2023

Speleogenomics of cave life in North America

Joseph B. Benito

SPELEOGENOMICS OF CAVE LIFE IN NORTH AMERICA

by

JOSEPH B. BENITO

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in
The Biotechnology Science and Engineering Program
to
The Graduate School
of
The University of Alabama in Huntsville

May 2023

In presenting this dissertation in partial fulfillment of the requirements for a doctoral degree from The University of Alabama in Huntsville, I agree that the Library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by my advisor or, in his/her absence, by the Chair of the Department or the Dean of the School of Graduate Studies. It is also understood that due recognition shall be given to me and to The University of Alabama in Huntsville in any scholarly use which may be made of any material in this dissertation.

Joseph B. Benito 12/12/2022

(student signature) (date)

DISSERTATION APPROVAL FORM

Submitted by Joseph B. Benito in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biotechnology Science and Engineering and accepted on behalf of the Faculty of the School of Graduate Studies by the dissertation committee.

We, the undersigned members of the Graduate Faculty of The University of Alabama in Huntsville, certify that we have advised and/or supervised the candidate on the work described in this dissertation. We further certify that we have reviewed the dissertation manuscript and approve it in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biotechnology Science and Engineering.

Matthew L. Niemiller 12/20/2022

Committee Chair (Date)

Zachary Culumber

Karen A. Ober

Megan L. Porter

Paul G. Wolf

Matthew L. Niemiller

Department Chair

Rainer Steinwandt Digitally signed by Rainer Steinwandt Date: 2023.01.07 14:29:37 -06'00'

Rainer Steinwandt

College Dean

Jon Hakkila Digitally signed by Jon Hakkila DN: cn=Jon Hakkila, o=UAH, ou=Associate Provost/Graduate Dean, email=jon.hakkila@uah.edu, c=US Date: 2023.04.18 15:40:14 -05'00' Adobe Acrobat version: 2020.005.30467

Graduate Dean

ABSTRACT

The Graduate School

The University of Alabama in Huntsville

Degree Doctor of Philosophy Program Biotechnology Science and Engineering

Name of Candidate Joseph B. Benito

Title Speleogenomics of Cave Life in North America

Cave organisms with their unusual morphologies including loss of eyes, lack of body pigmentation, and the compensatory improvement of nonvisual sensory systems have long intrigued and fascinated biologists. Due to their restricted distributions and life history traits, many troglomorphic (subterranean-adapted) species are considered imperiled and are high priority targets for protective management. With increasing anthropogenic threats to cave ecosystems, it is important to study cave-dwelling organisms, so that informed management decisions can be made on a regular basis. Detailed information on the genetics, taxonomy, distribution, and colonization history of cavernicoles are necessary to make rigorous ecological inference and develop respective recommendations for monitoring and protecting cave organisms. The inference of phylogenetic relationships among subterranean fauna can be challenging because of morphological homoplasy due to certain requirements of cave life. To overcome this, molecular data are needed to test morphology-based hypotheses regarding the systematic and biogeographic relationships of terrestrial and aquatic subterranean life. While recent genetic and phylogeographic analyses have greatly improved our understanding of evolutionary and biogeographic history of cave organisms, many questions either remain unanswered or poorly

investigated. In this dissertation I examined the current state of knowledge on cave ecology and molecular evolution and also discuss the advantages and possibilities that biospeleological investigations at the genomic level or "speleogenomics" provide to the understanding of these fascinating systems – with special emphasis in the areas of systematics, selection pressures, mt genome evolution, and phylogeography.

In particular, I investigated several evolutionary and biogeographic questions in two model organisms, the eastern North American crangonyctid amphipods and cave trechine beetles. I described the complete mitogenomes of four species of groundwater amphipods as well as a surface spring-dwelling species belonging to the family Crangonyctidae. I compared the base composition, codon usage, and gene order rearrangement, conducted comparative mitogenomic and phylogenomic analyses, examined evolutionary signals imprinted on mitogenome of surface-adapted amphipods and compared to their subterranean counterparts to show evidence of adaptive evolution. In addition, I elucidated the colonization history, biogeography, and systematics of cave trechine beetles distributed primarily in the Appalachians (APP), Interior Low Plateau (ILP), and Ozarks (OZK) karst regions of central and eastern North America using UCE phylogenomics to estimate divergence times and ancestral range distribution.

