mexican *Bursera*. *American Journal of Botany* **86**:1047-1057.

Bermingham, E. and J.C. Avise (1986) Molecular zoogeography of freshwater fishes in the southeastern United States. *Genetics* **113**(4):939-965.

Bernatchez, L. and C.C. Wilson (1998) Comparative phylogeography of Nearctic and Palearctic fishes.*Molecular Ecology* 7(4):431-452.

Brunsfeld, S.J., J. Sullivan, D.E. Soltis and P.S. Soltis (2001) Comparative phylogeography of Northwestern North America: a synthesis. In J. Silvertown and J. Antonovics (Eds.), Integrating ecological and evolutionary processes in a spatial context (p.319-339). Blackwell Science, Oxford.

Bohonak, A.J. (2002) IBD (Isolation By Distance): a program for analyses of isolation by distance. *Journal of Heredity* **93**:153-154.

Bucklin, A. (2000) Methods for Population Genetic Analysis of Zooplankton. Chapter 11. The Zooplankton Methodology Manual, International Council for the Exploration of the Sea. Academic Press, London:533 - 570.

Bucklin, A., R.C. Sundt and G. Dahle (1996) The population genetics of *Calanus finmarchius* in the North Atlantic. *Ophelia* **44**:29-45.

Bunje, P.M.E. (2005) Pan-European phylogeography of the aquatic snail *Theodoxus fluviatilis* (Gastropoda: Neritidae). *Molecular Ecology* **14**(14):4323–4340.

Burns, J.A. (1980) The brown lemming, *Lemmus sibiricus* (Rodentia, Arvicolidae), in the late Pleistocene of Alberta and its postglacial dispersal. *Canadian Journal of Zoology* **58**:1507-1511.

Burton, R.S. (1998) Intraspecific phylogeography across the point conception biogeographic boundary. *Evolution* **52**(3):734-735.

Burton, R.S. and B.N. Lee (1994) Nuclear and mitochondrial gene geneologies and allozyme polymorphism across a major phylogeographic break in the copepod *Tigriopus californicus*. *Proceedings of the National Academy of Sciences USA.* **91**:5197-5201.

Casgrain, P. and P. Legendre (2001) The R-Package for multivariate and spatial analysis and User's manual, Early draft. Département de sciences biologiques Université de Montréal. (http://www.fas.umontreal.ca/BIOL/legendre/) [accessed January, 2003].

Clement, M., D. Posada and K.A. Crandall (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**:1657-1659.

Cox, A.J. and P.D.N. Hebert (2001) Colonization, extinction, and phylogenetic patterning in a freshwater crustacean. *Molecular Ecology* **10**:371-356.

Crandall, K.A. (1996) Multiple interspecies transmissions of human and simian T-cell leukemia/lymphoma virus type I sequences. *Molecular Biology and Evolution* **13**(1):115-131.

Crandall, K.A., A.R. Templeton and C.F. Sing (1994) Intraspecific phylogenetics: problems and solutions. Models in Phylogeny Reconstruction.Systematics Association special volume no.52. R. W. Scotland, D. J. Siebert and D. M. Williams editors. Clarendon Press, Oxford: 273-297.

Creer, S., A. Malhotra, R.S. Thorpe and W.H. Chou (2001) Multiple causation of phylogeographical pattern as revealed by nested clade analysis of the bamboo viper (*Trimeresurus stjnegeri*) within Taiwan. *Molecular Ecology* **10**:1967-1981.

Crossman, E.J. and D.E. McAllister (1986) Zoogeography of freshwater fishes of the Hudson Bay drainage, Ungava Bay and the Arctic Archipelago. In C.H. Hocutt and E.O. Wiley (Eds.), The zoogeography of North American freshwater fishes, (p. 53-104). John Wiley and Sons, New York.

Demboski, J.R. and J. Sullivan (2003) Extensive mtDNA variation within the yellow-pine chipmunk, *Tamias amoenus* (Rodentia: Sciuridae), and phylogeographic inferences for northwest North America. *Molecular Phylogenetics and Evolution* **26**:389-408.

Donald, D.B. (1985) The wing length of *Sweltsa revelstoka* (Plecoptera:Chloroperlidae). *The Canadian Entomologist* **117**:233-239.

Dyke, A.S. and V.K. Prest (1987) Late Wisonsian and Holocene history of the Laurentide ice sheet. *Géographie Physique et Quaternaire* **41**:237-263.

Edmans, S. (2001) Phylogeography of the intertidal copepod *Tigriopus californicus* reveals substantially reduced population differentiation at northern latitudes. *Molecular Ecology* **10**:1743-1790.

Edmondson, W.T. and H.B. Ward (1959) Fresh-water biology. New York, Wiley. 1111pp.

Flint, R.F. (1971) Glacial and Quaternary geology.Wiley, New York.

Folmer, O., M. Black, W. Hoen, R. Lutz and R.Vrijenhoek (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metozoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**:294-299.

Guarnieri, M. (1996) A preliminary study of genetic variation within and among populations of *Diaptomus leptopus* (Copepoda:Calanoida): A DNA sequence analysis of the mitochondrial Cytochrome Oxidase I gene. MSc Thesis, Concordia University, Montreal 55pp.

Hairston Jr., N.G.B., R.A., Kearns, C.M., Engstrom, D.R. (1995) Age and survivorship of diapausing eggs in a sediment egg bank. *Ecology* **7**(6):1706-1711.

Hammer, M.F., T. Karafet and Rasanayagan (1998) Out of Africa and back again: nested cladistic analysis and human Y chromosome. *Molecular Biology and Evolution* **15**:427-441.

Harpending, H.C., M.A. Batzer, M. Gurven, L.B. Jorde, A.R. Rogers and S.T. Sherry (1998) Genetic traces of ancient demography. *Proceedings of the Natural Academy of Sciences USA* **95**:1961-1967.

Hellberg, M.E. (1994) Relationships between inferred levels of gene flow and geographic distance in a philopatric coral, *Balanophyllia elegans*. *Evolution* **48**:1829–1854.

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

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

Huelsenbeck, J.P. (1991). Tree-length distribution skewness: An indicator of phylogenetic information. *Systematic Zoology* **40**(3):257-270.

Huelsenbeck, J.P., J.P. Bollback and A.M. Levine (2002) Inferring the Root of a Phylogenetic Tree. *Systematic Biology* **51**(1):32-43.

Hughes, R.D. (1955) Geology of portions of Sunwapta and Southesk map areas, Jasper National Park, Alberta, Canada. *Alberta Society of Petroleum Geologists 5th Annual Conference Guide Book*: 69-116.

Hutchison, D.W. and A.R. Templeton (1999) Correlation of Pairwise Genetic and Geographic Distance Measures: Inferrring the Relative Influences of Gene Flow and Drift on the Distribution of Genetic Variability. *Evolution* **53**(6):1898-1914.

Kimura, M. (1981) Estimation of evolutionary distances between homologous nucleotide sequences. *Proceedings of the National Academy of Sciences of USA* **78**:454--458.

Knowles, L.L. and W.P. Maddison (2002) Statistical phylogeography. *Molecular Ecology* **11**:2623-2635.

Knowlton, N., L.A. Weight, L.A. Solorzano, D.K. Mills and E. Bermingham (1993) Divergence in proteins, mitochondrial DNA and reproductive compatibility across the isthmus of Panama. *Science* **260**:1629–1632.

Kuchta, S.R. and A.M. Tan (2005) Isolation by distance and post-glacial range expansion in the rough-skinned newt, *Taricha granulosa*. *Molecular Ecology* **14**:225-244.

Kumar, S. K. Tamura and M. Nei (2004) MEGA 2: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics* **5**:150-163.

Lee, C.E. and B. Frost (2002) Morphological stasis in the *Eurytemora affinis* species complex (Copepoda: Temoridae). *Hydrobiologia* **480**:111-128.

Legendre, P. and L. Legendre (1998) Numerical ecology. Elsevier Science BV, Amsterdam.

Machida, R.J., M.U. Miya, M. Nishida and S. Mishida (2002) Complete mitochrondrial DNA sequence of *Tigriopus japonicus* (Crustacea:Copepoda). *Marine Biotechnology* **4**:406-417.

Maddison, W.P. and D.R. Maddison (1992) MacClade version 3: Analysis of phylogeny and character evolution. Sinauer Associates, Sunderland, Massachusetts.

Maly, E. J., Maly, M.P. (1974) Dietary differences between two co-occuring calanoid copepod species. *Oecologia* **17**:325-333.

McLean, J.E., D.E. Hay and E.B. Taylor (1999) Marine population structure in an anadromous fish: Life-history influences patterns of mitochondrial DNA variation in the eulachon, *Thaleichthys pacificus*. *Molecular Ecology* **8**:143-158.

Mila, B., D.J. Girman, M. Kimura and T.B. Smith (2000) Genetic evidence for the effect of a postglacial population expansion on the phylogeography of a North American songbird. *Proceedings of the Royal Society of London B* **267**:1033-1040.

Nei, M. (1987) Molecular evolutionary genetics. Columbia, New York

Norris, R.M. and R.M. Webb (1990) Geology of California. New York.

Page, R.D.M. and E.C. Holmes (1998) Molecular evolution : a phylogenetic approach. Oxford ; Malden, MA, Blackwell Science.

Pielou, E.C. (1991) After the Ice Age : the return of life to glaciated North America. University of Chicago Press, Chicago.

Posada, D. and K.A.Crandall (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**:917-818.

Posada, D. and K.A. Crandall (2001) Intraspecific phylogenetics: Trees grafting into networks. *Trends in Ecology and Evolution* **16**(1):37-45.

Posada, D., K.A. Crandall and A.R. Templeton (2000) GeoDis: a program for the cladistic nested analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology* 9:487-488.

Rogers, A. R. (1995) Genetic evidence for a Pleistocene population explosion. *Evolution* **49**(4):608-615.

Rogers, A.R. and H. Harpending (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**(3):552-569.

Rozas, J., J.C. Sanchez-DelBarrio, X. Messeguer and R. Rozas (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**(18):2496-2497.

Rychlik, W., R.E. Rhoads (1989) A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. *Nucleic Acids Research* **17**(21):8543-8551.

Schultheis, A.S., L.A. Weight and A.C. Hendricks (2002) Gene flow, dispersal, and nested clade analysis among populations of the stonefly *Peltoperla tarteri* in the southern Appalachians. *Molecular Ecology* **11**:317-327.

Slatkin, M. (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47**(1):264-279.

Slatkin, M. and R.R. Hudson (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**:555-562.

SPSS Inc. (1999) SPSS Base 11.5 for Windows. SPSS Inc., Chicago IL

Staden, R. (1996) The Staden Sequence Analysis Package. *Molecular Biotechnology* **5**:233-241.

Swofford, D.L. (2001) PAUP* 4.0 - Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Assoc., Sunderland, MA.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **125**:585-595.

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

Templeton, A. R. (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology* **13**:789-809.

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., K.A. Crandall and C.F. Sing (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonucleases mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**:619-669.

Templeton, A.R., E. Routman and C. Philips (1995) Separating population structure from population history: a cladistic analysis the geographical distribution of mitochrondrial DNA haplotype in the Tiger salamander: *Ambystoma tigrinum*. *Genetics* **140**:767-782.

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* **132**:659-669.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin and D.G. Higgins (1997) The CLUSTAL_X windows interface; Flexible strategies for multiple-sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **27**:2682-2690.

Turgeon, J., and L. Bernatchez (2001) Clinal variation at microsatellite loci reveals historical secondary intergradation between glacial races of *Coregonus artedi* (Teleostei : Coregoninae). *Evolution* **11**:2274-2286.

Turgeon, J. and M.A. McPeek (2002) Phylogeographic analysis of a recent radiation of *Enallagma* damselflies (Odonata: Coenagrionidae). *Molecular Ecology* **11**(10):1989–2001.

Väinölä, R., J.K. Vainio, and J.U. Palo (2001) Phylogeography of "glacial relict" *Gammaracanthus* (Crustacea, Amphipoda) from boreal lakes and the Caspian and White seas." *Canadian Journal of Fisheries and Aquatic Science* **58**(11):2247-2257.

Ward, H.B. (1904) A biological reconnaissance of some elevated lakes in the Sierras and Rockies. *Transactions of the American Microscopical Society* **25**:127-152.

Watras, C.J. and J.F. Haney (1980) Oscillations in the reproductive condition of *Diaptomus leptopus* (Copepoda: Calanoida) and their relation to rates of egg-clutch production. *Oecologia* **45**:94-103.

Wilson, C.C. and P.D.N. Hebert (1998) Phylogeography and postglacial dispersal of lake trout (*Salvelinus namaycush*) in North America. *Canadian Journal of Fisheries and Aquatic Science* **55**:1010-1024.

Wilson, M.M. and H.C. Yeatman (1959) Free-Living Copepoda. In W.T. Edmonson and H.B. Ward (Eds.) Ward's and Whipple's Freshwater Biology (p.735-794). John Wiley and Sons, New York.

Wright, S. (1943) Isolation by distance. *Genetics* **28**:114–138.

Wright, S. (1946) Isolation by distance under diverse systems of mating. *Genetics* **31**:39-59.

Xia, X. and Z. Xie (2001) DAMBE: Data analysis in molecular biology and evolution. *Journal of Heredity* **92**:371-373.

Xia, X., Z. Xie, M. Salemi, L. Chen and Y. Wang (2003) An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution* **26**:1-7.

Chapter 3

Isolation and characterization of microsatellite markers for *Hesperodiaptomus shoshone.*

Introduction

Microsatellite markers are considered to be one of the most powerful tools for genetic studies of natural populations including the investigation of genetic structure and gene flow (Balloux and Lugon-Moulin, 2002). Microsatellite regions are prone to high mutation rates (Hancock, 1995) and therefore, show higher levels of variation between individuals within a species. Aside from their use in population and conservation biology, microsatellites have also proven useful in the study of genetic diseases (Ota *et. al,* 1999; Nsengimana *et al.,* 2005). However, the utility of microsatellites as genetic markers in many taxa is limited due to a paucity of genome sequence information needed to design oligonucletide primers for selective polymerase chain reaction (PCR) amplification of the microsatellite region. This is particularly true for many non-model organisms, and often requires screening of partial genomic DNA libraries to obtain nucleotide sequences for designing primers. To date, no microsatellites have been isolated and characterized for any of the North American diaptomid copepods. This chapter of the thesis focuses on the isolation and characterization of microsatellite markers for *Hesperodiaptomus shoshone* (S.A. Forbes, 1882).

Materials and Methods

Population and Species Description

Hesperodiaptomus shoshone is found in high altitude lentic environments of western North America, reproduces sexually and requires copulation before a clutch of eggs can be extruded (Watras and Haney, 1980). After emerging from the egg, the developing copepod passes through six naupliar stages (N1-N6) and five copepodite stages (C1-C5) before reaching the adult stage (Wilson and Yeatman, 1959). It is univoltine and produces only diapausing eggs (Maly and Maly, 1974).

Samples of *H. shoshone* for the genomic library construction were collected during the summer of 2002 from basin L8 (39° 01.72'N, 107° 03.83'W) of the Mexican Cut Nature Preserve located in the Elk Mountains of south-central Colorado. This watershed is owned by The Nature Conservancy and is managed by the Rocky Mountain Biological Laboratory. Four populations covering the broad natural distribution range of *H. shoshone*, namely Niwot Ridge (Mountain Research Station, University of Colorado in Boulder, Colorado), the Beartooth Mountains at the border of Wyoming/ Montana and the Sierra Nevada, California (Table 1) were also sampled and used for assessing the SSR polymorphisms. All samples were collected during the summers of 2002 and 2003 by the author except for the California population, which was collected by Andrew Kramer in 2003. Copepods were collected using a Wisconsin plankton tow net (120 μm) and immediately preserved in 95% ethanol following the procedure of Bucklin (2000).

Table 1
Coordinates and sample sizes of populations, of *Hesperodiaptomus shoshone* used for assessing the microsatellite polymorphism.

Population	Location	Lat/Long Coordinates	Sample Size
CO30	Niwot Ridge 5, Colorado	40° 01.71'N, 105° 33.45'W	10
CO14	Mexican Cut U2, Colorado	39° 01.55'N, 107° 03.84'W	9
WY1	Beartooth 8, Wyoming/Montana	44° 56.97'N, 109° 29.04'W	11
CA22	Lower Goethe Lake, California	37° 13.26'N, 118° 42.21'W	11

Molecular Methods

Isolation and characterisation of microsatellites were done by following a protocol modified from Dayanandan *et al.*, (1999).

DNA extraction

Sample of *H. shoshone* used for construction of the partial genomic library was extracted from a pool of adult copepods using the DNAeasy tissue kit (Qiagen, Mississauga, Ontario). Approximately 150 µl of crushed copepods were used in the extraction procedure. The DNA from individual copepods for assessing the microsatellite polymorphism was extracted following a protocol modified from Lee and Frost (2002) and Bucklin (2000). The copepods preserved in ethanol were soaked at 4°C overnight in 1.5 ml centrifuge tubes with 400 µl of TE (pH 8.5). On the following day, TE was removed and 100 µl of lysis buffer (10 mM Tris pH 8.3, 50 mM KCl, 0.5% Tween 20, and 200 µg/ml Proteinase K) was added before homogenizing the copepods with a pore sealed pipet tip. The homogenates were incubated at 65°C for 1 hour, and then 100 µl of Chloroform/Isoamyl alcohol (24:1) was added, gently inverted for 5 min and centrifuged at 7000 rpm for 5 min. The upper phase was transferred to a clean 1.5ml centrifuge tube containing 100µl of Isopropanol. The tubes were

gently inverted and the DNA was allowed to precipitate at room temperature for 45 min. The tubes were centrifuged at 14 000 rpm for 10 min and the supernatant was decanted. This was followed by 70% ethanol wash and the pellet was dried at room temperature overnight. The dried pellet was dissolved in 150µl of TE with RNAse (1 ug/ul) and incubated at 4°C for 4 hrs before using for PCR amplification.

Library construction

Extracted genomic DNA (50 µl) was digested with Sau3AI (Promega, Wisconsin) for 3 hrs at 37°C. The Sau3AI enzyme was then inactivated by heating the mixture for 10 minutes at 65°C. The digested DNA was concentrated by ethanol precipitation and dissolved in 30µl of TE. The M13 vector (MBI Fermentas) DNA was digested with the restriction enzyme *Bam*HI (MBI Fermentas) for 1 hour at 37°C and dephosporylated with 5U calf intestinal alkaline phosphatase (MBI Fermentas). The phosphatase was inactivated by adding 1.25 µl of 0.1 M EDTA and heating to 75 °C for 10 minutes. The digested DNA samples were electrophoresed in a 1% agarose gel with 1X TBE buffer. *H. shoshone* DNA fragments between 500 and 1000 bp and M13 DNA were excised from the gel and purified using the QIAquick gel extraction kit (Qiagen, Mississauga, Ontario).

The ligation of copepod and vector DNA was performed in a total volume of 21 µl with 4 µl of gel purified copepod DNA, 4µl of M13 DNA and 1µl of T4 DNA ligase at 16°C for 22 hours. Ligated products were transformed into XL1 BlueMRF competent bacterial cells (Strategene, California) following the manufacturer's recommendations. Transformed cells were plated on 137 mm culture plates with LB/tetracycline agar, Xgal, and IPTG and

incubated at 37°C overnight. A partial genomic library of approximately 2400 clones with a density of 300- 1200 plaques per plate was obtained.

Screening for microsatellite repeat sequences

Plates were blotted with nylon membranes (Hybond N+, Amersham) for 2 minutes. Membranes were baked at 80 °C for 2 hrs and washed sequentially in two aliquots of 2X SSC solution. Membranes were washed again with 5X SSC in a 50 °C shaking water bath for 1hr. The membranes were prehybridized overnight at 50 °C in a Tupperware™ container with 250 ml of hybridization medium (150 ml 20X SSC, 25 ml 100X Denhardts, 5 g BSA and 20 ml 10% SDS dissolved in a total volume of 500 ml dH_2O). Following prehybridization, the hybridization medium was replaced with 250 ml of fresh hybridization medium, and the labeled probe was added (prepared as described below) and incubated in a shaking water bath at 50°C for 3 hrs. Hybridized blots were washed twice with 100 ml of 0.1% SDS/6X SSC solution at room temperature, and soaked in 1000 ml of the same solution at 50 °C for 30 minutes. The membranes were rinsed with 6X SSC and exposed overnight to an autoradiographic film. The following oligonucleotides were used as probes: (AC)15/(TC)15 (AG)15/(TG)15 and a mixture of (ATC)10/(AAC)10/(AAT)10/(AAAT)7/(AT)15. The probes were removed from membranes by boiling in 0.1% SDS solution for 10 min before each successive hybridization. Each probe (20 pmol) was endlabeled with 1.11 x 10^6 Bq of [γ^{32}P] -ATP (Amerisham) using T4 polynucleotide kinase (Fermentas, Burlington, Ontario) at 37 °C for 1 hr in a total volume of 50 µl. The enzyme was inactivated by heating to 65 °C for 20 min. The labeled probe was purified using Qiagen Nucleotide Removal Kit (Qiagen, Mississauga, Ontario) and denatured at 90 °C before adding to the hybridization medium (see above).

Autoradiogaphic films were developed and positive clones were identified and picked using a sterile pipette tip. Clones were grown in 15 ml culture tubes containing 2.5 ml of LB/Tetracycline and 150 µl of overnight-grown XL1 BlueMRF bacterial cells in a shaking incubator at 37 °C for 6 hrs. A portion of each culture was transferred to 2 ml micro-centrifuge tubes and centrifuged at 14000 rpm for 5 minutes. The supernatant was decanted and the double stranded DNA was extracted from the pellet using the QIAprep Spin Miniprep kit (Qiagen, Mississauga, Ontario). The purified DNA was sequenced both directions with M13 forward and reverse primers using the ABI Prism Big DYE terminator sequencing kit (Applied Biosystems) on an ABI 310 automated DNA sequencer (Applied Biosystems).

Oligonucleotide primers complementary to regions flanking the repeat regions were designed and synthesized with Oligo 3.4 software (Rychlik and Rhoads, 1989). A DNA sequence (18 nucleotides) complementary to the M13 forward primer sequence was included at the 5' end of one of the primers of each pair. This allowed labeling of PCR products using fluorescently labelled universal M13 primers. The polymerase chain reaction (PCR) conditions were optimized for each primer pair to amplify microsatellites in different individuals. Amplification reactions were performed in a total volume of 25 µl with 250 µM of dNTP, 2.0- 2.5 mM $MgCl_2$, 2.5 ul 10X buffer (0.2 M Tris pH 9.5; 0.25 M KCl; 1mg/ml BSA, 5µl/ml Tween 20), and 0.4 pmol of each primer, 1.5 units of *Taq* DNA Polymerase, and 5 pmol (1ul) fluorescent dye-labelled M13 universal primer (either FAM, HEX, or TET). PCR conditions for primers were optimized as follows. After an initial denaturation for 1 minute at 94 °C, 35 cycles were done at 94 °C for 30 seconds (denaturation), annealing temperature of 40-60 °C for 15 seconds, and 72 °C for 30 seconds (extension). Optimal annealing temperatures for each primer are given in Table 2.

Genotyping

The PCR product (2-4 μl) was mixed with 15 μl of HiDi formamide and 0.1μl of GeneScan [TAMRA] size standard and electrophoresis was done on an ABI310 genetic analyzer. Allele sizes were determined using the ABI GENESCAN software.

Data Analysis

For each population and each polymorphic locus, the average number of alleles per locus (*A*), observed heterozygosity *(Ho),* and Nei's (1978) unbiased estimates of average heterozygosity *(H_E)* were calculated with Genetic Data Analysis (GDA) version 1.0 software (Lewis and Zaykin, 2001). The deviation from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium were examined for each population and all samples together using a Markov chain randomization test as implemented in Genepop 3.4 software (Raymond and Rousset, 1995). In order to detect the presence of non-amplifying ("null") alleles, the 1 -tailed probabilities of departure from Hardy Weinberg (HWE) equilibrium (heterozygote deficiency) were tested, using a Markov chain randomization test in Genepop 3.4 software (Raymond and Rousset, 1995).

Results

The partial genomic library of *H. shoshone* consisted of approximately 2400 clones with a density of 300-1200 colonies per plate. The first probe AC/TC, identified 9 positive clones, which upon sequencing revealed 7 microsatellites for which 4 primers were designed. Some positive clones with microsatellites were rejected due to the short microsatellite repeat numbers. The second probe, AG/TG, identified 13 positive clones and sequencing revealed 8 microsatellite sequences and primers were designed for 1. The probe mixture consisting of ATC/AAC/AAT/AAAT/AT identified 4 clones, which revealed 3 microsatellites, for which primers were designed for 1. Optimization of PCR conditions resulted in successful amplification with 6 of the 7 primer pairs (Table 2). The amplification with HSO4 was inconsistent, and could not be amplified reliably. The HSO6 primer pair produced multiple amplification products. Five loci produced amplification products consistent with a polymorphic single-locus pattern. The number of alleles per locus varied from 2 to 16 (Table 3).

Overall, the locus HSO1 showed the highest allelic diversity with an average of 5.75 alleles per population (Table 4). The lowest allelic diversity was found in the HSO7 locus with an average of 1.5 alleles. Observed heterozygosity *(Ho)* ranged from 0.1 to 0.89 and Nei's (1978) unbiased heterozygosity *(HE)* ranged from 0.1 to 0.86 (Table 4). The global test for Hardy Weinberg equilibrium revealed a significant departure from equilibrium at 1 of the 5 loci (locus HSO1). At the population level, a significant departure from Hardy Weinberg equilibrium was observed for 3 populations (locus HSO1: Mexican Cut U2 and Niwot Ridge 5, locus HSO5: Mexican Cut U2). Tests for heterozygote deficiency revealed significant heterozygote deficits in only 1 population (locus HSO1: Mexican Cut U2). No linkage

disequilibrium was observed, indicating that alleles of all five loci are segregating independently (Table 5).

Table 2

Primer sequences and optimal annealing temperatures (Ta) for seven microsatellite loci isolated from *Hesperodiaptomus shoshone.*

Locus	Primer sequences (5'-3')	T_a (°C)
HSO1	F: TCACTCCGTTGTTTTATGTTG R: CGTATTTGGCCCCTTTTACA	50-55
HSO2	F: GAATTCATTGAGCCGGGATAC R: GCCTAGGGGTACAATAGGGAAC	42
HSO3	F: CAAAAATAGGGGGTCACAAAG R: CAGACTCTTCTACATTTTCCCTC	45
HSO4	F: ATGTTAAGGAGGCCGCAC R: GAGACGCGCAGCCTGTAC	-
HSO5	F: GGCATGACCAGGGGATTG R: AGTTTTCATGTAAATCGCAACG	53
HSO6	F: TTCCCACGGTTTTGCGAG R: CTTAACCCCCTTTGGTACGGT	-
HSO7	F: GCTACTGGTACCCTTTGTGCT R: TCTTCACATGTGACAACGGAC	45

Table 3
Repeat pattern, cloned allele size (bp), and number of alleles detected (NA) for five polymorphic nuclear microsatellite loci isolated and optimized for *Hesperodiaptomus shoshone.*

Locus	Repeat Motif	Cloned allele size (bp)	NA
HSO1	(GT)17	350	16
HSO2	(AAATTT)6	171	2
HSO3	(TG)10	170	6
HSO5	$(TG)_7$	206	3
HSO7	(TAA)8	223	3

Table 4
Average number of alleles per locus, heterozygosity values and significance values for HWE for five polymorphic nuclear microsatellite loci in four populations of *Hesperodiaptomus shoshone*.. A: Number of alleles per locus, Ho: Observed heterozygosity, He: expected heterozygosity, HWE p: significance values from deviation from HWE, Hdef p: heterozygote deficiency significance levels. The following population abbreviations were used: CO14: Mexican Cut U2, Colorado; CO30: Niwot Ridge 5, Colorado; WY1: Beartooth 8, Wyoming; CA22: Lower Goethe Lake, California.

Locus	Population	A	Ho	He	HWE p	Hdef p
HSO1	CO14	4	0.2	0.43	**0.02**	**0.02**
	CO30	8	0.88	0.86	**0.03**	0.40
	WY1	6	0.45	0.47	0.31	0.58
	CA22	4	0.54	0.63	0.88	0.36
	Global	*5.75*	*0.52*	*0.61*	***0.03***	
HSO2	CO14	2	0.1	0.26	0.15	0.15
	CO30	2	0.33	0.38	1	1
	WY1	2	0.09	0.09	1	1
	CA22	2	0.18	0.39	0.28	0.27
	Global	*2*	*0.17*	*0.24*	*0.39*	
HSO3	CO14	4	0.6	0.7	0.11	1
	CO30	5	0.89	0.7	0.30	0.23
	WY1	4	0.66	0.71	0.22	-
	CA22	1	-	-	-	-
	Global	*3.5*	*0.50*	*0.51*	*0.13*	
HSO5	CO14	2	0.9	0.72	**0.04**	0.98
	CO30	3	0.88	0.58	0.07	-
	WY1	2	0.09	0.09	1	1
	CA22	3	0.18	0.17	1	1
	Global	*2.5*	*0.51*	*0.34*	*0.07*	
HSO7	CO14	1	-	-	-	-
	CO30	2	0.66	0.47	0.45	1
	WY1	2	0.27	0.45	0.43	0.43
	CA22	1	-	-	-	-
	Global	1.5	0.23	0.21	0.52	

Table 5
Linkage disequilibrium Chi^2 and P-values for each locus pair across all populations.

Locus pair	Chi^2	P-value
HSO1 & HSO2	9.516	0.301
HSO1 & HSO3	1.826	0.935
HSO2 & HSO3	7.268	0.297
HSO1 & HSO5	0.358	1.000
HSO2 & HSO5	4.326	0.827
HSO3 & HSO5	1.379	0.848
HSO1 & HSO7	2.013	0.733
HSO2 & HSO7	4.790	0.310
HSO3 & HSO7	6.046	0.196
HSO5 & HSO7	1.069	0.899

Discussion

Microsatellites are found in both prokaryotic (Field and Wills, 1996; Van Belkum *et al.,* 1998) and eukaryote genomes and are present in both coding (very rare) and non-coding regions. The distribution of microsatellites is not homogenous within a genome due to varying constrains imposed on coding vs. non-coding DNA (Toth *et al.,* 2000), historical processes (Wilder and Hollocher, 2001) and possible different functional roles of the repeats (Siwach *et al.,* 2006; Hammock and Young, 2004). The frequency of genomic microsatellites also varies per taxon, in terms of absolute numbers of microsatellite loci and preferential repeats (Hancock, 1999).

According to Zane *et al.,* (2002) arthropod DNA does not appear to be microsatellite rich and the general trend appears to be that the success rate for isolating positive repeats is approximately 2%. This is exceedingly low, indicating that the success rate achieved in this study was not atypical. The observed heterozygosity (Ho) values of 0.1 - 1 with a mean of 0.52 found in this study for *H. shoshone* are lower than those reported for other copepod species using microsatellite markers. For example, Ho values for the European freshwater diaptomid *Eudiaptomus gracilis* ranged between 0.53 and 0.87 with a mean of 0.72, and those for *Eudiaptomus graciliodes* ranged between 0.57 to 0.87 with a mean of 0.72 (Zeller, 2006). Genetic diversity of *H. shoshone* (mean H_E: 0.53) in the four populations examined was also lower than for *E. gracilis* and *E. graciliodes (*H_E: 0.85 and 0.84 respectively*)* (Zeller and Reusch, 2004). Low genetic diversity of *H. shoshone* might be due to the isolated temporary high altitude habitats occupied by this species.

Most of the populations and loci were in Hardy Weinberg equilibrium except for three populations and the HSO1 locus. Additionally, the Mexcican Cut U2 population in the HSO1

locus showed significant heterozygote deficiency. Several factors can lead to heterozygote deficiency, including null alleles, assortive mating, the Wahlund effect, selection against heterozygotes, inbreeding, or a combination of these. Null alleles are alleles that are not amplified (usually due to a mutation in one of the primer binding sites) and are commonly reported in microsatellite studies as being the source of heterozygosity deficit (Pemberton *et al.*, 1995).

Co-dominant inheritance cannot necessarily be assumed for microsatellite markers (Dobrowolski *et al.*, 2002). Controlled crosses addressing the inheritance of the isolated microsatellites should be performed for each marker but could not be addressed in this study for the following reason. After mating *H. shoshone* produces only resting eggs. Various extraction methods performed on the resting eggs failed to recover any DNA for subsequent PCR reactions. The conditions and cues for hatching of the resting eggs for *H. shoshone* are unknown and therefore it is impossible to culture progeny in the laboratory. Controlled crosses addressing the inheritance of a molecular marker have been performed previously for the intertidal copepod *Tigriopus californicus* (Harrison *et al.*, 2004) where Mendelian segregation was demonstrated for two inbred lines.

The five microsatellite loci isolated in this study may be useful for investigations of genetic variation within and among natural populations of *H. shoshone*.

References

Balloux, F. and N. Lugon-Moulin (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11**:155–165.

Bucklin, A. (2000) Methods for Population Genetic Analysis of Zooplankton. Chapter 11. In The Zooplankton Methodology Manual, International Council for the Exploration of the Sea. Academic Press, London p. 533 - 570.

Dayanandan, S., J. Dole, K. Bawa, and R. Kesseli (1999) Population structure delineated with microsatellite markers in fragmented populations of a tropical tree, *Carapa guianensis* (Meliaceae). *Molecular Ecology* **8**(10):1585-1592.

Dobrowolski, M.P., I.C. Tommerup, H.D. Blakeman and P.A. O'Brien (2002) Non-Mendelian inheritance revealed in a genetic analysis of sexual progeny of *Phytophthora cinnamomi* with microsatellite markers. *Fungal Genetics and Biology* **35**:197-212.

Field, D. and C. Wills (1996) Long, polymorphic microsatellites in simple organisms. *Proceedings: Biological Sciences* **263**(1367):209-215.

Hammock, E.A. and L.J. Young (2004) Functional microsatellite polymorphism associated with divergent social structure in vole species. *Molecular Biology and Evolution*: **21**(6):1057-1063

Hancock, J.M. (1995) The contribution of slippage-like processes to genome evolution. *Journal of Molecular Evolution* **41**:1038-1047.

Hancock, J.M. (1999) Microsatellites and other simple sequences: Genomic context and mutational mechanisms. In Microsatellites: Evolution and Applications. (D. B. Goldstein and C. Schlötterer, editors) Oxford University Press: 1-9.

Harrison, J.S., D.L. Peterson, J.R. Swain and S. Edmans (2004) Microsatellite DNA markers for the intertidal copepod *Tigriopus californicus. Molecular Ecology Notes* **4**:736-738.

Lee, C.E. and B. Frost (2002) Morphological stasis in the *Eurytemora affinis* species complex (Copepoda: Temoridae). *Hydrobiologia* **480**:111-128.

Lewis, P.O. and D. Zaykin (2001) GENETIC DATA ANALYSIS: computer program for the analysis of allelic data, version 1.0. Free program distributed by the authors from http://lewis.eeb.uconn.edu/lewishome/software.html.

Maly, E. J., Maly, M.P. (1974). Dietary differences between two co-occuring calanoid copepod species. *Oecologia* **17**:325-333.

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

Nsengimana, J., H. Renard and D. Goldgar (2005). Linkage analysis of complex diseases using microsatellites and single-nucleotide polymorphisms: application to alcoholism. *BMC Genetics* **6**(Suppl 1):10

Ota, M., N. Mizuki, Y. Katsuyama, G. Tamiya, T. Shiina, A. Oka, H. Ando, M. Kimura, K. Goto, S. Ohno and H. Inoko (1999). The Critical Region for Behçet Disease in the Human Major Histocompatibility Complex Is Reduced to a 46-kb Segment Centromeric of HLA-B, by Association Analysis Using Refined Microsatellite Mapping. *American Journal of Human Genetics* **64**:1406-1410.

Pemberton, J.M., J. Slate, D.R. Bancroft and J.A. Barrett (1995). Nonamplifying alleles at microsatellite loci a caution for parentage and population studies. *Molecular Ecology* **4**:349-352.

Raymond, M. and F. Rousset (1995) GENEPOP: Population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**: 248-249.

Rychlik, W. and R.E. Rhoads (1989) A computer program for choosing optimal oligonucleotides for filter hybridisation, sequencing and *in vitro* amplification of DNA. *Nucleic Acids Research* **17**(21):8543-8551.

Siwach, P., Pohaly, S.D., and S. Ganesh (2006) Genomic and evolutionary insights into genes encoding proteins with single amino acid repeats. *Molecular Biology and Evolution*: **23**(7):1357-1369.

Toth, G., Z. Gaspari and J. Jurka (2000) Microsatellites in Different Eukaryotic Genomes: Survey and Analysis. *Genome Research* **10**(7):967-981.

van Belkum, A., S. Scherer, L. van Alphen and H. Verbrugh (1998) Short-Sequence DNA Repeats in Prokaryotic Genomes. *Microbiology and Molecular Biology Reviews* **62**(2):275-293.

Watras, C.J. and J.F. Haney (1980) Oscillations in the reproductive condition of *Diaptomus leptopus* (Copepoda: Calanoida) and their relation to rates of egg-clutch production. *Oecologia* **45**:94-103.

Wilder, J. and H. Hollocher (2001). Mobile Elements and the Genesis of Microsatellites in Dipterans. *Molecular Biology and Evolution* **18**(3):384-392.

Wilson, M.M. and H.C. Yaetman (1959) Free-Living Copepoda. In Ward's and Whipple's Freshwater Biology. (W. T. Edmonson, editor) John Wiley and Sons, New York

Zane, L., L. Bargelloni and T. Patarnello (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology* **11**:1–16.

Zeller, M. (2006) Population genetic consequences of a diapause: A species comparison of two calanoid copepods (*Eudiaptomus* Spec.) PhD Thesis. Marx Plank Insitute.

Zeller, M. and T.B.H. Reusch (2004) Identification and characterization of 10 microsatellite primers for the calanoid freshwater copepods *Eudiaptomus gracilis* and *E. graciloides* using enriched genomic libraries. *Molecular Ecology Notes* **4**(3):355-357.

Chapter 4

Genetic structure of *Hesperodiaptomus shoshone* (Copepoda) in aestival ponds in the Rocky Mountains of Colorado based on nuclear microsatellite data.

Introduction

Aquatic ecosystems, one of the common features of alpine landscapes in Western North America, represent diverse habitats for many organisms and play a key role in the maintenance of biodiversity (Hinden *et al.,* 2005). These ecosystems are highly sensitivitive to environmental changes and, therefore, have recently become the focus of global climate change research (Fureder *et al.,* 2006). Due to extreme environmental conditions in alpine freshwater systems coupled with relatively simple food webs, these ecosystems react rapidly to environmental and anthropogenic changes (Psenner, 2002). Even slight impacts can significantly affect the physical and chemical properties and induce shifts in species composition and abundance (Hofer *et al.,* 2001). To date, most of the research has focused on alpine lakes, and smaller water bodies such as ponds and their benthic and planktonic communities have received relatively less attention.

Mountain valleys on the wet western slopes of the Rocky Mountains of Colorado are dotted with numerous wetlands, ponds and lakes, which occur in a diversity of geomorphologic and hydrologic settings (Wissinger *et al.,* 1999). Although multiple water-bodies occur in close proximity in the same valley, they may vary in size, nutrient content, hydroperiod and species composition. For example, the Mexican Cut Nature Preserve in the Elk Mountains of Colorado is a wetland complex composed of over 50 water basins. Species diversity and community composition is dependent on hydroperiod (vernal to permanent),

basin size and the presence of the vertebrate predator, the tiger salamander, *Ambystoma tigrinum nebulosum* (Green, 1825) (Wissinger *et al.,* 1999). Thus, water basins separated by a few meters might have distinct species assemblages. Similarly, individuals of a single species residing in neighbouring basins may be subjected to widely differing environments and potentially different selection regimes. Thus, local adaptation, population bottlenecks and genetic drift may have important implications in origin and maintenance of biodiversity in this area. These freshwater systems provide a unique opportunity to study the evolution of biological diversity and the partitioning of genetic diversity. The overall objective of the study presented in this chapter is to examine the genetic diversity at a local scale of a high altitude freshwater diaptomid copepod *Hesperodiaptomus shoshone* (S.A. Forbes, 1882) within and among neighbouring watersheds of the Elk Mountains of Colorado.

Hesperodiaptomus shoshone is a large diaptomid copepod restricted to high altitude regions of North America including the Rocky Mountains and the Sierra Nevada (Wilson and Yeatman, 1959). Although this species occurs in lakes and ponds throughout its range, in Colorado it is mainly found in temporary bodies of water that are subject to greater environmental fluctuations. Diaptomid copepods are an important component of alpine freshwater communities, and most often the large raptorial *Hesperodiaptomus* species are of keystone importance in these environments (Anderson, 1980; Stoddard, 1987; Paul *et al.,* 1995). Moreover, copepods provide an ideal system for biological diversity studies including the analysis of partitioning of genetic diversity. They occur in large numbers, are easy to sample for genetic analysis and occupy discrete habitats, making the task of defining the populations simple.