Abstract Approval:

Committee Chair Matthew L. Niemiller

Department Chair Matthew L. Niemiller

Graduate Dean Jon Hakkila

ACKNOWLEDGEMENTS

I would first like to thank Dr. Matthew Niemiller, my dissertation advisor. I was and remain very humbled by his deep knowledge in the field of cave biology. I would not know half of what I know about cave biology had it not been for his influence. I selected Dr. Paul Wolf, Dr. Zachary Culumber, Dr. Megan Porter, and Dr. Karen Ober to round out my committee, and I owe them a great deal of gratitude, as well. I thank them for their words of encouragement and commitment to me. I thank my former program coordinator, Dr. Joseph Ng for his advice during the initial days of the PhD program. I thank Dr. David Young from the Alabama Supercomputer Center for providing assistance with troubleshooting and program installation on the high-performance computing cluster. I thank my lab alumni members Katherine Dooley and Kayla Wilson for their assistance in collecting amphipod specimens from local caves and creating a wonderful lab environment.

I am grateful for the help and support I received from Dr. Denise Niemiller during our caving trips and being a mentor in our teaching labs. I thank my past principal investigators and employers, Dr. Shawn Levy and Dr. David Willoughby for their guidance during my previous employment. I am grateful to my Master's advisor, Dr. Kalle Levon for his guidance and support to pursue higher education in the US. I thank my family for their overwhelming love, support, and prayers. I especially want to thank Dad, Kaiser, Mom, Pushparani, Brother, Sartho, and Sister, Hailma. Above any other person, I owe so much to my loving wife, Priyasha and daughter, Shauna. They deserves all the credit for my happiness and I love them more than anything. Finally, and above all, I thank the Lord for everything. I have been blessed with so much by so many, and I am forever grateful.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

GENERAL INTRODUCTION

Cave organisms with their unusual morphologies including loss of eyes, lack of body pigmentation, and the compensatory improvement of nonvisual sensory systems (Culver *et al*. 1995; Culver and Pipan 2009) have fascinated biologists since the description of olm (*Proteus anguinus*), the first cave vertebrate in 1768 (Laurenti 1768). Due to their restricted distributions and life history traits, many troglomorphic (subterranean-adapted) species are considered imperiled and are high priority targets for protective management (Culver *et al*. 2000; Niemiller and Zigler 2013; Niemiller *et al*. 2017). Troglomorphic species are mostly endemic to a single cave or habitat (Reddell 1994; Culver *et al*. 2000; Christman *et al*. 2005; Gao *et al*. 2018) and are often characterized by small populations (Mitchell 1970). The fauna of most of the world's caves remain unknown or incompletely surveyed (Howarth 1983; Whitten 2009; Gibert and Deharveng 2002; Deharveng 2000; Encinares and Lit 2014; Gilgado *et al*. 2015).

With increasing anthropogenic threats to cave ecosystems, it is important to study cave-dwelling organisms, so that informed management decisions can be made on a regular basis. Cave ecosystems encounter various human impacts all over the world including land cover modification (Culver 1986; Trajano 2000; Howarth *et al*. 2007; Silva *et al*. 2015), mining (Elliott 2000; Silva *et al*. 2015; Sugai *et al*. 2015), groundwater pollution (Aley 1976; Notenboom *et al*. 1994; Graening and Brown 2003; Whitten 2009), water impoundments (Lisowski 1983; Ubick and Briggs 2002; Olson 2005), invasive species (Elliott 1992; Reeves 1999; Taylor *et al*. 2003; Howarth *et al*. 2007; Wynne *et al*. 2014), climate change (Chevaldonné and Lejeune 2003; Badino 2004; Mammola *et al*. 2018), and

recreational use (Culver 1986; Howarth and Stone 1993; Pulido-Bosch *et al*. 1997). These threats have significant implications for conservation of caves which are highly sensitive habitats serving as hotspots of subterranean biodiversity and endemism (Culver *et al*. 2000; Culver and Sket 2002; Eberhard *et al*. 2005). Thus, detailed information on the genetics, taxonomy, distribution, and colonization history of cavernicoles are necessary to make rigorous ecological inference and develop respective recommendations for monitoring and protecting cave organisms (Wynne *et al*. 2018).