Copepods and zooplankton in general, are considered to disperse frequently over large distances (see review by Bohonak and Jenkins, 2003). However, genetic studies often reveal high genetic differentiation among nearby populations, indicating restricted gene flow (Boileau *et al.,* 1992, De Meester, 1996, Staton *et al.,* 2003). At present there is no general consensus to explain the low gene flow / high dispersal paradox for zooplankton (Bohonak and Jenkins, 2003). Boileau *et. al.* (1992) first suggested that the observed high genetic variability is a result of "persistent founder effects", where colonisation events by a small number of individuals result in large stable populations resistant to subsequent invasions. The monopolization hypothesis proposed by De Meester *et. al.* (2002) expanded upon this idea and suggested that the presence of diapausing stages and rapid local adaptation should increase the persistent founder effects. Although dispersal is frequent, the propagules do not thrive or establish themselves readily in the new populations. Both hypotheses assume that dispersal in zooplankton is a common occurrence. In contrast, Bohonak and Jenkins (2003) suggested that passive dispersal is infrequent in freshwater invertebrates, which exacerbates the persistent founder effects and they expanded the monopolization hypothesis to include aquatic invertebrates in general. Copepod populations restricted to small unpredictable alpine environments should show less resistance to invasions compared to large populations found in lakes. Thus, gene flow within watersheds should be more common than among watersheds in montane habitats.

Reliable estimates of genetic differentiation are crucial for understanding the connectivity among populations and important for developing conservation strategies. Genetic structuring reflects the number of individuals exchanged between populations. Low gene flow in populations will influence the evolutionary process by either permitting local

adaptation and ultimately speciation or reducing the fitness by fixation of deleterious mutations, which could lead to extinction (Higgins and Lynch, 2001). On the other hand, high gene flow can prevent local adaptation or fixation of alleles and will slow down the process of speciation (Barton and Hewitt, 1985).

Recent studies have shown that populations of the large hesperodiaptomids such as *H. shoshone* do not always recover from anthropogenic disturbances such as fish stocking (Parker *et al.*, 1996; McNaught, 1999). Reintroduction efforts of *H. shoshone* in the Sierra Nevada have also yielded mixed results (Andrew Kramer, personal communication). Understanding the genetic structure of *H. shoshone* is important for gaining insights into how they recover from natural and anthropogenic disturbances in these environments. This chapter of the thesis will address the following questions:

1) What is the local scale genetic structure of *H. shoshone* populations in alpine and sub alpine ponds?

2) Does the genetic diversity of *H. shoshone* show patterns associated with isolation by distance?

Methods

Species chosen for the study

The diaptomid copepod, *Hesperodiaptomus shoshone* reproduces sexually, and females must copulate before a clutch of eggs is extruded (Watras and Haney, 1980). The life cycle includes six naupliar stages and five copepodid stages before the final moult into adulthood, after which growth stops (Edmonson and Ward, 1959). *Hesperodiaptomus shoshone* is found in the Rocky mountains from British Columbia to Colorado, and in the Sierra Nevada. This species is univoltine and matures early in the summer in Colorado (Maly and Maly, 1974).

Sampling locations

Sampling was conducted in the Elk Mountains of central Colorado. Ponds (<1000 m^2) were sampled using a Wisconsin plankton tow net (120 μm) during the summers of 2003-2004. Eight representative watersheds and 12 ponds were used for the analysis (Table 1, Figure 1). The location of sampling sites within the Mexican Cut watershed is given in Figure 2.

Mexican Cut Nature Preserve

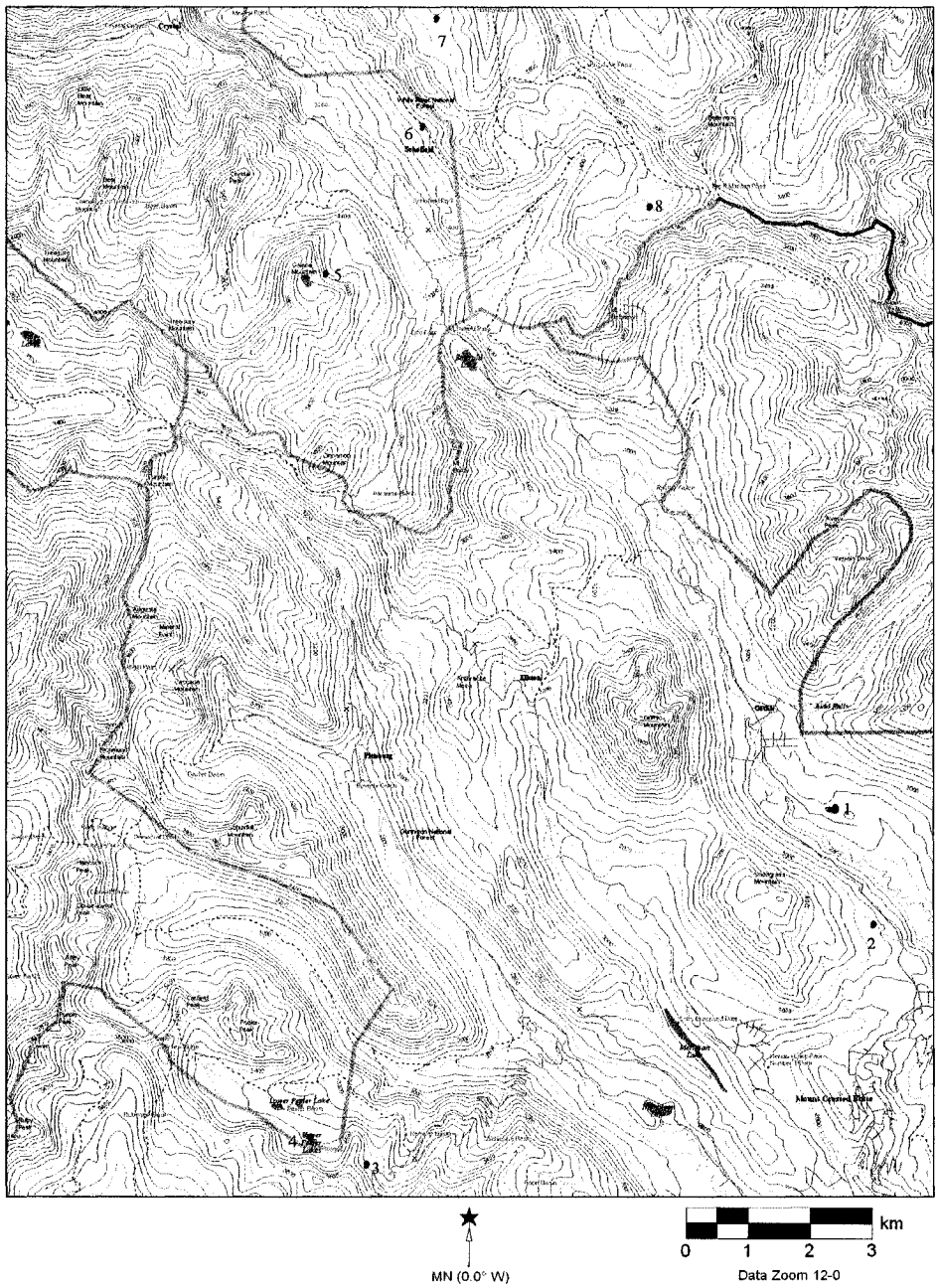
The Mexican Cut is the first natural area designated by the Nature Conservancy in Colorado. Alternating layers of limestone and quartzite eroded by glacial action form a series of shelves that hold over 50 water basins. The numerous interconnected glacial basins are fed by snowmelt and rainfall. Although many of the basins are only a few meters apart, they often contain quite distinct assemblages of invertebrates (Wissinger *et al.,* 1999). The ponds

included in this study however have similar invertebrate communities (see Wissinger *et al.*, 1999). These semi-permanent ponds are often dry for a short period of time in late summer then refill in autumn before the ice cover forms. All basins occur at about the same elevation, have similar microclimatic conditions and share the same pool of potential colonists. Ponds that contain *H. shoshone* are identified with an x in Figure 2. *Hesperodiaptomus shoshone* does not occur in permanent ponds with adults of the salamander, *Ambystoma tigrinum*. Although the salamanders are mostly confined to permanent basins, eggs and larvae do occur some years in the temporary ponds. The characteristics of ponds located at the Mexican Cut including densities of the salamander larvae that actively feed on *H. shoshone* are given in Table 2.

Table 1
The basins and ponds sampled in the Elk Mountains of central Colorado used in the genetic structure analysis of *Hesperodiaptomus shoshone*. N = sample size used in the study.

Basin Name	Pond	Altitude	Longitude	Latitude	N
Mexican Cut	U1	3460m	39° 01.73'N	107° 04.06'W	26
	U2	3460m	39° 01.68'N	107° 04.04'W	32
	L11	3400m	39° 01.794'N	107° 03.837'W	30
	L8	3400m	39° 01.731'N	107° 03.844'W	29
	L15	3395m	39° 01.64'N	107° 03.71'W	30
Hasley basin	Pond 8	3541m	39° 03.955'N	107° 02.867'W	30
Redwell Basin	Pond 2	3410m	38° 53.587'N	107° 03.619'W	30
Peeler Basin	Pond 2	3357m	38° 53.753'N	107° 04.206'W	30
Schofield Pass	Pond 1	3713m	39° 02.75'N	107° 02.75'W	30
Kettle ponds	K4	2860m	38° 56.70'N	106° 58.60'W	30
Snodgrass	Pond 2	2902m	38° 55.62'N	106° 57.96'W	28
East Maroon Pass	-	3606m	39° 0.69'N	106° 56.48'W	30

Figure 1
Sampling locations of *Hesperodiaptomus shoshone* populations in Elk Mountains of central Colorado. Refer to figure 2 for more details of the Mexican Cut Nature Preserve.



1. Kettle pond (K4)	5. Mexican Cut
2. Snodgrass	6. Schofield Pass
3. Redwell Basin	7. Hasley Basin (#8)
4. Peeler Basin (#2)	8. East Maroon Pass

Figure 2
The distribution of basins in the Mexican Cut Nature preserve (Sampling location 5 in Figure 1) with populations of *Hesperodiaptomus shoshone* sampled in 2002/2003. The dotted lines represent the two elevations at which the ponds occur at the Mexican Cut (3460m and 3400m). The x's represent ponds that contain *H. shoshone*. The ponds used in the analyses are: U1, U2, L11, L8 and L15. The arrows indicate the direction of the water flow between the basins. Modified from Wissinger *et al.* (1999).

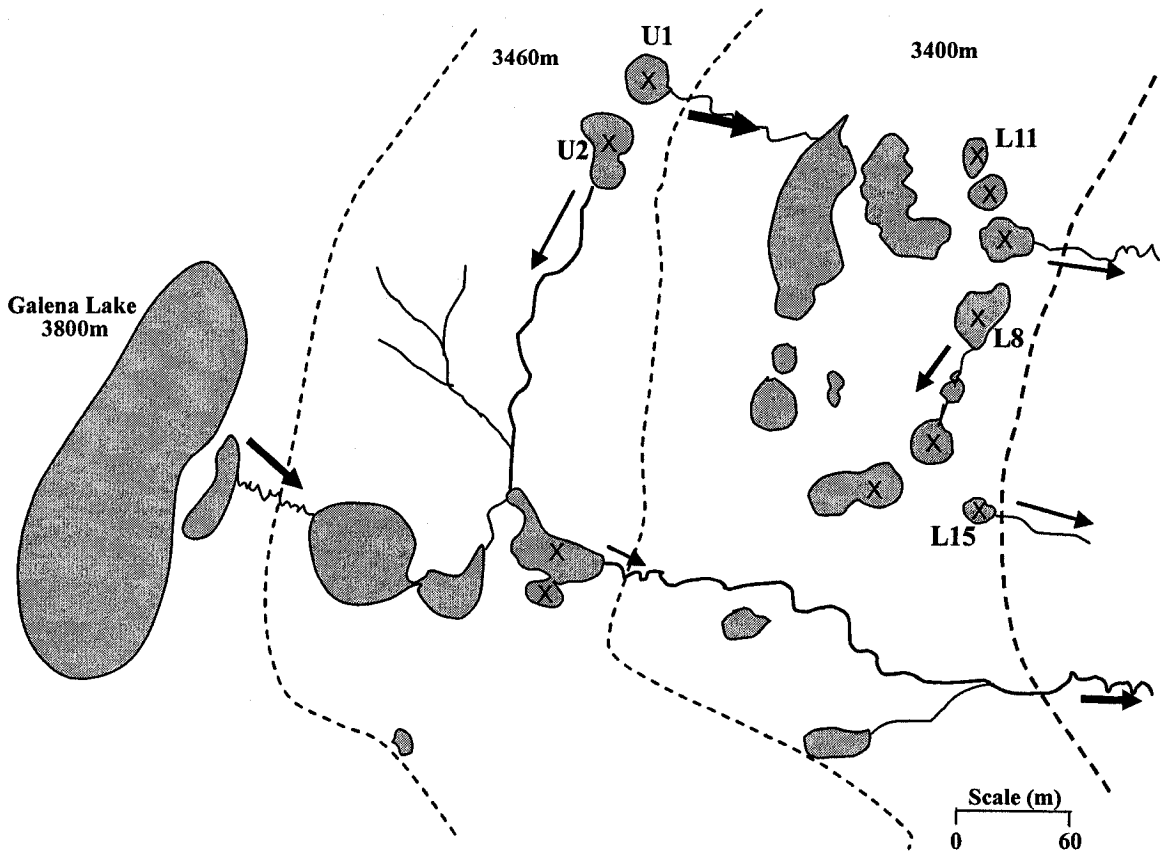


Table 2
Basin characteristics (area, depth, hydroperiod) and distribution and abundance of *Ambystoma tigrinum* hatchling larvae based upon multiple mark recapture censuses from 1990 for ponds at the Mexican Cut Nature Preserve (Wissinger *et al.*, 1999). Area is the maximum coverage (m^2) and depth is maximum depth (m) when the basins are completely filled. Wet represents the average number of ice-free days with water above the substrate.

	L8	**L11**	**L15**	**U1**	**U2**
Area (m^2)	928	213	64	600	750
Depth (m)	0.7	1	0.6	0.7	0.7
Wet	112	115	74	100	102
A. tigrinum hatchling larvae numbers	450	40	0	0	0

Molecular analysis

DNA from 26-30 individuals from each pond (Table 1) was extracted using a modified method from Bucklin (2000) and Lee and Frost (2002) described in detail in Chapter 1. Although individuals were screened for the five loci isolated in Chapter 3 only three were retained for analyses in this chapter. Individuals were genotyped using three polymorphic loci (HSO1, HSO2, and HSO7). Oligonucleotide primer sequences, PCR and genotyping conditions are given in Chapter 3 of this thesis.

Data analysis

Descriptive Statistics and Hardy Weinberg equilibrium

For each population and at each polymorphic locus, the average number of alleles (*A*), frequency of private alleles, observed heterozygosity *(Ho)* and Nei's (1978) unbiased estimates of average heterozygosity *(H_E)* were calculated using Genetic Data Analysis (GDA) version 1.0 software (Lewis and Zaykin, 2001). Linkage disequilibrium between loci and deviations from Hardy-Weinberg genotype frequency equilibrium (HWE) were tested using GENEPOP software (Raymond and Rousset, 1995). The significance of departure from HWE and linkage disequilibrium was estimated by Fisher exact tests where p-values were estimated by applying a Markov chain method with the following parameters (10 000 dememorizations; 200 batches; 5 000 interactions per batch) as described in Raymond and Rousset (1995). Correlations between altitude and the two measure of genetic diversity A and H_E were performed in SPSS 11.5 (1999).

Population genetic structure

Two mutational models are commonly used for modeling microsatellite evolution which form the basis for different statistics used for the analysis of genetic subdivision of populations.

1) The infinite alleles model (IAM; Kimura and Crow, 1964)

This model assumes that each mutation creates a unique allele (there are an infinite number of states that an allele can mutate to). In this model no two alleles are likely to share a sequential history or new mutations cannot be predicted on the basis of earlier ones.

2) The stepwise mutation model (SMM; Kimura and Otha, 1978)

Microsatellites mutate largely by slippage, increasing or decreasing allele size by one or more repeat units. In this mutational model alleles have a potentially recoverable history based on allele size and it assumes that new mutations can be predicted on the basis of earlier ones.

Much attention has been given to the performance of different statistical methods for assessing population structure using microsatellite markers, with differences between the statistics that assume the SMM and IAM model. To date a consensus has not been reached as to which performs best (Neigel, 2002) therefore both classes of statistics were employed.

Most methods for estimating population differentiation are related to F-statistics, originally developed by Wright (1951) using the infinite alleles mutation model (IAM). F_{st} can be defined as

$$F_{st} = \frac{(H_t - H_s)}{H_t}$$

Where H_t is the heterozygosity of the total population and H_S is the average heterozygosity over all subpopulations. F_{ST} can theoretically range from 0 (no genetic

differentiation) to 1(complete fixation). Many different methods have been developed to estimate unbiased parameters analogous to F_{ST}.

Slatkin's (1995) R_{st}, an F_{ST} analogue, is an alternative method for estimating subdivision and gene flow from microsatellite data based on the stepwise mutational model (SMM). R_{ST} differs from F_{ST} that it considers allele sizes, which according to the stepwise mutation model contain information about the relationships among alleles (Slatkin, 1995). Whether the R_{ST} model is a better estimate of gene flow than the F_{ST} depends on whether microsatellite loci more closely follow the stepwise mutation model or the infinite mutation model.

R_{ST} (Slatkin 1995) can be defined as:

$$R_{ST} = \frac{(S - S_w)}{S}$$

Where S is equal to the average sum of squares of the differences in allele size between populations, and Sw is the average sum of squares of the differences in allele size within populations. R_{ST} is the fraction of the total variance in allele size that is between populations. Most microsatellites are thought to evolve according to a model that is intermediate to the SMM and IAM, although microsatellite mutations are considered to follow the SMM more closely than the IAM (Fu and Chakraborty, 1998).

Population genetic structure was estimated with both Θ, an estimator of F_{ST} (Weir and Cockerham, 1984) and Rho, an unbiased estimator of Slatkin's R_{ST} (1995) using the programs FSTAT (Goudet, 1995) and RST-CALC (Goodman, 1997) respectively. R_{ST} assumes that populations are of equal size and that all loci have equal variances, assumptions not met in most populations. On the other hand, Rho estimates are weighted so that each population, regardless of size, contributes equally to the estimate of the statistic. For pairwise Rho and

F_{ST} estimates, the level of statistical significance was assessed through a permutation procedure (1 000 permutations) and adjusted for multiple simultaneous tests using the Bonferroni correction (Holm, 1979: Rice, 1989).

Although several genetic distance measures have been developed based upon IAM and SMM, only a few remain in common use (reviewed in Nei, 1987; Kalinowski, 2002), Nei's (1978) unbiased standard genetic distance or Ds continues to be one of the most commonly used genetic distances assuming the IAM, while delta μ^2 distance, developed by Goldstein *et al.* (1995) for use with microsatellite data, represents a commonly used genetic distance measure based upon SMM .

Nei's (1978) unbiased standard genetic distance Ds was calculated, using Popgene version 1.31 software (Yeh *et al.,* 1999) between each population pair. Calculation of delta μ^2 (Goldstein *et al.,* 1995) was conducted using RST-CALC (Goodman, 1997). Nei's genetic distance and delta μ^2 were used to construct UPGMA (unweighted pair group method with arithmetic average) dendrograms using PAUP* 4.0 (Swofford, 2001).

Hierarchal F statistics

The *F*-statistic is a widely used measure for elucidating the pattern and extent of genetic variation within and among populations of animal and plant species.

The *F* statistics (Wright, 1951; 1965) for the hierarchy of populations within species (F_{ST}), for individuals within the species (F_{IT}), and for populations (F_{IS}) were calculated by the method of Weir and Cockerham (1984), as implemented in the program F_{STAT} (Goudet, 1995). For subdivided populations and over multiple loci the following estimates can also be made to estimate population structure:

Measure of genetic diversity within populations: $F_{IS} = \frac{(H_S - H_O)}{H_S}$

Measure of genetic diversity in the total population: $F_{IT} = \frac{(H_T - H_O)}{H_T}$

Where

H_O: Average observed heterozygosity within a subpopulation over all loci

H_S: Average expected heterozygosity within subpopulations over all loci

H_T: The average of the expected heterozygosity

Isolation by distance

Pairwise estimates of genetic distance and geographic distances were used to infer patterns of isolation by distance (Slatkin and Maddison, 1990; Slatkin, 1993). In this study, pairwise Rousset's distances (1997) (Fst/(1-Fst) were plotted against geographic distances in kilometers, and the correlation coefficient was estimated. Pairwise geographic distances between populations were calculated while accounting for the curvature of the earth geographic distances in the R package 3.0 (Casgrain and Legendre, 2001). Because sample points are not independent, the significance of the correlation coefficient was estimated with Mantel tests (Mantel, 1967), 30 000 replicates, as performed in IBD (Isolation by Distance) software (Bohonak, 2002).

Results

Five microsatellite loci developed for *Hesperodiaptomus shoshone* (Chapter 3) were initially surveyed for this study. Two of the 5 loci were not included in the analyses either due to lack of polymorphism (HSO5), or because the occurrence of multiple fragments caused difficulty in inferring the homology of amplified DNA fragments and assigning to corresponding alleles to the locus (HSO3) in individuals of some populations examined.