While recent genetic and phylogeographic analyses have greatly improved our understanding of evolutionary and biogeographic history of cave organisms (reviewed in Juan *et al*. 2010), many questions either remain unanswered or poorly investigated. The inference of phylogenetic relationships among subterranean fauna can be challenging because of morphological homoplasy due to certain requirements of cave life (Cooper *et al*. 2002; Proudlove and Wood 2003; Lefébure *et al*. 2006). To overcome this, molecular data are needed to test morphology-based hypotheses regarding the systematic and biogeographic relationships of terrestrial and aquatic subterranean life (Loria *et al*. 2011; Hedin *et al*. 2018; Katz *et al*. 2018; Leray *et al*. 2019; Hart *et al*. 2020; Derkarabetian *et al*. 2022; Dooley *et al*. 2022; Grant *et al*. 2022). Molecular data aid morphological study and vice versa (Dayrat 2005; Page *et al*. 2005b), because a complete view of evolutionary history can only be attained by accessing the strengths of both morphological characters and molecular data (Hillis and Wiens 2000; Lee 2004).

In parallel to the modern sequencing technologies and developments, there has been a rapid increase in mitochondrial (mt) genome data availability for many animal groups, but particularly crustaceans. Indeed, such expansion of available genetic resources for a

particular species is the most commonly invoked reason for undertaking mt genome sequencing. Whole mt genomes have been used for the same wide array of research goals as individual mt genes including molecular systematics (at both deep and shallow taxonomic scales), population genetics/phylogeography (Ma *et al*. 2012), diagnostics (Nelson *et al*. 2012), and molecular evolutionary studies (Castro *et al*. 2002; Salvato *et al*. 2008; Shao *et al*. 2003). In addition, whole mt genome sequencing also allows the study of comparative and evolutionary genomics questions such as the frequency and type of gene rearrangements (Cameron *et al*. 2011; Dowton *et al*. 2009), evolution of genome size etc (Shao *et al*. 2009). The small size of the mt genome makes it a practical genome study system in crustaceans which nuclear genome sequencing will not equal in the near future. Similarly, a recently developed class of markers surrounding ultraconserved DNA elements (UCEs; Faircloth *et al*. 2012) can be used in conjunction with sequence capture and massively parallel sequencing to generate large amounts of orthologous sequence data among a taxonomically diverse set of species. UCEs are numerous in a diversity of metazoan taxa (Ryu *et al*. 2012), and over 5000 have been identified in amniotes (Stephen *et al*. 2008; Faircloth *et al*. 2012). Although UCEs are highly conserved across distantly related taxa, their flanking regions harbor variation that increases with distance from the conserved core (Faircloth *et al*. 2012). The conserved region allows easy alignment across widely divergent taxa, while variation in the flanking region is useful for comparative analyses. Faircloth *et al*. (2012) suggested that because variation within UCE flanking regions is abundant, UCEs are likely be useful for investigations at shallow evolutionary timescales.

With third-generation sequencing technology rapidly approaching, it is more feasible to obtain whole mt genome and large multilocus data sets to infer evolutionary relationships. These enormous quantities of data have ignited the development of several new programs for phylogenetic inference for these highly heterogeneous data sets. From multiple sequence alignment to species tree construction, these new methods are changing the way we gather, manipulate, analyze data, and interpret results. In this dissertation I examine the current state of knowledge on cave ecology and molecular evolution and also discuss the advantages and possibilities that biospeleological investigations at the genomic level or "speleogenomics" provide to the understanding of these fascinating systems – with special emphasis in the areas of systematics, selection pressures, mt genome evolution, and phylogeography.

In this dissertation, I investigate several evolutionary and biogeographic questions in two model organisms, the North American crangonyctid amphipods and cave trechine beetles. Amphipods (Class Malacostraca: Order Amphipoda) are one of the most ecologically diverse crustacean groups including over 10,000 species (Arfianti *et al*. 2018; Horton *et al*. 2020), occurring in a diverse array of aquatic and even terrestrial environments globally, from aphotic groundwater aquifers and hadal depths to freshwater streams and lakes in temperate and tropical forests, among other habitats (Bousfield 1983; Barnard and Karaman 1991). We compared the mitogenomes of surface and subterranean amphipods, including the 13 mitochondrial PCGs involved in the OXPHOS pathway to understand the potential molecular mechanisms of energy metabolism in this diverse crustacean group. Cave trechines are prominent in many terrestrial subterranean habitats in Asia, Europe, and North America, and they are an ideal model system to study the

evolution and biogeography of subterranean life (Faille *et al*. 2010, 2014; Ribera *et al*. 2010; Rizzo *et al*. 2013; Chen *et al*. 2021). They are small, predatory ground beetles (3–8 mm long) almost all of which lack eyes, are flightless, and are depigmented with long, slender bodies, elongated appendages, and sensory setae ('aphaenopsian'; Barr 2004; Ober *et al*. 2022). North American cave trechine beetles include 162 taxa in six genera distributed primarily in the Appalachians (APP), Interior Low Plateau (ILP), and Ozarks (OZK) karst regions of central and eastern North America (Barr 2004; Ober *et al*. 2022). We generated the first molecular phylogenetic framework for the study of the origin, diversification, and biogeography of cave trechines in eastern North America using UCE phylogenomics.

In Chapter I, I describe the complete mitogenomes of four species of groundwater amphipods, including *Stygobromus pizzinii*, *Stygobromus allegheniensis*, *Stygobromus tenuis potomacus*, and *Bactrurus brachycaudus*, as well as a surface spring-dwelling species, *Crangonyx forbesi* belonging to the family Crangonyctidae. In Chapter II, I compare the base composition, codon usage, gene order rearrangement, conduct comparative mitogenomic and phylogenomic analyses, and examine evolutionary signals imprinted on mitogenome of surface-adapted amphipods and compare to their subterranean counterparts to show evidence of adaptive evolution. Lastly, I elucidate the colonization history, biogeography, and systematics of cave trechine beetles using UCE phylogenomics to estimate divergence times and ancestral range distribution in Chapter III. The three chapters are briefly summarized below.

CHAPTER 1. THE MITOCHONDRIAL GENOMES OF FIVE SPRING AND GROUNDWATER AMPHIPODS OF THE FAMILY CRANGONYCTIDAE (CRUSTACEA: AMPHIPODA) FROM EASTERN NORTH AMERICA

The following chapter is a slightly modified version of the publication:
Benito, J. B., Porter, M. L., & Niemiller, M. L. (2021). The mitochondrial genomes of five spring and groundwater amphipods of the family Crangonyctidae (Crustacea: Amphipoda) from eastern North America. *Mitochondrial DNA Part B*, *6*(6), 1662-1667.

1.1 Introduction

Amphipods of the family Crangonyctidae are a diverse group of crustaceans, many of which are associated with groundwater-related habitats, such as cave streams, small springs, seeps, hyporheic zones, and wells. This family is particularly diverse in groundwater of the eastern United States with over 100 described species in three genera: *Bactrurus*, *Crangonyx*, and *Stygobromus* (Holsinger 1977, 1978; Zhang 1997; Konemann and Holsinger 2002). Much of this diversity is troglomorphic, lacking eyes, pigment, and frequently with attenuated bodies (Holsinger 1977, 1978). The mitogenomes of several amphipods have been sequenced over the last decade (Krebes and Bastrop 2012; Pons *et al*. 2014; Stokkan *et al*. 2015; Aunins *et al*. 2016; Romanova *et al*. 2016). For the family Crangonyctidae, however, only three mitogenomes are available and all are for species in the genus *Stygobromus* (Aunins *et al*. 2016).