The levels of genetic diversity and polymorphism among loci used were variable (Table 3). Locus HSO1 showed high levels of polymorphism with 2 to 8 alleles per population and expected heterozygosities ranged from 0.16 to 0.77. Loci HSO2 and HSO7 showed low levels of polymorphism with 1 to 3 alleles per population and expected heterozygosities ranged from 0 to 0.45. The H_E across all loci was highest in the Snodgrass population with a value of 0.49. The low H_E values were found in Lower pond 15 (Mexican Cut), Hasley Basin (#8) and Lower pond 11 (Mexican Cut) with values of 0.20, 0.26 and 0.26 respectively. Observed heterozygosity ranged from 0.10 in pond L8 (Mexican Cut) to 0.56 in the Snodgrass population. Mean Ho and H_E across all loci were 0.26 and 0.32 respectively (Table 3).

Of the twelve populations studied, those from Snodgrass and Redwell Basin demonstrated the highest allelic diversity with an average of 4 alleles per locus. The lowest allelic diversity was found in lower pond 8, Mexican Cut with an average of 1.67 alleles per locus. Two populations were fixed for alleles, the Mexican cut (Lower 11) population was fixed for both HSO2 and HSO7 loci and the Mexican Cut (L8) was fixed for the HSO7 locus (Figure 3). Private alleles were identified in four of the twelve populations (Table 4). One private allele was found in each of the East Maroon Pass, Hasley Basin (pond #8), and Kettle

pond (#K4) populations, while four were found in Redwell Basin (#2). The locus HSO1 private alleles account for more than 15% of the allelic frequency in two populations; 1) allele 356, Kettle pond K4, with an allele frequency of 35%, 2) allele 382, Redwell Basin (#2), with an allele frequency 20%. Figure 3 shows the allele frequencies for each locus and each population sampled. No significant correlations were observed between altitude and H_E (R^2= 0.022, p= 0.642) or altitude and allelic diversity (R^2= 0.006, p= 0.797).

Table 3
Summary of genetic variation at three microsatellite loci scored from *Hesperodiaptomus shoshone* populations in Gunnison, Colorado: sample size (N), observed and expected heterozygosity (H_O and H_E); number of alleles per locus (A). He in bold indicate significant deviation from HWE (at 0.05 significance level). He followed by * indicate significant heterozygote deficit.

Pop.	**HSO1**			**HSO2**			**HSO7**			**Avg. over all loci**			
	Ho	He	A	Ho	He	A	Ho	He	A	N	Ho	He	A
Schofield	0.26	**0.67***	7	0.30	0.44	2	0.13	0.18	2	30	0.23	**0.43**	3.66
Hasley B.	0.40	**0.59***	6	0.03	0.03	2	0.16	0.16	2	30	0.20	**0.26**	3.33
Peeler B.	0.37	**0.44***	5	0.30	0.30	2	0.06	0.07	2	30	0.24	0.27	3.00
Redwell B.	0.67	0.63	8	0.23	0.35	2	0.10	0.09	2	30	0.33	0.36	4.00
MC (#U2)	0.28	**0.40***	5	0.38	0.41	2	0.06	0.06	2	32	0.24	0.29	3.00
Kettle pond	0.57	0.61	5	0.1	**0.30***	2	0.03	0.03	2	30	0.23	**0.31**	3.00
MC (#L8)	0.10	0.16	2	0.20	0.29	2	0	0	1	29	0.10	0.15	1.67
MC (#L11)	0.69	**0.77***	5	0	0	1	0	0	1	30	0.22	0.26	2.33
MC (#L15)	0.12	**0.24***	2	0.4	0.33	2	0.03	0.03	2	30	0.18	0.20	2.00
Snodgrass	0.56	0.58	7	0.6	0.46	2	0.5	0.45	3	28	0.56	0.49	4.00
MC(#U1)	0.58	**0.73***	7	0.12	**0.24***	2	0.23	0.31	2	26	0.31	**0.42**	3.67
E. Maroon	0.50	**0.71***	5	0.23	0.26	2	0.1	0.09	2	30	0.28	**0.35**	3.00
all pops	0.42	**0.57**	5.33	0.24	**0.28**	1.9	0.12	0.12	1.91	355	0.26	0.32	3.05

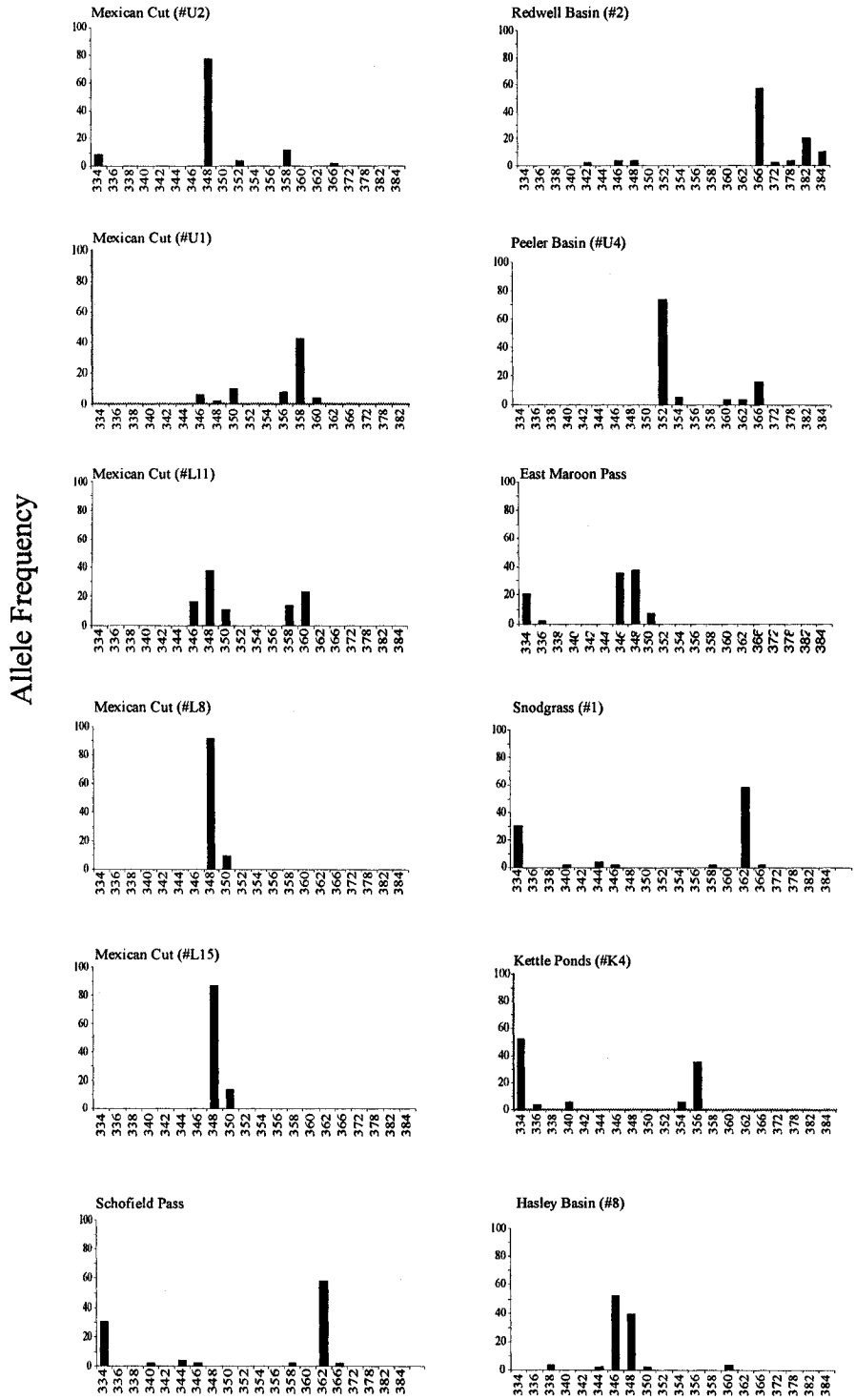
Table 4
Private alleles and their frequency of occurrence in *Hesperodiaptomus shoshone* populations sampled in 12 populations of the Gunnison, Colorado

.

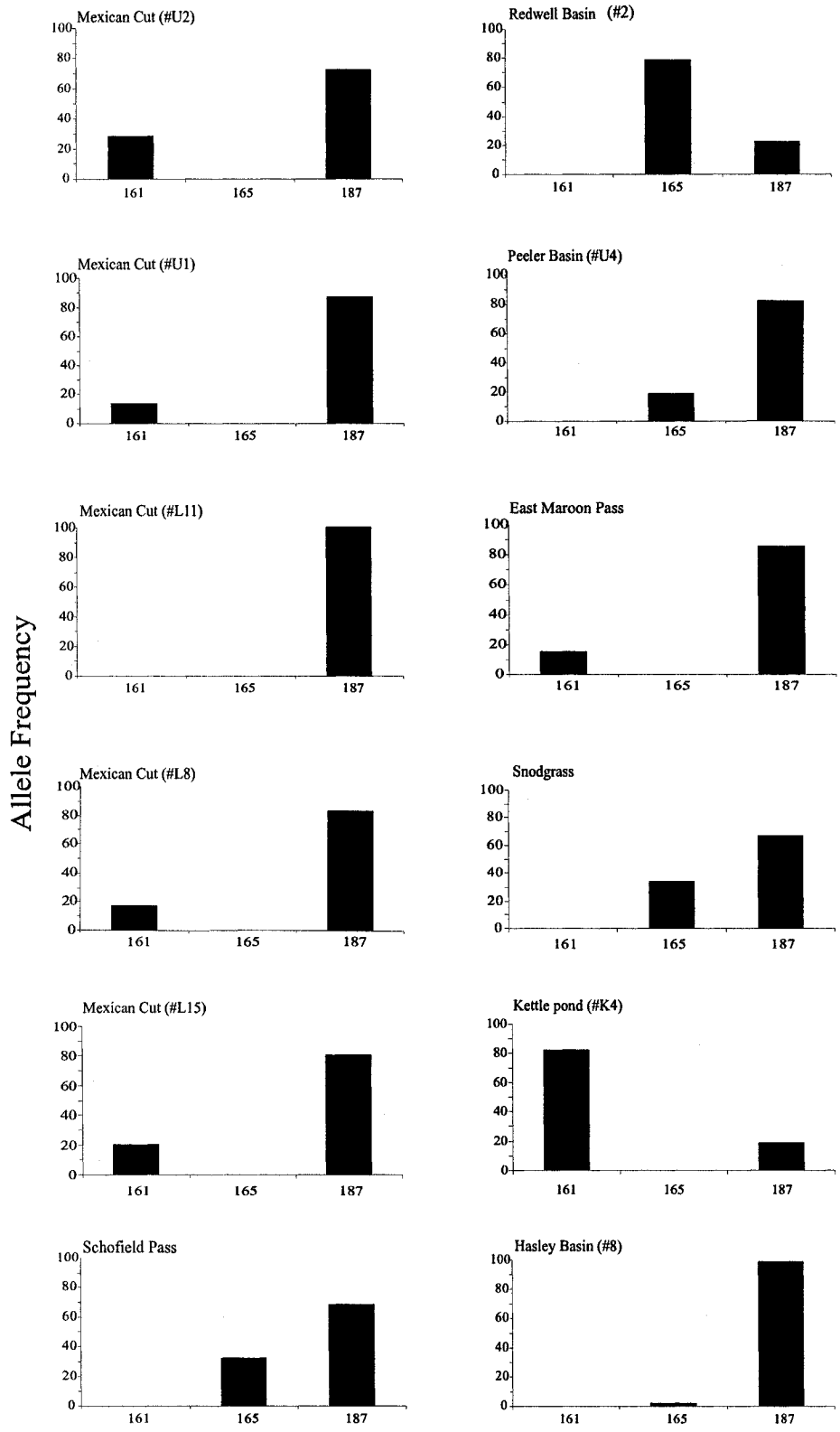
Locus	Population	Allele	Frequency
HSO1	Kettle pond (K4)	356	0.35
	Redwell Basin (#2)	342	0.02
	Redwell Basin (#2)	384	0.10
	Redwell Basin (#2)	372	0.02
	Redwell Basin (#2)	382	0.20
	Hasley Basin (#8)	338	0.03
HSO7	East Maroon Pass	234	0.05

Figure 3 Microsatellite allele size (bp) frequencies across loci and populations of *Hesperodiaptomus shoshone.* a) HSO1 locus, b)HSO2 locus and c) HSO7 locus

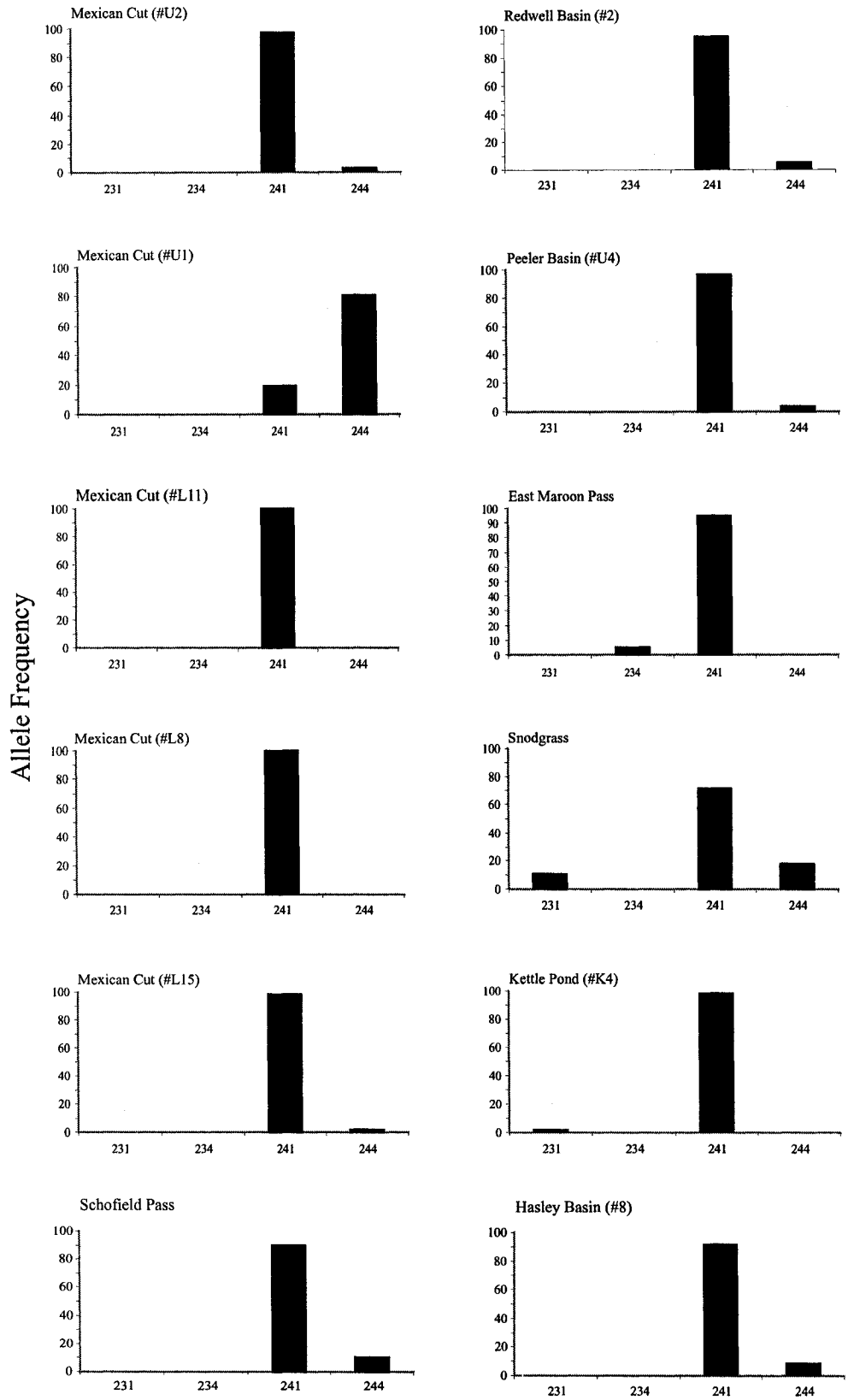
a) HSO1 Locus



b) HSO2 Locus


Mexican Cut (#U2)
Redwell Basin (#2)
Mexican Cut (#U1)
Peeler Basin (#U4)
Mexican Cut (#L11)
East Maroon Pass
Mexican Cut (#L8)
Snodgrass
Mexican Cut (#L15)
Kettle pond (#K4)
Schofield Pass
Hasley Basin (#8)
100
80
60
40
20
0
161
165
187
Allele Frequency
Allele size

c) HSO7 Locus


Mexican Cut (#U2)
Redwell Basin (#2)
Mexican Cut (#U1)
Peeler Basin (#U4)
Mexican Cut (#L11)
East Maroon Pass
Mexican Cut (#L8)
Snodgrass
Mexican Cut (#L15)
Kettle Pond (#K4)
Schofield Pass
Hasley Basin (#8)
Allele Frequency
231
234
241
244
Allele Size

Most loci and populations showed no significant departure from Hardy Weinberg equilibrium expectations (Table 3). A significant departure from the Hardy Weinberg equilibrium was observed only for the HSO1 locus in most populations and tests for heterozygote deficiency revealed significant heterozygote deficits in eight of the twelve populations for this locus. Linkage disequilibrium was observed for four populations, Peeler, Redwell, Snodgrass and East Maroon pass between the HSO1 and HSO2 loci and for Peeler between the HSO1 and HSO7 locus (Table 5). No linkage disequilibrium was observed for other populations, indicating that all three loci are most likely segregating independently of each other.

The R_{ST} estimate averaged over all three loci was 0.332, indicating substantial genetic structure (Wright, 1978; Hartl and Clark, 1997), whereas R_{ST} between populations ranged from -0.0006 to 0.751 (Table 5). Overall F_{ST} estimates were 0.379 and comparable to R_{ST} estimates, but showed less variability and ranged from -0.0144 to 0.583 between population pairs (Table 6).

Table 5
Linkage analysis results by population and locus for 12 populations of *Hesperodiaptomus shoshone* based upon 3 microsatellite loci. Statistically significant values ($p<0.05$) are indicated in bold.

Population	Locus pair		P value
Schofield Pass	HSO1	HSO2	0.149
	HSO1	HSO7	0.591
	HSO2	HSO7	0.148
Hasley Basin	HSO1	HSO2	0.175
	HSO1	HSO7	0.087
	HSO2	HSO7	1
Peeler Basin	HSO1	HSO2	**<0.001**
	HSO1	HSO7	**0.047**
	HSO2	HSO7	1
Redwell Basin	HSO1	HSO2	**0.005**
	HSO1	HSO7	0.332
	HSO2	HSO7	1
Mexican Cut U2	HSO1	HSO2	0.615
	HSO1	HSO7	1
	HSO2	HSO7	1
Kettle pond K4	HSO1	HSO2	0.213
	HSO1	HSO7	0.323
	HSO2	HSO7	0.231
Mexican Cut L8	HSO1	HSO2	0.492
	HSO1	HSO7	Not possible
	HSO2	HSO7	Not possible
Mexican Cut L11	HSO1	HSO2	Not possible
	HSO1	HSO7	Not possible
	HSO2	HSO7	Not possible
Mexican Cut L15	HSO1	HSO2	0.478
	HSO1	HSO7	1
	HSO2	HSO7	0.402
Snodgrass	HSO1	HSO2	**0.025**
	HSO1	HSO7	0.357
	HSO2	HSO7	0.644
Mexican Cut U1	HSO1	HSO2	0.599
	HSO1	HSO7	0.077
	HSO2	HSO7	0.706
East Maroon Pass	HSO1	HSO2	**<0.001**
	HSO1	HSO7	0.485
	HSO2	HSO7	1

Table 6

Pairwise Θ estimates of F_{ST} (a) and estimates of R_{ST} (b) based upon three microsatellite loci for twelve populations of *Hesperodiaptomus shoshone* in the Elk Mountains of Colorado. Significant pairwise Θ estimates of F_{ST} and estimates of R_{ST} are indicated in bold.

a) Fst estimates

	1	2	3	4	5	6	7	8	9	10	11
1- Schofield Pass											
2- Hasley Basin #8	**0.0741**										
3- Peeler Basin	**0.2965**	**0.3869**									
4- Redwell Basin	**0.3014**	**0.5014**	**0.4349**								
5- Mexican Cut U2	**0.2227**	**0.251**	**0.4173**	**0.4833**							
6- Kettle pond K4	**0.4308**	**0.546**	**0.5333**	**0.4977**	**0.4436**						
7- Mexican Cut L8	**0.2985**	**0.3243**	**0.5321**	**0.582**	**0.0338**	**0.5895**					
8- Mexican Cut L11	**0.1542**	**0.1035**	**0.3474**	**0.4921**	**0.1849**	**0.5303**	**0.2543**				
9- Mexcican Cut L15	**0.266**	**0.2879**	**0.4917**	**0.5458**	**0.0212**	**0.5445**	-0.0144	**0.216**			
10- Snodgrass	**0.2064**	**0.3187**	**0.3173**	**0.322**	**0.3487**	**0.3956**	**0.447**	**0.3049**	**0.4122**		
11- Mexican Cut U1	**0.3909**	**0.4488**	**0.4763**	**0.5435**	**0.4767**	**0.5328**	**0.5734**	**0.437**	**0.5355**	**0.3154**	
12- East Maroon Pass	**0.078**	**0.0445**	**0.3211**	**0.4314**	**0.1398**	**0.398**	**0.2213**	**0.0814**	**0.1864**	**0.2442**	**0.4055**

b) Rst estimates

	1	2	3	4	5	6	7	8	9	10	11
1- Schofield Pass											
2- Hasley Basin #8	0.1389										
3- Peeler Basin	**0.189**	**0.3157**									
4- Redwell Basin	**0.6235**	**0.751**	**0.5469**								
5- Mexican Cut U2	-0.0042	**0.1637**	**0.1415**	**0.5954**							
6- Kettle pond K4	**0.306**	**0.5371**	**0.4996**	**0.6031**	**0.3015**						
7- Mexican Cut L8	0.0252	**0.1096**	**0.1829**	**0.6805**	0.014	**0.4144**					
8- Mexican Cut L11	**0.2513**	**0.2438**	**0.1277**	**0.7082**	**0.2186**	**0.5923**	**0.2035**				
9- Mexciican Cut L15	0.0082	**0.1201**	**0.1689**	**0.6607**	-0.0006	**0.3852**	-0.0141	**0.2076**			
10- Snodgrass	0.0367	**0.083**	**0.0197**	**0.2745**	0.013	**0.1667**	0.0208	0.0476	0.0193		
11- Mexican Cut U1	**0.3545**	**0.436**	**0.4223**	**0.6206**	**0.3898**	**0.5385**	**0.4495**	**0.4926**	**0.4283**	**0.2711**	
12- East Maroon Pass	0.0735	**0.0979**	**0.2529**	**0.6547**	0.0722	**0.3323**	0.0441	**0.216**	0.0506	0.0442	**0.4534**

The genetic distances ranged from 0.005 between Mexican Cut lower 15 pond and lower 8 pond to 1.487 between upper pond 1 and upper pond 2 of the Mexican Cut for Nei's standard genetic distances (Table 7). While distances ranged from 0.0018 between ponds L15 and L8 of the Mexican Cut, to 2.674 between East Maroon Pass and Redwell Basin pond 2 for delta μ^2 distances. Table 7 shows lower matrices for both genetic distances for each population pair.