To provide better representation for the family Crangonyctidae and provide mitogenomic resources for future phylogenetic studies, we describe herein the complete

mitogenomes of four species of groundwater amphipods, including *Stygobromus pizzinii* (Shoemaker, 1938), *Stygobromus allegheniensis* (Holsinger, 1967), *Stygobromus tenuis potomacus* (Holsinger, 1967), and *Bactrurus brachycaudus* (Hubricht & Mackin, 1940), as well as a surface spring-dwelling species, *Crangonyx forbesi* (Hubricht & Mackin, 1940). Mitogenomic data presented are the first available for the genera *Bactrurus* or *Crangonyx*. Aunins *et al*. (2016) described the mitogenome of *S. tenuis potomacus* previously, but we present the mitogenome sampled from a different population offering the first intraspecific comparison of amphipod mitogenomes. We also provide a comparative analysis of structure and gene order of other amphipod mitogenomes available.

1.2 Materials and Methods

Specimens of *Stygobromus pizzinii* and *S. tenuis potomacus* were collected from Pimmit Run Seepage Spring (38.929432 °N; -77.118613 °W), Arlington County, Virginia, in May 2015. Specimens of *Bactrurus brachycaudus* were collected from Fogelpole Cave (38.198055 °N; -90.129722 °W), Monroe County, Illinois, in October 2014. Specimens of *Stygobromus allegheniensis* were collected from Caskey Spring (35.50319 °N; -77.85139 °W), Berkeley County, West Virginia, in September 2013. Specimens of *Crangonyx forbesi* were collected from a unidentified spring in Monroe County, Illinois, in 2013. All specimens were preserved in 100% ethanol in the field. Specimens and DNA extracts are maintained in the CaveBio collection at The University of Alabama in Huntsville. Whole genomic DNA was isolated using the Qiagen DNA Easy Blood and Tissue kit and libraries were prepared using the Illumina TruSeq DNA Library Prep Kit (Illumina Inc., California).

Libraries were then paired-end sequenced (2 × 150bp) on an Illumina HiSeq 4000 platform at the Vincent J. Coates Genomics Sequencing Laboratory at the University of California, Berkeley (supported by NIH S10 OD018174 Instrumentation Grant).

We assessed the quality of the raw reads using FastQC v0.11.5 (Andrews 2010), and the reads were trimmed and filtered using Trimmomatic v0.33 (Bolger *et al*. 2014). De-novo assembly was carried out using NOVOPlasty v2.6.3 assembler (Dierckxsens *et al*. 2017). We then annotated the protein-coding genes, transfer RNAs (tRNAs), and ribosomal RNAs (rRNAs) for each of the five mitogenomes using the mitochondrial genome annotation web server MITOS (Bernt *et al*. 2013). The annotation of tRNAs were further confirmed using the Mitochondrial tRNA finder MiTFi with Amphipoda models to search specific mitochondrial tRNA (Juhling *et al*. 2012, Romanova *et al*. 2020). The locations of start and stop codons of protein coding genes were further confirmed using NCBI ORFfinder (Wheeler DL *et al*. 2003) and by visual comparison to other published amphipod mitogenomes. The location of the control region was confirmed by the presence of a large intergenic spacer region with a string of thymines found immediately after *rrnS* and before *trnI*. However, the control region was located in a different region within the *C. forbesi* mitogenome. The secondary structures of tRNAs were inferred using MITFI (Juhling *et al*. 2012), a built-in module in MITOS. We downloaded from GenBank the annotated mitogenomes of 18 related amphipods that occupy the groundwater and spring habitats and three isopods as outgroup for comparative analyses (Table 1). The amino acid sequences of 13 protein-coding genes of the five new mitogenomes, 18 previously published amphipod mitogenomes, and three isopod mitogenomes were aligned using MAFFT version 7 (Katoh *et al*. 2013). Poorly aligned regions were eliminated using

Gblocks version 0.91b (Castresana 2000). The best partitioning strategy and best-fit evolutionary models for each partition were inferred using PartitionFinder version 2.1.1 (Lanfear *et al*. 2012). Phylogenetic relationships of the 23 amphipod mitogenomes and three isopod mitogenomes using the concatenated 13 protein-coding gene alignment were determined using Bayesian inference in MrBayes v3.2 (Ronquist *et al*. 2012). All analyses were conducted using the PhyloSuite v1.1.15 (Zhang *et al*. 2019).