UPGMA dendograms based on two genetic distances used in this study were not congruent (Figure 4). However, some similarities are observed. The UPGMA dendogram based on Nei's genetic distances revealed a cluster of populations located in the Mexican Cut and Hasley Basin, which are geographically very close together. In this analysis, pond U1 from the Mexican Cut did not group with other ponds in the Mexican Cut and occupied a basal position on the dendrogram. A similar Mexican cut cluster was observed with the delta μ^2 genetic distances but pond L11 and U1 from the Mexican Cut did not cluster within this group. In the delta μ^2 genetic distance based dendrogram, the Redwell basin population occupied the basal position.

Table 7

Pairwise standard genetic distances (Nei, 1978) (a) and delta μ^2 genetic distances (b) based upon three microsatellite loci for twelve populations of *Hesperodiaptomus shoshone* in the Elk Mountains of Colorado.

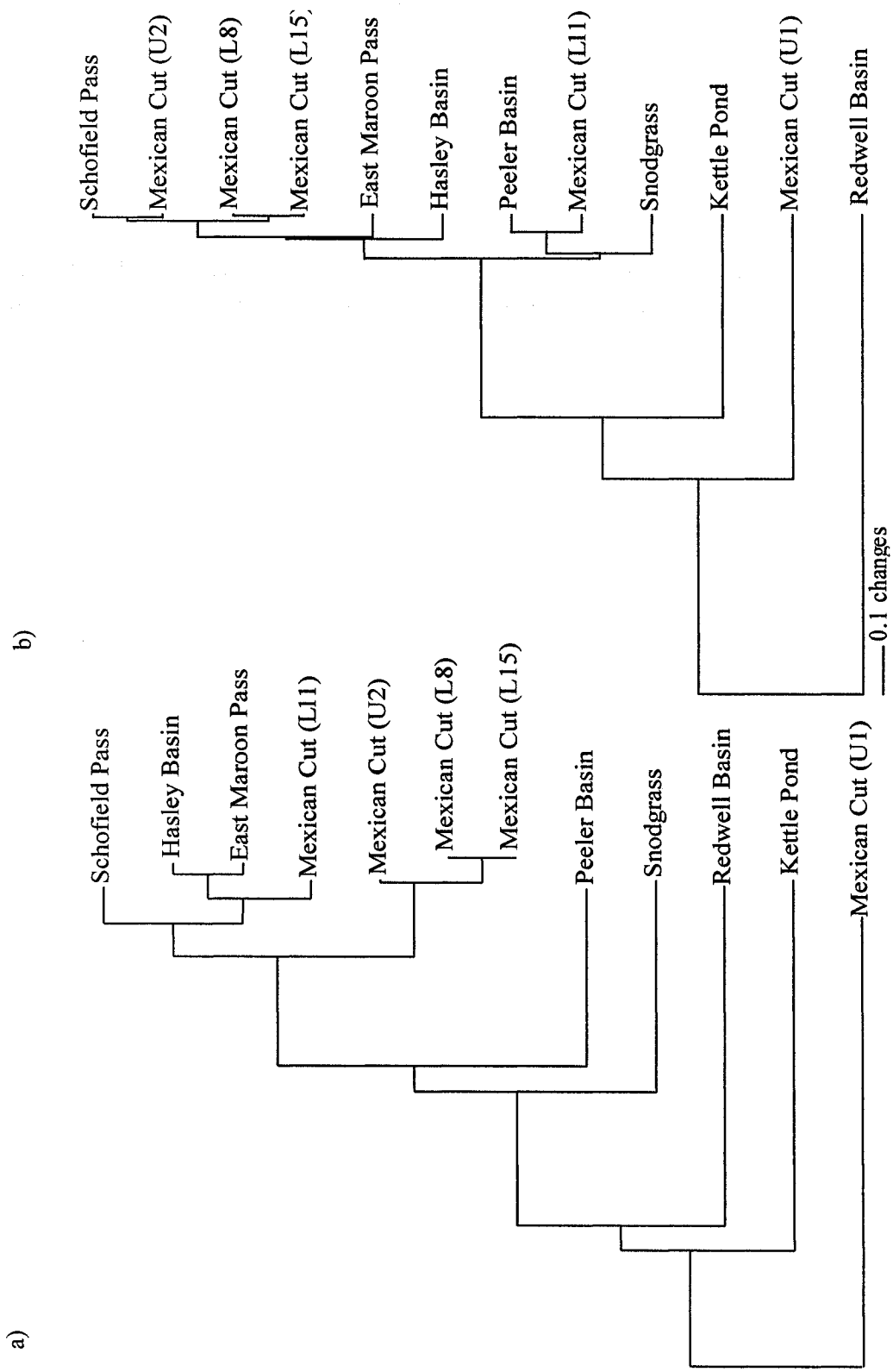
a) Nei's genetic distance

	1	2	3	4	5	6	7	8	9	10	11
1- Schofield Pass											
2- Hasley Basin #8	0.96222										
3- Peeler Basin	0.77314	0.76999									
4- Redwell Basin	0.71399	0.54774	0.64636								
5- Mexican Cut U2	0.83977	0.87055	0.71845	0.54988							
6- Kettle pond K4	0.54115	0.50757	0.52107	0.49086	0.64903						
7- Mexican Cut L8	0.83579	0.87563	0.69554	0.52386	0.99274	0.55845					
8- Mexican Cut L11	0.90933	0.95885	0.80994	0.57144	0.91414	0.54542	0.91465				
9- Mexcican Cut L15	0.83805	0.87631	0.69969	0.52899	0.99373	0.57561	1.00277	0.91709			
10- Snodgrass	0.77191	0.73076	0.72732	0.65121	0.66937	0.55866	0.64341	0.75248	0.64835		
11- Mexican Cut U1	0.50585	0.58007	0.51922	0.22722	0.49727	0.32386	0.47171	0.60647	0.47999	0.59964	
12- East Maroon Pass	0.94262	0.97974	0.78397	0.57594	0.92192	0.66155	0.90981	0.9622	0.91392	0.76573	0.55944

b) Delta mu ^2 genetic distance

	1	2	3	4	5	6	7	8	9	10	11
1- Schofield Pass											
2- Hasley Basin #8	0.1225										
3- Peeler Basin	0.2165	0.246									
4- Redwell Basin	1.9974	2.5515	1.248								
5- Mexican Cut U2	0.0125	0.1476	0.1584	1.7897							
6- Kettle pond K4	0.665	1.2955	1.2987	2.4669	0.655						
7- Mexican Cut L8	0.034	0.0621	0.1498	2.0365	0.0245	0.8751					
8- Mexican Cut L11	0.2213	0.0935	0.0741	1.908	0.1884	1.5285	0.1044				
9- Mexciican Cut L15	0.0213	0.0768	0.1495	1.9772	0.0133	0.8143	0.0018	0.1229			
10- Snodgrass	0.1894	0.3286	0.1223	1.3987	0.1051	0.8255	0.1234	0.202	0.1206		
11- Mexican Cut U1	0.7997	0.837	0.9261	2.5839	0.9313	2.1583	0.9724	0.9862	0.9382	1.4623	
12- East Maroon Pass	0.1196	0.1118	0.3807	2.6748	0.1179	0.8526	0.063	0.2432	0.0741	0.2297	1.3677

Figure 4 UPGMA dendogram depicting a) standard genetic distances (Nei, 1978) and b) delta μ^2 genetic distances between twelve *Hesperodiaptomus shoshone* populations based on three polymorphic nuclear microsatellite loci.



Genetic differentiation

F_{IS}, F_{IT}, and F_{ST} values by locus ranged from 0.043 - 0.402, 0.458 - 0.622, and 0.369 – 0.433 respectively (Table 8). Values of F_{IT} and F_{ST} estimates did not differ among the three loci however, the F_{IS} value was lower for the HSO7 locus compared to HSO1 and HSO2. The overall estimates of F_{IT} and F_{ST} were 0.492, and 0.379 respectively. F_{IS} values computed by population ranged from -0.119 in the Snodgrass population to 0.465 in the Schofield Pass population.

Table 8
F coefficients calculated for three polymorphic *Hesperodiaptomus shoshone* nuclear microsatellite loci based on 355 individuals from twelve Colorado populations.

Locus	F_{IS}	F_{IT}	F_{ST}
HSO1	0.228	0.508	0.363
HSO2	0.150	0.477	0.384
HSO7	0.043	0.458	0.433
Overall	0.181	0.492	0.379

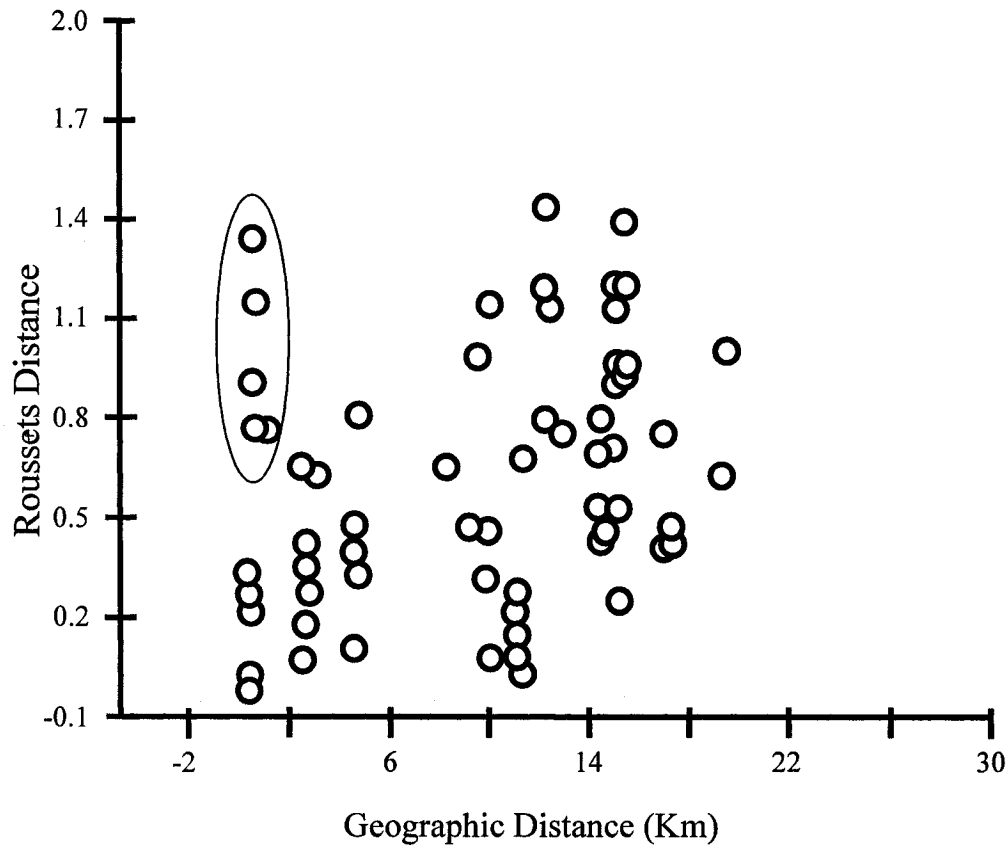
Table 9
F_{IS} coefficients calculated for each population for three polymorphic *Hesperodiaptomus shoshone* nuclear microsatellite loci based on 355 individuals.

Population	F_{IS}
Schofield Pass	0.465
Hasley Basin	0.236
Peeler Basin	0.099
Redwell Basin	0.073
Mexican Cut U2	0.179
Kettle Pond K4	0.268
Mexican Cut L8	0.315
Mexican Cut L11	0.100
Mexican Cut L15	0.082
Snodgrass	-0.119
Mexican Cut U1	0.286
East Maroon Pass	0.221

Isolation by Distance

Genetic differentiation between population pairs increased significantly with geographical distance (R^2= 0.112, p = 0.028; Figure 5). However, the scatter plot does not show a consistent pattern across the geographical distances, especially at short geographic distances, which is reflected by a lower R^2 value of 0.112. Since one of the ponds of the Mexican Cut (pond Upper 1) was very divergent compared to other geographically proximate ponds (e.g. Upper 2 (U2) see circled populations in Figure 5), the analysis was repeated without this basin. The R^2 between pairwise estimates of genetic and geographic distances doubled after removal of the pond U1 population from the analysis (R^2= 0.269, p = 0.001).

Figure 5
Isolation by Distance for twelve populations of *Hesperodiaptomus shoshone* for three microsatellite loci (R^2= 0.112, p = 0.028). The circled population pairs represent relationships between pond U1 and other ponds of the Mexican Cut.



Discussion

The populations of *H. shoshone* examined in this study showed relatively low levels of genetic diversity. Using microsatellite markers, the observed heterozygosity (H_O) values of 0.1 – 0.56 with a mean of 0.26 are much lower than those reported for other diaptomid species. For example, H_O values for the European freshwater diaptomid *Eudiaptomus gracilis* ranged between 0.53 and 0.87 with a mean of 0.72, and those for *Eudiaptomus graciliodes* ranged between 0.57 and 0.87 with a mean of 0.72 (Zeller *et al.*, 2006). Genetic diversity of *H. shoshone* (mean H_E: 0.320, Table 3) in the populations examined was also lower than for *E. gracilis* and *E. graciliodes* (H_E: 0.85 and 0.84 respectively) (Zeller *et al.*, 2006). Zeller *et al.* (2006) sampled copepods from large permanent European lakes, therefore the low genetic diversity observed in *H. shoshone* may be the consequence of the isolated temporary alpine habitats they occupy. No other studies of freshwater diaptomids based upon microsatellite markers exist for comparison. The direct comparisons of diversity levels should be interpreted cautiously because different microsatellite loci may generate different levels of variation.

Boileau (1989) examined genetic variation for seven species of *Hesperodiaptomus* in arctic pond populations with allozyme markers. The genetic diversity was lower than reported in this study with expected heterozygosities ranging between 0.043 and 0.138 and a mean of 0.09 (Table 3). Although it is difficult to compare the absolute values of genetic diversity measures among markers, comparative studies between allozyme and microsatellite markers revealed that microsatellites show higher levels of genetic variation than allozymes with respect to both observed and expected heterozygosities (Terauchi and Konuma, 1994; Sanchez *et al.*, 1996; Sun *et al.*, 1998). Thus, the observed low genetic diversity may not only

be a characteristic of *H. shoshone* but could also a general trend within the genus. However, these conclusions need to be validated with additional molecular analysis of *H. shoshone* occupying various habitats such as large lakes.

A study by Bohonak (1998) based on the same system of ponds in Colorado found low heterozygosity values for *Branchinecta coloradensis* compared to other fairy shrimp populations. Bohonak (1998) concluded that the low heterozygosity might be consistent with a relatively recent colonization from unglaciated populations from lower elevations. However, in this study no consistent patterns between altitude and genetic diversity (H_E) were observed (R^2=0.022, p=0.642).

Another factor possibly influencing the reduction in genetic diversity in some ponds is predation. Previous studies on other species have demonstrated that populations known to have experienced a reduction in population size often show reduced genetic diversity (Bouzat *et al.,* 1998). Two Mexican Cut Nature Preserve ponds, lower 8 and 11 are fixed and homozygous for two loci (Figures 3 b and c). Allelic diversity among *H. shoshone* was lowest in lower pond 8. Both of these ponds also contain the highest density of *Ambystoma tigrinum* hatchling larvae of all the ponds examined at the Mexican Cut (Table 2). These actively feed on *H. shoshone.* Predation has been considered to be an important factor influencing species composition in high altitude zooplankton communities (Dodson, 1970; Sprules, 1972). Vertebrate predators such as fish and salamanders affect zooplankton assemblages by size-selective predation of larger species, causing a size shift towards smaller species (Brooks and Dodson, 1965; Anderson, 1972) or the extirpation of local populations all together (Parker *et al.,* 1996).

After bottleneck events, reduction of both heterozygosity and allelic diversity is often observed (England *et al.,* 2004). Although related, allelic diversity and heterozygosity describe different aspects of the genetic diversity harboured by a population. Wright (1931) showed that heterozygosity is lost at a rate dependent on the effective population size during the bottleneck. However, short, severe bottlenecks are predicted to reduce allelic diversity more rapidly than heterozygosity (Nei *et al.,* 1975) because rare alleles, which contribute little to average heterozygosity, are easily lost in a bottleneck. Lower pond 8 of the Mexican Cut has the highest density of salamander larvae and also the lowest genetic and allelic diversity observed (Table 2, Table 3). The predation pressure is so intense that some years *H. shoshone* is eliminated from the water body by August before the pond dries (personal observation).

Hesperodiaptomus shoshone produces only diapausing eggs in these ponds (Maly and Maly, 1974) which hatch in the following growing season. The resting eggs accumulate in the sediments and are comparable to the seed banks of plants. The presence of egg banks can have important effects on the rate of evolution by lengthening the generation time which can lead to a slowing of the rate of response to natural selection (Hairston and De Stasio, 1988) or by enhancing the maintenance of genetic variation (Ellner and Hairston 1994). The presence of an egg bank should, therefore rescue populations from the severity of population size bottlenecks. However, as the number of resting eggs becomes depleted, the effects of predation should be more noticeable. Diapausing eggs are depleted from the sediments over time through hatching, predation and mortality (Hairston *et al.,* 1995 Parker *et al.,* 1996). In addition, low population numbers from increased predation pressure should further reduce the deposition of eggs into the egg bank. It is possible that under continuous

predation pressure, the egg bank is not sufficient to rescue the genetic diversity of *H. shoshone* in these ponds.

Observed values of heterozygosity were generaly lower than Nei's (1978) unbiased expected heterozygosity (H_E) measure (Table 3). The genotypic frequencies showed a significant deficit of heterozygotes relative to Hardy Weinberg expectations for 8 populations for the HSO1 locus and 3 populations for the HSO2 locus (Table 3). The departure of genotypic frequencies from Hardy Weinberg Expectations might be due to a variety of causes. An excess of heterozygotes might be due to the presence of overdominant selection or the occurrence of outbreeding. Whereas a deficit of heterozygosity could be observed if there is inbreeding in the population (assortative mating), the locus is under selection, the Wahlund effect occurs (population subdivision into separate breeding units), underdominant selection and selection-induced microscale differentiation is happening (Epperson, 1990; Knowles, 1991), or if null alleles exist. Although several of these scenarios are unlikely, they cannot be completely ruled out.

Although no detailed analysis of the breeding system has been performed in *H. shoshone* a heterozygote deficit was consistently observed for one of the three loci which indicates that inbreeding is not a likely explanation as inbreeding should affect all loci equally. On the other hand, the inbreeding coefficients calculated by the hierarchal F statistics suggest otherwise (F_{IT} =0.492 and F_{IS} ranging from -0.119 to 0.465 per population, Table 8 and Table 9). Redwell Basin was the only population with a negative F_{IS} value of -0.119. Wright's hierarchal F statistics (Wright, 1951), F_{IS} and F_{IT} represent correlations of alleles within individuals relative to the group and total population, respectively, and theoretically can range -1.0 to 1.0 (Sugg *et al.*, 1996). Positive F_{IS} values represent a reduced

level of heterozygosity relative to the pond population, while positive values of F_{IT} represent reduced heterozygosity within the entire network of ponds (Storz, 1999). Thus, positive F_{IS} and F_{IT} values are usually interpreted as indicators of inbreeding (Dobson *et al.,* 1997). Negative values, on the other hand, represent excess levels of heterozygosity and are often interpreted as evidence of inbreeding avoidance or outcrossing (Storz, 1999). Mating experiments have shown size selective assortative mating for this species (Marszalek, unpublished data) and evidence for mate choice in females but no inbreeding avoidance experiments have been conducted to date. Mating experiments on the harpacticoid copepod *Tigriopus californicus* showed inbreeding avoidance and outbreeding preference when females were given a choice of males (Palmer and Edmans, 2000). Although, assortative mating might cause inbreeding, it is possible that the small ponds especially those subjected to predation and population fluctuations coupled with low gene flow (see below) experience inbreeding due to the isolated nature of the habitats they occupy.

A heterozygote deficit can also be produced when different populations are combined and treated as a single population in the genetic analysis (Wahlund effect; Wahlund, 1928). Although this may appear unlikely for this study, due to the small size of ponds, some evidence suggest that subpopulations could be maintained even in such habitats. Marszalek (2002) and Aguilera Salas (2002) found that *H. shoshone* occur in aggregations that are maintained throughout the season which offer a reproductive advantage to the individuals within them. Observations also suggest the possibility of two cohorts in some basins at the Mexican Cut (Marszalek, unpublished data) which could cause the Wahlund effect.

The population subdivision could also lead to a statistically significant deviation from linkage equilibrium. Only four populations Peeler Basin, Redwell Basin, Snodgrass and East

Maroon pass showed statistically significant levels of linkage disequilibrium (LD) for the HSO1 and HSO2 locus (Table 5). However, only one of the four populations (East Maroon Pass), with significant levels of linkage disequilibrium showed reduced heterozygosity values (Table 3) making this explanation unlikely. LD might also reflect an actual physical association between loci on chromosomes (Hartl and Clark, 1997), but since only four populations show deviations from LD it is an unlikely explanation.

Finally, it is possible that heterozygotes may have been missed due to null (nonamplifying) alleles, possibly produced by mutations in the flanking region of the microsatellite loci preventing one or both of the primers from binding (Holm *et al.,* 2001). Although null alleles are commonly reported (Shaw *et al.,* 1999; McGoldrick *et al.,* 2000), the effects of null alleles in datasets on population genetic estimates have received relatively little attention. Genetic studies focussing on segregation patterns of alleles within a family would be valuable to discern any null alleles.

Population genetic structure

The distribution of genetic variation among *H. shoshone* populations was examined by using two measures F_{ST} and R_{ST}. These measures differ in their assumptions about the mode of mutation; F_{ST} is the appropriate measure of differentiation if mutations occur equally to all alleles (IAM), whereas R_{ST} is better if mutations occur more commonly in alleles with similar numbers of repeats (SMM). In practice, the best measure to use is far from clear (Paetkau et al., 1997, Valsecchi *et al.,* 1997). Balloux and Lougon-Moulin (2002) showed that F_{ST} is more efficient in the case of high levels of gene flow whereas R_{ST} better reflects population differentiation under low gene flow. However F_{ST} is predicted to perform better with smaller

sample sizes or fewer loci (Gaggiotti *et al.,* 1999). The traditional measure of genetic divergence among groups, F_{ST}, represents allelic correlations within the group relative to the total population and can attain a theoretical minimum of zero (indicating complete panmixia) and a maximum of 1.0 (complete genetic differentiation; Sugg *et al.,* 1996; Storz, 1999). Negative values are also possible (Wright, 1951), when subpopulation level estimates of heterozygotes are higher than the total population values. The overall F_{ST} value of 0.379 and overall R_{ST} value of 0.332 indicate that on average about 35.5 % of the genetic variation can be accounted for among subpopulations. According to Wright's (1978) guidelines F_{ST} values > 0.25 indicate very great genetic differentiation and suggest very low gene flow between the populations examined.

A significant isolation by distance pattern was observed at the scale of the present study (Figure 5). Isolation by distance is expected between populations if gene flow is prevalent but dependent on geographical distance because the exchange of individuals should be more frequent among nearby populations than populations located further away (Hutchinson and Templeton, 1999). The microsatellite data in this study supports this hypothesis, with Rousset's genetic distance generally increasing between geographically distant samples than those in close proximity. The variance observed at short geographic distances could be the result of genetic drift or signatures of past founder effects. Persistent founder effects should be largely independent of geographical distance (De Meester *et al.,* 2002, Zeller *et al,* 2006) and completely isolated populations are expected to exhibit no isolation by distance patterns. For example, when a very divergent pond from the Mexican Cut, Upper 1, was removed from the analysis the IBD relationship almost doubled (R^2= 0.112

to R^2= 0.269). The present data indicates that gene flow is restricted to nearby populations but at the same time genetic drift in some populations may obscure this relationship.

The failure of populations to cluster by geographical area in the UPGMA analysis further supports the importance that genetic drift and founder effects may play in these populations (Figure 4). Bohonak (1998) found similar results in the analysis of the genetic structure of *Branchinecta coloradensis* from the same set of ponds.

Some interesting patterns in genetic structure were observed within the Mexican Cut Nature Preserve (Figure 2). The pond complex is the most important component of the Mexican Cut ecosystem and the available long-term physical and biological datasets provide the opportunity to examine various hypotheses. As mentioned previously, the presence of *Ambystoma tigrinum* larvae may cause reductions in the genetic diversity of *H. shoshone.* From the allele frequencies and F_{ST} and R_{ST} values, dispersal at the Mexican cut occurs between physically interconnected ponds. The pond lower 8, which flows down to lower 15 share both alleles and F_{ST} and R_{ST} estimates are -0.0144 and -0.0141 respectively indicating high gene flow. On the other hand, ponds U1 and U2 of the Mexican Cut, which are separated by less than 20 m share only 2 of 8 alleles for the HSO1 locus. F_{ST} and R_{ST} estimates for this population pair are 0.4767 and 0.3898 respectively indicating lack of gene flow. The most surprising fact is the presence of high genetic structuring at the Mexican Cut. During the period of last four decades, the ponds have been sampled and examined by many biologists and it would be fair to assume that this sampling could promote the dispersal of eggs or individuals between the ponds via boots or plankton nets.

Boileau *et al.* (1992) found that the extent of genetic variation observed among conspecific populations of different species of copepods was inconsistent with predictions

based on their local distribution patterns. For example, in Churchill, Manitoba populations *H. victoriaensis* (Reed, 1958) have a restricted distribution and show less genetic variation than *Leptodiaptomus tyrrelli* (Poppe, 1888) which has broader distribution patterns. Boileau *et al.* (1992) suggested that the level of genetic differentiation among populations does not need to reflect their dispersal ability. Zeller *et al.* (2006) on the other hand examined the genetic structure of two freshwater copepods and found that, at the scale of this study, microsatellite F_{ST} estimates were in the 0.001-0.1 range but genetic structure was randomly distributed at this geographic scale. The species which produces resting stages had lower F_{ST} values, suggesting a larger gene flow possibility.

In conclusion, the present study shows that substantial genetic differences are common among copepods populations in localized alpine habitats presumably due to low gene flow and genetic drift. The temporary habitats in alpine regions promote the genetic divergence of populations possibly leading to the evolution of "species". The genetic processes underlying the genetic differentiation of copepods may play a similar role in other organisms in other aquatic ecosystems and may have a significant influence on the evolution of biological diversity in alpine ecosystems.

References

Aguilera Salas, M.P. (2002) Distribution pattern of *Hesperodiaptomus shoshone* and *Leptodiaptomus coloradensis* and its effects on their mating and reproductive potential. MSc Thesis. Concordia University, Montreal, Quebec. 74pp.

Anderson, R.S. (1972) Zooplankton composition and change in an alpine lake. *Int. Ver. Theor. Angew. Limnol. Verh.* **18**: 264-268.

Anderson, R.S. (1980) Relationships between trout and invertebrate species as predators and the structure of the crustacean and rotiferan plankton in mountain lakes. In W.C. Kerfoot (Ed.), Evolution and Ecology of Zooplankton Communities (p.635-641). University Press of New England, Hanover.

Balloux, F. and N. Lugon-Moulin (2002) The estimation of population differentiation with microsatellite markers. *Molecular Ecology* **11**:155–165.

Barton, N.H. and G.M. Hewitt (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16**:113–148.

Bohonak, A.J. (1998) Genetic population structure of the fairy shrimp *Branchnecta coloradensis* (Anostraca) in the Rocky Mountains of Colorado. *Canadian Journal of Zoology* **76**:2049-2057.

Bohonak, A.J. (2002) IBD (Isolation By Distance): a program for analyses of isolation by distance. *Journal of Heredity* **93**:153-154.

Bohonak, A. and D. Jenkins (2003) Ecological and evolutionary significance of dispersal by freshwater invertebrates. *Ecology Letters* **6**:783-796.

Boileau, M.G. (1989) Biochemical genetic variation in the freshwater copepoda: evolutionary consequences of passive dispersal in freshwater zooplankton. Phd Thesis. University of Windsor, Windsor, Ontario. 167pp.

Boileau, M.G., P.D.N. Hebert and S.S. Schwartz (1992) Non-equilibrium gene frequency divergence: persistent founder effects in natural populations. *Journal of Evolutionary Biology* **5**:25-39.

Bouzat, J.L., H.H. Cheng, H.A. Lewin, R.L. Westemeier, J.D. Brawn and K.N. Paige (1998) Genetic evaluation of a demographic bottleneck in the greater prairie chicken. *Conservation Biology* **12**(4):836-843.

Brooks, J.L. and S.I. Dodson (1965) Predation, body size, and composition of plankton. *Science* **150**:28-35.

Bucklin, A. (2000) Methods for Population Genetic Analysis of Zooplankton. Chapter 11. In R. Harris, P. Wiebe, J. Lenz, H. Skjoldal, and M. Huntley (Eds.), The Zooplankton Methodology Manual, International Council for the Exploration of the Sea (p 533 – 570). Academic Press, London.

Casgrain, P. and P. Legendre (2001) The R-Package for multivariate and spatial analysis and User's manual, Early draft. Département de sciences biologiques Université de Montréal. (http://www.fas.umontreal.ca/BIOL/legendre/) [accessed January, 2003]."

De Meester, L. (1996) Local genetic differentiation and adaptation in freshwater zooplankton populations: patterns and processes. *Ecosience* **3**(4):385-399.

DeMeester, L., A. Gómez, B. Okamura and K. Schwenk (2002) The Monopolization Hypothesis and the dispersal–gene flow paradox in aquatic organisms. *Acta Oecologica* **23**: 121-135.

Dobson, F.S., R.K. Chesser, J.L. Hoogland, D.W. Sugg and D.W. Foltz (1997) Do Black-Tailed Prairie Dogs Minimize Inbreeding? *Evolution* **51**(3):970-978.

Dodson, S.I. (1970) Complementary feeding niches sustained by size-selective predation. *Limnology and Oceanography* **15**(1):131-137.

Edmondson, W.T. and H.B. Ward (1959) Fresh-water biology. New York, Wiley. 1111pp.

Ellner, S., N.G. Hairston Jr. (1994) Role of overlapping generations in maintaining genetic variation in a fluctuating environment. *The American Naturalist* **143**(3):403-417.

England, P.R., G.H.R. Osler, L.M. Woodworth, M.E. Montgomery, D.A. Briscoe and R. Frankham (2004) Effects of intense versus diffuse population bottlenecks on microsatellite genetic diversity and evolutionary potential. *Conservation Genetics* **4**(5):595-604.

Epperson, B.K. (1990) Spatial autocorrelation of genotypes under directional selection. *Genetics* **124**:757-771.

Fu, Y.X. and R. Chakraborty (1998) Simultaneous estimation of all the parameters of a stepwise mutation model. *Genetics* **150**:487-497.

Fureder, L., R. Ettinger, A. Boggero, B. Thaler and H. Thies (2006) Macroinvertebrate diversity in alpine lakes: Effects of altitude and catchment properties. *Hydrobiologia* **562**:123-144.

Gaggiotti, O.E., O. Lange, K. Rassmann and C. Gliddon (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology* **8**:1513-1520.

Goldstein, D.B., A.R. Linares, L.L. Cavalli-Sforza and M.W. Feldman (1995) An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139**:463-471.

Goodman, S.J. (1997) RST CALC: A collection of computer programs for calculating unbiased estimates of genetic differentiation and determining their significance for microsatellite data. *Molecular Ecology* **6**:881-885.

Goudet, J. (1995) FSTAT Version 1.2: A computer program to calculate F-statistics. *Journal of Heredity* **86**:485-486.

Hairston Jr., N.G., and B.T. De Stasio Jr.(1988) Rate of evolution slowed by a dormant propagule pool. *Nature* **336**:239-242.

Hairston Jr., N.G., R.A. Kearns and D.R. Engstrom (1995) Age and survivorship of diapausing eggs in a sediment egg bank. *Ecology* **7**(6):1706-1711.

Hartl, D. and A. Clark (1997) Principles of population genetics. Third edition.

Higgins, K. and M. Lynch (2001) Metapopulation extinction caused by mutation accumulation. *Proceedings of the National Academy of Sciences of the USA* **98**:2928–2933.

Hinden, H., B. Oertli, N. Menetrey, L. Sager and J.B. Lachavanne (2005) Alpine pond biodiversity: what are the related environmental variables? *Aquatic Conservation: Marine and Freshwater Ecosystems* **15**(6):613-624.

Hofer, R., R. Lackner, J. Kargl, B. Thaler, D. Tait, L. Bonetti, R. Vistocco and G. Flaim (2001) Organo-chlorine and metal accumulation in fish (*Phoxinus phoxinus*) along a northsouth transect in the Alps. *Water Air and Soil Pollution* **125**:189-200.

Holm, L.E., V. Loeschcke and C. Bendixen (2001) Elucidation of the molecular basis of a null allele in a rainbow trout microsatellite. *Marine Biotechnology* **3**:555-560.

Holm, S. (1979) A simple sequentially rejective multiple test procedure. *Scand J Stat* **6**:65-70.

Hutchinson, D.W. and A.R. Templeton (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* **53**(6):1898-1914.

Kalinowski, S.T. (2002) Evolutionary and statistical properties of three genetic distances. *Molecular Ecology* **11**:1263-1273.

Kimura, M. and J. Crow (1964) The number of alleles that can be maintained in a finite populations. *Genetics* **49**:725–738.

Kimura, M. and T. Otha (1978) Stepwise mutation model and distribution of allelic frequencies in a finite populations. *Proceedings of the National Academy of Sciences of the USA.* **75**:2868–2872.

Knowles, P. (1991) Spatial genetic structure within two natural stands of black spruce (*Picea mariana* (Mill.). *Silvae Genetica* **40**:13-19.

Lee, C.E. and B. Frost (2002) Morphological stasis in the *Eurytemora affinis* species complex (Copepoda: Temoridae). *Hydrobiologia* **480**:111-128.

Lewis, P.O. and D. Zaykin (2001) GENETIC DATA ANALYSIS: computer program for the analysis of allelic data, version 1.0 (d16c). Free program distributed by the authors from http://lewis.eeb.uconn.edu/lewishome/software.html."

Maly, E.J., M.P. Maly (1974) Dietary differences between two co-occuring calanoid copepod species. *Oecologia* **17**:325-333.

Mantel, N. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**:209-220.

Marszalek, M.A. (2002). Proximate factors influencing the spatial distribution of a high altitude copepod : *Hesperodiaptomus shoshone.* MSc Thesis. Concordia University, Montreal, Quebec. 89pp.

McGoldrick, D.J., D. Hedgecock, J. English, P. Baoprasertkul and R.D. Ward (2000) The transmission of microsatellite alleles in Australian and North American stocks of the Pacific oyster (*Crassostrea gigas*): selection and null alleles. *Journal of Shellfish Research* **19**:779-788.

McNaught, A.S., D.W. Schindler, B.R. Parker, A.J. Paul, R.S. Anderson, D.B. Donald and M. Agbeti (1999) Restoration of the food web of an alpine lake following fish stocking. *Limnology and Oceanography* **44**:127-136.

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

Nei, M. (1987). Molecular evolutionary genetics. Columbia, New York.

Nei, M., T. Maruyama and R. Chakraborty (1975) The bottleneck effect and genetic variability in populations. *Evolution* **29**(1):1-10.

Neigel, J.E. (2002) Is FST obsolete? *Conservation Genetics* **3**:167-173.

Paetkau, D., L.P. Waits, P.L. Clarkson, L. Craighead and C. Strobeck (1997) An Empirical evaluation of genetic distance statistics using microsatellite data from bear (Ursidae) Populations. *Genetics* **147**(4):1943–1957.

Palmer, C.A. and S. Edmands (2000) Mate choice in the face of both inbreeding and outbreeding depression in the intertidal copepod *Tigriopus californicus*. *Marine Biology* **136**(4):693-698.

Parker, B.R., F.M. Wilhelm and D.W. Schindler (1996) Recovery of *Hesperodiaptomus arcticus* populations from diapausing eggs following elimination by stocked salmonids. *Canadian Journal of Zoology* **74**:1292–1297.

Paul, A.J., P.R. Leavitt, D.W. Schindler and A.K. Hardie (1995) Direct and indirect effects of predation by a calanoid copepod (subgenus: *Hesperodiaptomus*) and nutrients in a fishless alpine lake. *Canadian Journal of Fisheries and Aquatic Science* **52**:2628-2638.

Psenner, R. (2002) Alpine Waters in the interplay of global change. Complex links – simple effects? In K.W. Steininger and H. Weck-Hannemann, E. Elgar (Eds.), Global Environmental Change in Alpine Regions. Impact, Recognition, Adaptation, and Mitigation (p.14-40) Cheltenham.

Raymond, M. and F. Rousset (1995) GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**:248-249.

Rice, W.R. (1989) Analyzing tables of statistical tests. *Evolution* **43**:223-225.

Rousset, F. (1997) Genetic differentiation and estimation of gene flow from F statistics under isolation by distance. *Genetics* **145**:1219-1228.

Sanchez, J.A., C. Clabby, D. Ramos, G. Blanco, F. Flavin, E. Vazquez and R. Powell (1996) Protein and microsatellite single locus variability in *Salmo salar* L. (Atlantic salmon). *Heredity* 77:423-432.

Shaw, P.W., G.J. Pierce and P.R. Boyle (1999) Subtle population structuring within a highly vagile marine invertebrate, the veined squid *Loligo forbesi*, demonstrated with microsatellite DNA markers. *Molecular Ecology* **8**:407-417.

Slatkin, M. (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* **47**(1):264-279.

Slatkin, M. (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**:457-462.

Slatkin, M. and W.P. Maddison (1990) Detecting isolation by distance using phylogenies of genes. *Genetics* **126**(1):249-260.

Sprules, W.G. (1972) Effects of size-selective predation and food competition on high altitude zooplankton communities. *Ecology* **53**:375-386.

SPSS Inc. (1999) SPSS Base 11.5 for Windows. SPSS Inc., Chicago IL

Staton, J.L., B.E. Taylor, N.V. Schizas, R. Wetzer, T.C. Glenn and B.C. Coull (2003) Mitochondrial gene diversity of *Skistodiaptomus mississippiensis* in impoundments of the Upper Coastal Plain near Aiken, South Carolina, USA. *Archiv fur Hydrobiologie* **158**(2):215-231.

Stoddard, J.L. (1987) Microcrustacean communities of high-elevation lakes in the Sierra Nevada, California. *Journal of Plankton Research* **9**:631-650.

Storz, J.F. (1999) Genetic consequences of mammalian social structure. *Journal of Mammalogy* **80**:553-569.

Sugg, D.W., R.K. Chesser, F.S. Dobson and J.L. Hoogland (1996) Population genetics meets behavioral ecology. *Trends in Ecology and Evolution* **11**:338-342.

Sun, G.L., O. Diaz, B. Salomon and R. Bothmer (1998) Microsatellite variation and its comparison with allozyme and RAPD variation in *Elymus fibrosus* (Schrenk) Tzvel. (Poaceae) *Hereditas* **129**(3):275-282.

Swofford, D.L. (2001) PAUP* 4.0 - Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Assoc, Sunderland, MA.

Terauchi, R. and A. Konuma (1994) Microsatellite polymorphism in *Dioscorea tokoro*, a wild yam species. *Genome* **37**:794-801.

Valsecchi, E., P. Palsboll, P. Hale, D. Glockner-Ferrari, M. Ferrari, P. Clapham and *et al.* (1997) Microsatellite genetic distances between oceanic populations of the humpback whale (Megaptera *novaeangliae*) *Molecular Biology and Evolution* **14**:355-362.

Wahlund, S. (1928) Composition of populations from the perspective of the theory of heredity. *Hereditas* **11**:65-105.

Watras, C.J. and J.F. Haney (1980) Oscillations in the reproductive condition of *Diaptomus leptopus* (Copepoda: Calanoida) and their relation to rates of egg-clutch production. *Oecologia* **45**:94-103.

Weir, B.S. and C.C. Cockerham (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38**:1358-1370.

Wilson, M.M. and H.C. Yeatman (1959) Free-Living Copepoda. In W.T. Edmonson and H.B. Ward (Eds.) Ward's and Whipple's Freshwater Biology (p.735-794). John Wiley and Sons, New York.

Wissinger, S.A., A.J. Bohonak, H.H. Whiteman and W.S. Brown (1999) Subalpine wetlands in Colorado: Habitat Permanence, Salamander Predation, and Invertebrate Communities. Chapter 31. In D.P. Batzer, R.B. Rader and S.A. Wissinger (Eds.), Invertebrates in Freshwater Wetlands of North America: Ecology and Management (p. 798-790). John Wiley and Sons, New York.

Wright, S. (1931) Evolution in Mendelian populations. *Genetics* **16**(2):97-159.

Wright, S. (1951) The genetical structure of populations. *Annals of Eugenetics* **15**:323-354.

Wright, S. (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* **19**:395-420.

Wright, S. (1978) Evolution and the Genetics of Populations, vol. 4, Variability Within and Among Natural Populations. University of Chicago Press, Chicago, IL.

Yeh, F.C., R.C. Yang, T.B.J. Boyle, Z.H. Ye and J.X. Mao (1999) POPGENE, the user friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada.

Zeller, M., T.B.H. Reusch and W. Lampert (2006) A comparative population genetic study on calanoid freshwater copepods: Investigation of isolation-by-distance in two *Eudiaptomus* species with a different potential for dispersal. *Limnology and Oceanography* **51**(1):117-124.

General Conclusion

High altitude ecosystems are both spatially and temporally variable providing diverse habitats for many organisms and play a key role in the maintenance of biodiversity. The island like nature and fragility of these habitats makes them especially vulnerable to environmental change and has therefore become the focus of numerous conservation studies. Understanding evolutionary relationships and biodiversity patterns in high altitude animal and plant taxa is the first essential step for formulating effective conservation, monitoring, and restoration programs in these threes environments. Freshwater hesperodiaptomid copepods, an important component of high altitude aquatic systems of Western North America have proven to be an ideal group of organisms for examining the evolution and maintenance of biological diversity in these mountain environments.

The overall goal of my thesis was to study the molecular evolution of *Hesperodiaptomus shoshone* (S.A. Forbes, 1882) by analyzing genetic relationships at various hierarchal levels to gain insights into the evolution of biological diversity in high altitude aquatic ecosystems of Western North America.

In chapter 1, *H. shoshone* was studied in a broader evolutionary context, with the objective of reconstructing the phylogeny of the genus *Hesperodiaptomus*. This study employed molecular characters and elucidated the evolutionary relationships among species of the genus *Hesperodiaptomus*. Although the monophyly of genera is expected in cladistic analyses, it has never been explicitly tested in North American diaptomid

copepods. The molecular based phylogenetic trees supported the monophyletic origin of the genus and revealed two main clades showing congruence with morphological characters. Additionally, the DNA sequences resulting from this study provide a molecular database that can be used as DNA barcodes for species with difficult identifications.

In Chapter 2, I examined the phylogeography of the high altitude copepod, *H. shoshone*. The objectives were to determine the patterns of genetic differentiation of populations of *H. shoshone* based on mtDNA and rDNA data, to asses the influence of historical events on dispersal, vicariance and diversification throughout the range of this species in North America. The results of this study revealed a genetically structured distribution and revealed two reciprocally monophyletic groups. These two lineages showed strong concordance with major geographic areas of Northwestern North America. The North lineage was found to be restricted to Alberta, Idaho and Wyoming, while the South lineage was confined to California and Colorado. Judging from their levels of sequence divergence, the clades most likely represent long separated lineages and suggest cryptic speciation in this taxon.

In Chapter 3 and 4, I characterized and used microsatellite markers to study the genetic diversity of *H. shoshone* at a local scale within and among neighbouring watersheds of the Elk Mountains of Colorado. Genetic studies of populations at the local scale provide insight into the current micro-evolutionary processes that influence the current observed diversity. Within the South lineage, the Colorado populations of *H. shoshone* showed high level of genetic divergence and depauperate genetic diversity suggesting restricted gene flow and high genetic drift between populations.

Overall, the present study elucidates evolutionary relationships among hesperodiaptomid copepods and shows genetic structuring at various hierarchal levels. The restricted gene flow and the isolated nature of these high altitude habitats occupied by hesperodiaptomids promotes genetic diversification and specification in this group of organisms. It also provides the first necessary step in unravelling the biogeography and evolutionary history of North American diaptomids. Further studies encompassing diverse aquatic organisms will be valuable for furthering our understanding of the modes of speciation in these high altitude aquatic ecosystems.

Appendix

Appendix 1

Alignment of nuclear ribosomal DNA sequences of 9 species of *Hesperodiaptomus* used in Chapter 1 of the thesis.

```
                          10         20         30         40         50
                           .          .          .          .          .
H.kenai                   AACAAGGTTTCCGTAGGTGAACCTGCGGAAGGATCATTATCGATTACTAC
H.hirsutus                ..................................................
H.shoshone                ..................................................
H.eiseni                  ..................A...............................
H.arcticus                ..................................................
H.californiensis          ..................................................
H.nevadensis              ..................................................
H.victorianensis          ...............................................A...
H.franciscanus            ..................................................

                          60         70         80         90         100
                           .          .          .          .          .
H.kenai                   CGATGCTTAGGCAGAGGGAATTTCGTAATAACCTCGGGCTTCGCTGCCTC
H.hirsutus                ..................................................
H.shoshone                ..................................................
H.eiseni                  ..............................-..........T........
H.arcticus                ..............................-...................
H.californiensis          ..............................-...................
H.nevadensis              ..................................................
H.victorianensis          ..............................-...................
H.franciscanus            ............A.................-...................

                          110        120        130        140        150
                           .          .          .          .          .
H.kenai                   GGGACGTCGGCTTGCCTGCCGTTGTCTGATAAGTCGGGATTGGGGGCTTG
H.hirsutus                ....................C......C....T..............
H.shoshone                ............C......................T.....C.........
H.eiseni                  .....A......C......................T.....C..A......
H.arcticus                .....A......C......................T.....C..A......
H.californiensis          .....A......C......................T.....C..A......
H.nevadensis              .....A......C......T...............T.....C..A......
H.victorianensis          .....A..A..........T...............T.....C..A......
H.franciscanus            .....C......C..........................CC...C......

                          160        170        180        190        200
                           .          .          .          .          .
H.kenai                   CGGCTACCTGGCCGCTTGCTCTCGCCTGACTGCCCG-ATCGGAGGTGCTG
H.hirsutus                ...................................A-G.............
H.shoshone                ....................................-.............
H.eiseni                  ..................TC....C.........-.............
H.arcticus                ....................C....C.........-.............
H.californiensis          ....................C....C.........-.............
H.nevadensis              ....................C....C.........-.............
H.victorianensis          T...................C....C.G.......-.............
H.franciscanus            ............T..A.TTG.C......G......TTG............
```

Appendix 1 cont.

```
                        210        220        230        240        250
                         .          .          .          .          .
H.kenai                 TGCGCCGTTGTCGCCCCATGGGGTGAAAGGCGGTAATTCAGCTGTCAAAG
H.hirsutus              ................T................................
H.shoshone              ................TC..............C.............
H.eiseni                ................TC................................
H.arcticus              ................TC................................
H.californiensis        ................TC................................
H.nevadensis            ..............T.TC................................
H.victorianensis        ..............T.TC.A..............................
H.franciscanus          ..............TTC....C............C.....T.A....

                        260        270        280        290        300
                         .          .          .          .          .
H.kenai                 AGCCTCGCCTCCGGCTCGGGTTATGTGCTCCCCCCATGACGGGCTTGTGA
H.hirsutus              ....................T..............TT..........
H.shoshone              ..................................C.TT..........
H.eiseni                ..................................C.TT..........
H.arcticus              ....................................TT..........
H.californiensis        ....................................T...........
H.nevadensis            ....................................T..........G
H.victorianensis        ............T.......................T...........
H.franciscanus          ............-.........C............T...........

                        310        320        330        340        350
                         .          .          .          .          .
H.kenai                 AAAACGGTACATTCCTAA-AGAAAATGACTGGCTTTACGCCGGAAACCAC
H.hirsutus              ..............T.-.-.....A.........AG.............
H.shoshone              ..............T-T..............A................
H.eiseni                ................-.A..............A..............
H.arcticus              ..............T.-.A...-CA.........A..............
H.californiensis        .G............T.-.A....CA.........A..............
H.nevadensis            .G............T.-.A...-CA.........A..............
H.victorianensis        .G.........A.A...-.A...-CA.........A....T.........
H.franciscanus          .G..............C..............C.AG.............

                        360        370        380        390        400
                         .          .          .          .          .
H.kenai                 AAAAACAACACTATACAGTGGATCATTTGGCTCGGGGGTCGATGAAGAAC
H.hirsutus              ..................................................
H.shoshone              ..................................................
H.eiseni                ..................................................
H.arcticus              ..................................................
H.californiensis        ..................................................
H.nevadensis            ..................................................
H.victorianensis        ..................................................
H.franciscanus          ..................................................
```

Appendix 1 cont.

```
                          410        420        430        440        450
                           .          .          .          .          .
H.kenai                   GGAGCTAACTCCGTGACGCAATGTGAACTGCAGGACACATGAACATCGAC
H.hirsutus                ..................................................
H.shoshone                ..................................................
H.eiseni                  ..................................................
H.arcticus                ..................................................
H.californiensis          ..................................................
H.nevadensis              ..................................................
H.victorianensis          ..................................................
H.franciscanus            ..................................................

                          460        470        480        490        500
                           .          .          .          .          .
H.kenai                   ATTTTGAACGCATATTGCGGACATGTGCTTGTTGCATGTCCATGTTTGGG
H.hirsutus                ..................................................
H.shoshone                ..................................................
H.eiseni                  ..................................................
H.arcticus                ..................................................
H.californiensis          ..................................................
H.nevadensis              ..................................................
H.victorianensis          ..................................................
H.franciscanus            ..................................................

                          510        520        530        540        550
                           .          .          .          .          .
H.kenai                   TCAGGGTCGGTTGAAACATCCATCAAAACTCTTGCGAGTTTTGCCCTGGG
H.hirsutus                ..................................................
H.shoshone                ......................G.................C........
H.eiseni                  ......................G..........................
H.arcticus                ......................G..........................
H.californiensis          ......................G..........................
H.nevadensis              ......................G..........................
H.victorianensis          ......................G..........................
H.franciscanus            ......................G...TC......G....C........

                          560        570        580        590        600
                           .          .          .          .          .
H.kenai                   TCGTCCTGTCGGCTTTGGCCGGCGGGTGGCCTCAAGTTGACTTGTGCAGT
H.hirsutus                ..............C.....................................
H.shoshone                ..............C.....................................
H.eiseni                  ..............C.....................................
H.arcticus                ..............C.....................................
H.californiensis          ..............C.....................................
H.nevadensis              .............CC.....................................
H.victorianensis          .............CC.....................................
H.franciscanus            .............CC.....................................
```

Appendix 1 cont.

```
                          610        620        630        640        650
                           .          .          .          .          .
H.kenai                   GCTCTGCCAAAGACTCTTGGGTCGTAAAAAGCCTGGGTCTGGGTGGGACA
H.hirsutus                ...........................................
H.shoshone                ...........................................
H.eiseni                  ...........................................
H.arcticus                ...........................................
H.californiensis          ...........................................
H.nevadensis              ...........................................
H.victorianensis          ...........................................
H.franciscanus            ............................A..............T..

                          660        670        680        690        700
                           .          .          .          .          .
H.kenai                   AGCCTGCACAGGGGATCTACCTGGAGGCGCTATGCCTACCGGGTTCCAAC
H.hirsutus                ............................T.........T..........A
H.shoshone                ............................G.......T...........
H.eiseni                  ....................TC.A....T.A....T..T..........A
H.arcticus                ......................A......A....T..T..........A
H.californiensis          ....................TC.A....T.A.......T..........A
H.nevadensis              ....................TC........G....T...T.........A
H.victorianensis          ....................T..A....TTA....T..T..A.......A
H.franciscanus            ....................TC......T.A...................

                          710        720        730
                           .          .          .
H.kenai                   CA-TTTATCTTTCGACCTGACATCAAGCAAGACTAC
H.hirsutus                .T-............................
H.shoshone                ..-..A.........................
H.eiseni                  .T-..A.........................
H.arcticus                .T-..A.........................
H.californiensis          .T-..A.........................
H.nevadensis              .T-..A.........................
H.victorianensis          ..-..A.........................
H.franciscanus            ..A............................
```

Appendix 2

Alignment of mtDNA cytochrome oxidase I sequences of 48 haplotypes of *Hesperodiaptomus shoshone* used in Chapter 2 of the thesis.

```
                10        20        30        40        50        60
                .         .         .         .         .         .
HAL2      -GGATGGACCAAATTTATAACGTGGTTGTCACAGCTCATGCTTTCATTATAATTTTTTTC
HAL3      -.....................................G..................
HCA4      -.....................................C.....A..T..............
HCA1      G.....................................C.....A..T..............
HCA2      --....................................C.....A..T..............
HCA6      --....................................C.....A..T..............
HCA7      -.....................................C.....A..T..............
HCA8      G.....................................C.....A..T..............
HCA9      -.....................................C.....A..T..............
HCA11     G.....................................C.....A..T..............
HCA13     -.....................................C.....A..T..............
HCA14     -.....................................C.....A..T..............
HCA18     -.....................................C.....A..T..............
HCO1      G.....................................G..................
HCO3      G.....................................G..................
HCO4      G.....................................G..................
HCO5      G.........................T.....C.....G..................
HCO7      G.....................................G..................
HCO8      G.....................................G..................
HCO11     G.....................................C.....G..................
HCO12     G.....................................C.....G..................
HCO13     G.....................................G..................
HCO14     G.....................................G..................
HCO15     G.....................................C.....G..................
HCO16     G.....................................G..................
HCO17     G.....................................C.....G..................
HCO18     G.....................................C.....G..................
HCO20     G.....................................G..................
HPF3      -.....................................C.....A..T..............
HID1      -.....................................G..................
HID2      -.....................................G..................
HID3      -.....................................G..................
HID8      -.....................................G..................
HID9      -.....................................G..................
HID10     -.....................................G..................
HWY1      -.....................................G.....................
HWY2      -.....................................G.....................
HWY3      ----..................................G..A..................
```

Appendix 2 cont.

```
                  70         80         90         100        110        120
                   .          .          .          .          .          .
HAL2      ATGGTAATACCTATTCTAATCGGCGGATTTGGTAATTGACTAGTGCCTCTTATGCTAGGA
HAL3      ....................................T.....................
HCA4      ..........A...T....T..G.............................AT....G
HCA1      ..........A...T....T..G..G..........................AT....G
HCA2      ..........A...T....T..G..G..........................AT....G
HCA6      ..........A...T....T..G.............................AT....G
HCA7      ..........A...T....T..G.............T.............AT....G
HCA8      ..........A...T....T..G.............................AT....G
HCA9      ..........A...T....T..G.............T.............AT....G
HCA11     ..........A...T....T..G.............T.............AT....G
HCA13     ..........A...T....T..G.............................AT....G
HCA14     ..........A...T....T..G.............................AT....G
HCA18     ..........A...T....T..G.............T.............AT....G
HCO1      ..A.....G..A...T....T..G..G........C...T....A..C......T.....
HCO3      ..A.....G..A...T....T..A..G........C...T....A..C......T.....
HCO4      ..A.....G..A...T....T..A..G........C...T....A..C......T.....
HCO5      ..A.A......A...T....T..G..G............T.G..A.........T....G
HCO7      ..A.....G..A..CT....T..A..............T....A..C......T.....
HCO8      ..A.....G..A..CT....T..A..............T....A..C......T.....
HCO11     ..A.....G..A...T....T..G..G...............A..C......T.....
HCO12     ..A.....G..A...T....T..A..G...............A..C......T.....
HCO13     ..A.....G..A...T....T..A..............T....A..C......T.....
HCO14     ..A.....G..A...T....T..A..............T....A..C......T.....
HCO15     ..A.....G..A...T....T..G...................A..C.....AT.....
HCO16     ..A.....G..A...T....T..A..............T....A..C......T.....
HCO17     ..A........A...T....T..G..G...........T....A..C......T....G
HCO18     ..A........A...T....T..A..G...........T....A.........T....G
HCO20     ..A.....G..A...T....T..A..............T....A..C......T.....
HPF3      ..........A...T....T..G.............................AT....G
HID1      ....................................T...................G
HID2      ....................................T...................G
HID3      ............C.......................T...................G
HID8      ....................................T...................G
HID9      ............C.......................T...................G
HID10     ............C.......................T...................G
HWY1      ........G...............................T.......A.........
HWY2      ........G...............................T.......A..........G
HWY3      ....................................T.......A..........G
```

Appendix 2 cont.

```
                 130        140        150        160        170        180
                  .          .          .          .          .          .
HAL2    GCAAGAGATATGGCTTTCCCCCGGATAAACAATATAAGGTTTTGATTCTTAATTCCTGCT
HAL3    ............................................................
HCA4    ..............C.....G..T..G..T.....G..A........TC....C..G..C
HCA1    ..............C.....G..T..G..T.....G..A........TC....C..G..C
HCA2    ..............C.....G..T..G..T.....G..A........TC....C..G..C
HCA6    ..............C.....G..T..G..T.....G..A........TC....C..G..C
HCA7    ..............C.....G..T..G..T..C.....A........TC....C..G..C
HCA8    ..............C.....G..T..G..T..C.....A........TC....C..G..C
HCA9    ..............C.....G..T..G..T..C.....A........TC....C..G..C
HCA11   ..............C.....G..T..G..T..C.....A........TC....C..G..C
HCA13   ..............C.....G..T..G..T..C..G..A........TC....C..G..C
HCA14   ..............C.....G..T..G..T..C..G..A........TC....C..G..C
HCA18   ..............C.....G..T..G..T..C..G..A........TC....C..G..C
HCO1    .....C.....A..A.....A..T..G..T..................TC.......G..A
HCO3    .....C.....A..A.....A..T..G..T..................TC.......G..C
HCO4    .....C.....A..A.....A..T..G..T..................TC.......G..A
HCO5    ........C.....A.....A..T.....T...........C..G..T........G..C
HCO7    ...........A..C.....A..T.....T..................T........G..C
HCO8    ...........A..C.....A..T.....T..................T........G..C
HCO11   ..T........A..C.....A..T..G..T........A........TC.......G..C
HCO12   ...........A..C.....A..T..G..T..................TC.......G..C
HCO13   ........C..A..C.....A..T..G..T..................TC.......G..C
HCO14   ........C..A..C.....A..T..G..T..................TC.......G..C
HCO15   ...........A..C.....A..T..G..T..................TC.......G..C
HCO16   ...........A..C.....A..T..G..T..................TC.......G..C
HCO17   ..............C.....A..T.....T..............G..T.....C..G..C
HCO18   ........C.....C.....A..T.....T...........C..G..T........G..C
HCO20   ........C..A..C.....A..T..G..T..................TC.......G..C
HPF3    ..............C.....G..T..G..T.....G..A........TC....C..G..C
HID1    ......................A.................................C......
HID2    ......................A.................................C......
HID3    ......................A.................................C......
HID8    ......................A.................................C......
HID9    ..........................................................C......
HID10   ......................A.................................C......
HWY1    ..........................................................C......
HWY2    ..........................................................C......
HWY3    ......................A.................................C......
```

Appendix 2 cont.

```
                 190        200        210        220        230        240
                  .          .          .          .          .          .
HAL2      TTAGTAATACTGTTATCTAGCTCCCTGGTAGAAAGCGGAGCCGGAACAGGGTGAACAGTA
HAL3      ............................................................
HCA4      ........GT.AC....C.....G.....G..G..A..G..T..G............T
HCA1      ........GT.AC....C.....G.....G..G..A..G..T..G............T
HCA2      ........GT.AC....C.....G.....G..G..A..G..T..G............T
HCA6      ........GT.AC....C.....G.....G..G..A..G..T..G............T
HCA7      ........GT.AC....C.....G.....G..G..A..G..T..G............T
HCA8      ........GT.AC....C.....G.....G..G..A..G..T..G............T
HCA9      ........GT.AC....C.....G.....G..G..A..G..T..G............T
HCA11     ........GT.AC....C.....G.....G..G..A..G..T..G............T
HCA13     ........GT.AC....C.....G.....G..G..A..G..T..G............T
HCA14     ........GT.AC....C.....G.....G..G..A..G..T..G............T
HCA18     ........GT.AC....C.....G.....G..G..A..G..T..G............T
HCO1      .........T.AC.G..C..T..G.....G..G..G.....T..................
HCO3      .........T.AC.G..C..T..G.....G..G..G.....T..................
HCO4      .........T.AC.G..C..T..G.....G..G..G.....T..................
HCO5      ..G..G...........C..T..G.....G..G..G..G..G..G...........G..G
HCO7      ...........AC....C..T..G.....G..G..G.....T..G.....A..G.....G
HCO8      ...........AC.G..C..T..G.....G..G..G.....T..G........G.....G
HCO11     ...........AC.G..C..T..G.....G.....G.....T..G...............
HCO12     .....G.....AC....C..T..G.....G..G..G.....T..G...............
HCO13     ........G..AC.G..C..T..G.....G..G..G.....T..G...............
HCO14     ...........AC.G..C..T..G.....G..G..G.....T..G...............
HCO15     ...........AC....C..T..G.....G..G..G.....T..G...............
HCO16     ...........AC.G..C..T..G.....G..G..G.....G..G...............
HCO17     ..G..G.....A.....C..T..G.....G..G..G.....G..G...........G...
HCO18     ..G..G...........C..T..G.....G..G..G..G..G..G...........G...
HCO20     ...........AC.G..C..T..G.....G..G..G.....T..G........G......
HPF3      ........GT.AC....C.....G.....G..G..A..G..T..G............T
HID1      ..................................T.............A........C
HID2      ...........A......................T.....................C
HID3      ..................................T.............A........C
HID8      ...........A......................T.....................C
HID9      ..................................T.............-........C
HID10     ..................................T.............A........C
HWY1      ..........................................T..............C
HWY2      ..........................................T..............C
HWY3      ..................................T.....T..............C
```

Appendix 2 cont.

```
                  250        260        270        280        290        300
                   .          .          .          .          .          .
HAL2       TATCCACCTTTATCTAGTAACATCGCGCACGCTGGCAGTTCAGTAGACTTTGCAATTTTT
HAL3       ..........................................................
HCA4       ..C..T...C....C..G..T.....C.....G..G..A..C..G..............
HCA1       ..C..T...C....C..G..T.....C.....G..G..A..C..G..............
HCA2       ..C..T...C....C..G..T.....C.....G..G..A..C..G..............
HCA6       ..C..C...C....C..G..T.....C.....G..G..A..C..G..............
HCA7       ..C..C..CC....C..G..T.....C.....G..G..A..C..G..............
HCA8       ..C..T...C....C..G..T.....C.....G..G..A..C..G..............
HCA9       ..C..T..CC....C..G..T.....C.....G..G..A..C..G..............
HCA11      ..C..T...C....C..G..T.....C.....G..G..A..C..G..............
HCA13      ..C..T...C....C..G..T.....C.....G..G..A..C..G..............
HCA14      ..C..T........C..G..T.....C.....G..G..A..C..G..............
HCA18      ..C..T...C....C..G..T.....C.....G..G..A..C..G..............
HCO1       ..C..C..CC.......C..T.....T.....G..G..A.....G..............
HCO3       ..C..C..CC.......C..T.....T.....G..G..A.....G..............
HCO4       ..C..C..CC.......C..T.....T.....G..G..A.....G..............
HCO5       ..C..T..CC.......G..T.....C.....G..G..A..C..G..............
HCO7       ..C..C..CC.......C..T.....C.....G..G..A.....G..............
HCO8       ..C..C..CC.......C..T.....C.....G..G..A.....G..............
HCO11      ..C..C..CC.......A..T.....C.....G..G..A.....G..............
HCO12      ..C..C..CC..........T.....C.....G..G..A.....G..............
HCO13      ..C..C..CC.......C..T.....C.....G..G..A.....G..............
HCO14      ..C..C..CC.......C..T.....C.....G..G..A.....G..............
HCO15      ..C..C..CC..........TG....C.....G..G..A.....G..............
HCO16      ..C..C..AC.......C..T.....C.....G..G..A.....G..............
HCO17      ..C..C..CC.......C..T.....C.....A..G..A..C..G..............
HCO18      ..C..C..CC.......G..T.....C.....G..A..A..C..G..............
HCO20      ..C..C..CC.......C..T.....C.....G..G..A.....G..............
HPF3       ..C..T..CC....C..G..T.....C.....G..G..A..C..G..............
HID1       ..........................................................
HID2       ..........................................................
HID3       ..........................................................
HID8       ..........................................................
HID9       ..........................................................
HID10      .....C....................................................
HWY1       ..............C...........................................
HWY2       ..............C...........................................
HWY3       ..............C...........................................
```

Appendix 2 cont.

```
                  310        320        330        340        350        360
                   .          .          .          .          .          .
HAL2      TCTTTGCACCTGGCGGGGGTAAGCTCTATCTTAGGCGCTGTTAATTTTATCAGAACCCTA
HAL3      ............................................................
HCA4      ..C..A..TT............G.....T.....G.....A........T..G..T..T
HCA1      ..C..A...T............G.....T.....G.....A........T..G..T..T
HCA2      ..C..A..TT............G.....T.....G.....A........T..G..T..T
HCA6      ..C..A..TT............G.....T.....G.....A........T..G..T..T
HCA7      ..C..A...T............G.....T.....G.....A........T..G..T..T
HCA8      ..C..A...T............G.....T.....G.....A........T..G..T..T
HCA9      ..C..A...T............G.....T.....G.....A........T..G..T..T
HCA11     ..C..A...T............G.....T.....G.....A........T..G..T..T
HCA13     ..C..A...T............G.....T.....G.....A........T..G..T..T
HCA14     ..C..A...T............G.....T.....G.....A........T..G..T..T
HCA18     ..C..A...T............G.....T.....G.....A........T..G..T..T
HCO1      .....A..T........C.....G.....T..G..G.....A........T..G..T..T
HCO3      ...C.A..T........C.....G.....T..G..G.....A........T..G..T..T
HCO4      .....A..T........C.....G.....T..G..G.....A........T..G..T..T
HCO5      .....A..TT..........G..G.....T.....A.....A........T.....T..T
HCO7      .....A..TT.......T..C..G.....T..G..G.....A........T..G..T..T
HCO8      .....A..TT.......T..C..G.....T..G..G.....A........T..G..T..T
HCO11     ..C..A...T..........G..G.....T.....G.....A........T..G..T..T
HCO12     ..C..A...T.......T..G..G.....T..G..G.....A........T..G..T..T
HCO13     .....A..T...........G..G.....T.....G.....A........T..G..T..T
HCO14     .....A..TT..........G..G.....T.....G.....A........T..G..T..T
HCO15     ..C..A............T..G..G.....T.....G..C..A............G..T..T
HCO16     .....A..T...........G..G.....T.....G.....A........T.....T..T
HCO17     .....A...T..........G..A.....T.....A.....A........T.....T..T
HCO18     .....A..TT..........G..G.....T.....A.....A........T.....T..T
HCO20     .....A..T...........G..G.....T.....G.....A........T..G..T..T
HPF3      ..C..A...T............G.....T.....G.....A........T..G..T..T
HID1      ........T.....................................................
HID2      ........T.....................................................
HID3      ........T.....................................................
HID8      ........T.....................................................
HID9      ........T....................T.................................
HID10     ........T.....................................................
HWY1      ........T...........G........T.................................
HWY2      ........T....................T.................................
HWY3      ........T....................T.................................
```

Appendix 2 cont.

```
                  370        380        390        400        410        420
                   .          .          .          .          .          .
HAL2       GGGAATTTACGAGCATTTGGGATAATCCTGGACCGGATACCGTTATTCGCTTGAGCCGTT
HAL3       ............................................................
HCA4       ......C.G..G........A.....T..A.....T.....T..G..T........T...
HCA1       ......C.G..G........A.....T..A.....T.....T..G..T........T...
HCA2       ......C.G..G........A.....T..A.....T.....T..G..T........T...
HCA6       ......C.G..G........A.....T..A.....T.....T..G..T........T...
HCA7       ......C.G..G........A.....T..A.....T.....T..G..T........T...
HCA8       ......C.G..G........A.....T..A.....T.....T..G..T........T...
HCA9       ......C.G..G........A.....T..A.....T.....T..G..T........T...
HCA11      ......C.G..G........A.....T..A.....T.....T..G..T........T...
HCA13      ......C.G..G........A.....T..A.....T.....T..G..T........T...
HCA14      ......C.G..G........A.....T..A.....T.....T..G..T........T...
HCA18      ......C.G..G........A.....T..A.....T.....T..G..T........T...
HCO1       ......C.G..G........A..G.....A.....C..G..T.....T........T..G
HCO3       ......C.G..G........A..G.....A.....C..G..T.....T........T..G
HCO4       ......C.G..G........A..G.....A.....C..G..T.....T........T..G
HCO5       ..C...C.G..G........T........A.....A..G........T........T..A
HCO7       ......C.G.....................A.....T..G..T.....T........T..G
HCO8       ......C.G..G..................A.....T..G..T.....T........T..G
HCO11      ......C.G..G........A........A.....T..G..T.....T........T..G
HCO12      ......C.G..G........A........A.....T..G..T.....T........T..G
HCO13      ......C.G..G........A........A.....T..G..T.....T........T..G
HCO14      ......C.G..G........A........A.....T..G..T.....T........T..G
HCO15      ......C.G..G........A........A.....T..G..T.....T........T..G
HCO16      ......C....G........A........A........G..T.....T........T..G
HCO17      ..C...C.G.....................A.....T..G........T........T...
HCO18      ..C...C.G..G........T........A.....A..G..A.....T........T...
HCO20      ......C.G..G.....C..A........A.....T..G..T.....T........T..G
HPF3       ......C.G..G........A.....T..A.....T.....T..G..T........T...
HID1       ..........................................A..................
HID2       ..........................................A..................
HID3       ..........................................A..................
HID8       ..........................................A..................
HID9       ..........................................A........C.........
HID10      ..........................................A..................
HWY1       ............................................................
HWY2       ............................................................
HWY3       ............................................................
```

Appendix 2 cont.

```
                  430        440        450
                   .          .          .
HAL2       TTAATTACTGCAGTATTACTCCTTCTATCTCTGCC
HAL3       ...................................
HCA4       .............T.....T..C...........
HCA1       .............T.....T..C...........
HCA2       .............T.....T..C...........
HCA6       .............T.....T..C...........
HCA7       .............T.....T..C...........
HCA8       .............T.....T..C...........
HCA9       .............T.....T..C...........
HCA11      .............T.....T..C...........
HCA13      .............T.....T..C...........
HCA14      .............T.....T..C...........
HCA18      .............T.....T..C...........
HCO1       .............T.....T...............
HCO3       .............T.....T...............
HCO4       .............T.....T...............
HCO5       .........G.....T.....T.............
HCO7       .............T.....T...............
HCO8       .........C.....T.....T.............
HCO11      .............T.....T...T...........
HCO12      ...........G..T.....T...............
HCO13      .............T.....T...............
HCO14      .............T.....T...............
HCO15      .............T.....T...............
HCO16      .............T.....T...............
HCO17      .............T.....T...T...........
HCO18      .........G.....T.....T.............
HCO20      .............T.....T...............
HPF3       .............T.....T..C...........
HID1       ...................................
HID2       ..................T................
HID3       ...................................
HID8       .........A.........................
HID9       ...................................
HID10      ...................................
HWY1       ........................T..........
HWY2       ........................T..........
HWY3       ........................T..........
```

Appendix 3

Alignment of nuclear rDNA sequences of 48 haplotypes of *Hesperodiaptomus shoshone* and the outgroup *H. kenai* used in Chapter 2 of the thesis.

```
                    10         20         30         40         50
                     .          .          .          .          .
Hap1        CTTAAATTCGGCGGGTAGTCTTGCTTGATGTCAGGTCGAAAGATTAATGG
Hap2        ..................................................
Hap3        ..................................................
Hap4        ..................................................
Hap5        ..................................................
Hap6        ..................................................
Hap7        ..................................................
Hap8        ..................................................
Hap9        ..................................................
Hap10       ..................................................
Hap11       ..................................................
Hap12       ..................................................
Hap13       ..................................................
Hap14       ..................................................
H.kenai     ................TA.........................A.....

                    60         70         80         90         100
                     .          .          .          .          .
Hap1        TTGGAACCCGGTAGGCATCGCGCCTCCAGGTAGATCCCCTGTGCAGGCTT
Hap2        ..................................................
Hap3        ..................................................
Hap4        ..................................................
Hap5        ..................................................
Hap6        ..................................................
Hap7        ..........A.......................................
Hap8        ..........A.......................................
Hap9        ..........A.......................................
Hap10       ..........A.......................................
Hap11       ..........A.......................................
Hap12       ..........A.......................................
Hap13       ..........A.......................................
Hap14       ..........A.......................................
H.kenai     ................A.................................
```

Appendix 3 cont.

```
                    110        120        130        140        150
                     .          .          .          .          .
Hap1        GTCCCACCCAGACCCAGGCTTTTTACGACCCAAGAGTCTTTGGCAGAGCA
Hap2        ..................................................
Hap3        ..................................................
Hap4        ..................................................
Hap5        ..................................................
Hap6        ..................................................
Hap7        ..................................................
Hap8        ..................................................
Hap9        ..................................................
Hap10       ..................................................
Hap11       ..................................................
Hap12       ..................................................
Hap13       ..................................................
Hap14       ..................................................
H.kenai     ..................................................

                    160        170        180        190        200
                     .          .          .          .          .
Hap1        CTGCACAAGTCAACTTGAGGCCACCCGCCGGCCGAAGCCGACAGGACGAC
Hap2        ..................................................
Hap3        ..................................................
Hap4        ..................................................
Hap5        ..................................................
Hap6        ..................................................
Hap7        ..................................................
Hap8        ..................................................
Hap9        ..................................................
Hap10       ..................................................
Hap11       ..................................................
Hap12       ..................................................
Hap13       ..................................................
Hap14       ..................................................
H.kenai     ..................................................

                    210        220        230        240        250
                     .          .          .          .          .
Hap1        CCAGGGCGAATCTCGCAAGAGTTTCGATGGATGTTTCAACCGACCCTGAC
Hap2        ..................................................
Hap3        ..................................................
Hap4        ..................................................
Hap5        ..................................................
Hap6        ..................................................
Hap7        ..........A.......................................
Hap8        .......A..A.......................................
Hap9        .......A..A.......................................
Hap10       .......A..A.......................................
Hap11       .......A..A.......................................
Hap12       .......A..A.......................................
Hap13       .......A..A.......................................
Hap14       .......A..A.......................................
H.kenai     .......A..A............T...........................
```

Appendix 3 cont.

```
                  260        270        280        290        300
                   .          .          .          .          .
Hap1       CCAAACATGGACATGCAACAAGCACATGTCCGCAATATGCGTTCAAAATG
Hap2       ...........................................................
Hap3       ...........................................................
Hap4       ...........................................................
Hap5       ...........................................................
Hap6       ...........................................................
Hap7       ...........................................................
Hap8       ...........................................................
Hap9       ...........................................................
Hap10      ...........................................................
Hap11      ...........................................................
Hap12      ...........................................................
Hap13      ...........................................................
Hap14      ...........................................................
H.kenai    ...........................................................

                  310        320        330        340        350
                   .          .          .          .          .
Hap1       TCGATGTTCATGTGTCCTGCAGTTCACATTGCGTCACGGAGTTAGCTCCG
Hap2       ...........................................................
Hap3       ...........................................................
Hap4       ...........................................................
Hap5       ...........................................................
Hap6       ...........................................................
Hap7       ...........................................................
Hap8       ...........................................................
Hap9       ...........................................................
Hap10      ...........................................................
Hap11      ...........................................................
Hap12      ...........................................................
Hap13      ...........................................................
Hap14      ...........................................................
H.kenai    ...........................................................

                  360        370        380        390        400
                   .          .          .          .          .
Hap1       TTCTTCATCGACCCCCGAGCCAAATGATCCACTGTATAGTGTTGTTTTTG
Hap2       ...........................................................
Hap3       ...........................................................
Hap4       ...........................................................
Hap5       ...........................................................
Hap6       ...........................................................
Hap7       ...........................................................
Hap8       ...........................................................
Hap9       ...........................................................
Hap10      ...........................................................
Hap11      ...........................................................
Hap12      ...........................................................
Hap13      ...........................................................
Hap14      ...........................................................
H.kenai    ...........................................................
```

Appendix 3 cont.

```
                410        420        430        440        450
                 .          .          .          .          .
Hap1    TGGTTTCCAGCGTATAGCCGGTCATTTTCAATAGGAATGTACCGCTTTTC
Hap2    .............................................
Hap3    .............................................T.....
Hap4    .............................................T.....
Hap5    .............................................T.....
Hap6    ........G.........A.........................T.....
Hap7    ........G.....A....A..T......T.A............T.....
Hap8    ........G.........A..........................
Hap9    ........G.........A..........................
Hap10   ........G.........A..........................
Hap11   ........G.........A....................A.....
Hap12   ........G.........A....................A.....
Hap13   ........G.........A.........................T.....
Hap14   ........G.........A.........................T.....
H.kenai ........G.....A....A.........TT............T....

                460        470        480        490        500
                 .          .          .          .          .
Hap1    ACAAGCCCAACATGGGGGGAGCACATAACCCGAGTCGGAGGCGAGGCTCT
Hap2    .................................................
Hap3    .................................................
Hap4    .................................................
Hap5    .................................................
Hap6    .................................................
Hap7    ...........G......................C................
Hap8    ..G........G......................C................
Hap9    ..G........G......................C................
Hap10   ...........G......................C................
Hap11   ..G........G......................C................
Hap12   ..G........G......................C................
Hap13   ..G........G......................C................
Hap14   ..G........G......................C................
H.kenai ........G..........................C................


                510        520        530        540        550
                 .          .          .          .          .
Hap1    TTGACAGCTGGATTACCGCCTTTCACCCCGAGGGGCGACAACGGCGCACA
Hap2    .................................................
Hap3    .........................................T.........
Hap4    .................................................
Hap5    .................................................
Hap6    .................................................
Hap7    .................................................
Hap8    .................................................
Hap9    .................................................
Hap10   .................................................
Hap11   .................................................
Hap12   .................................................
Hap13   .................................................
Hap14   .................................................
H.kenai ..........A..................AT...................
```

Appendix 3 cont.

```
                    560        570        580        590        600
                     .          .          .          .          .
Hap1        GCACCTCCGATCGGGCAGTCAGGCGAGAGCAAGCGGCCAGGTAGCCGCAA
Hap2        ..................................................
Hap3        ..................................................
Hap4        ..................................................
Hap5        ..................................................
Hap6        ..................................................
Hap7        ..................................................
Hap8        ..................................................
Hap9        ..................................................
Hap10       ..................................................
Hap11       ..................................................
Hap12       ..................................................
Hap13       ..................................................
Hap14       ..................................................
H.kenai     ..................................................

                    610        620        630        640        650
                     .          .          .          .          .
Hap1        GCCCCCAATCCCAACTTATCAGACAACGGCAGGCGAGCCGATGTCCCGAG
Hap2        ............G.....................................
Hap3        ............G.....................................
Hap4        ............G.....................................
Hap5        ............G.........................A...........
Hap6        ............G.........................C...........
Hap7        ......G.........................A.....C...........
Hap8        ......G.....................................TC.....
Hap9        ......G.....................................C.....
Hap10       ......G.....................................C.....
Hap11       ......G.....................................C.....
Hap12       ......G.....................................TC.....
Hap13       ......G.....................................C.....
Hap14       ......G.........................A.....C...........
H.kenai     ............G.....................C...........

                    660        670        680        690        700
                     .          .          .          .          .
Hap1        GCAGCAAAGCCCGAGGTTATTACGAAATTCCCTCTGCCTAAGCATCGGTA
Hap2        ..................................................
Hap3        ..................................................
Hap4        ..................................................
Hap5        ..................................................
Hap6        ..................................................
Hap7        .....G............................................
Hap8        .....G............................................
Hap9        .....G............................................
Hap10       .....G............................................
Hap11       .....G............................................
Hap12       .....G............................................
Hap13       .....G............................................
Hap14       .....G............................................
H.kenai     .....G............................................
```

Appendix 3 cont.

```
                    710        720        730
                     .          .          .
Hap1        GTAATCGATAATGATCCTTCCGCAGGTTCACCTACGGAA
Hap2        ........................................
Hap3        ........................................
Hap4        ........................................
Hap5        ........................................
Hap6        ........................................
Hap7        ........................................
Hap8        ........................................
Hap9        ........................................
Hap10       ........................................
Hap11       ........................................
Hap12       ........................................
Hap13       ........................................
Hap14       ........................................
H.kenai     ........................................
```