1.3 Results and Discussion

We assembled and annotated the complete mitogenomes of *Stygobromus pizzinii* (15,176 bp, GenBank accession no. MN175620), *Stygobromus tenuis potomacus* (14,712 bp, GenBank accession no. MN175621), *Stygobromus allegheniensis* (15,164 bp, GenBank accession no. MN175622), *Bactrurus brachycaudus* (14,661 bp, GenBank accession no. MN175619), and *Crangonyx forbesi* (15,469 bp, GenBank accession no. MN175623). All mitogenomes contained 13 protein-coding genes, 22 tRNA genes, one small ribosomal RNA (*rrnS*) gene, and one large ribosomal RNA (*rrnL*) gene, and a non-coding control region representative of the Kingdom Animalia. In all mitogenomes, the heavy strand encoded a total of 23 genes, whereas the light strand encoded 14 genes. Similar to the ancestral gene order of pancrustaceans (crustaceans and hexapods; Kilpert and Podsiadlowski 2006; Yang and Yang 2008), the light strand encoded the same four protein-coding genes (*nad1*, *nad4*, *nad4l*, and *nad5*) in all five mitogenomes. AT content of the *Stygobromus* and *Crangonyx* mitogenomes ranged 67.2–69.1%, consistent with the mitogenomes of other amphipods (Aunins *et al*. 2016; Romanova *et al*. 2016). However, AT content of the *B. brachycaudus* mitogenome was slightly lower at 63.9%.

Both *C. forbesi* and *S. allegheniensis* had more intergenic spacers (18 and 16, respectively) than *B. brachycaudus*, *S. pizzinii*, or *S. tenuis potomacus*, which all possessed just eight intergenic spacers. These intergenic spacers were found primarily between protein-coding genes, but at times also between tRNA genes. The high number of spacers in *C. forbesi* and *S. allegheniensis* could be a result of gene rearrangement (Rodovalho *et al*. 2014). However, additional studies characterizing the mitochondrial genomes of other Crangonyctidae species is needed to better understand the possible association of these intergenic spacers with gene rearrangements.

Gene order in *S. tenuis potomacus*, *S. pizzinii*, *S. allegheniensis*, and *B. brachycaudus* was almost identical to the ancestral pancrustacean gene order (Pons *et al*. 2014, Boore *et al*. 1998), except for the transposition of a few tRNA genes. However, *C. forbesi* exhibited quite distinctive gene rearrangements. *Stygobromus pizzinii*, *S. tenuis potomacus*, *S. allegheniensis*, and *B. brachycaudus* shared the conserved gene order of *trnF–nad5–trnH–nad4–nad4l* with other amphipod mitogenomes. Surprisingly, *C. forbesi* had a transposition of *trnH–nad4–nad4l* downstream after *nad6–cytb–trnS2*. The gene order in *C. forbesi* also differed from the ancestral pancrustacean gene order of *nad1–trnL1–rrnL–trnV–rrnS* with a transposition of *trnV* upstream between *trnM* and *nad2* and a transposition of *nad1* upstream between *trnT* and *trnM*. In addition, *S. tenuis potomacus*, *S. pizzinii*, *S. allegheniensis*, and *B. brachycaudus* shared the conserved gene order of *trnC-trnY-trnQ*; however, *C. forbesi* exhibited a transposition of *trnC* downstream after *trnl-trnG*. In addition, *C. forbesi* displayed several other transposed tRNA genes when compared to other amphipods.

ATN start codons, including ATT, ATC, ATG, and ATA, were the most frequently used start codons of most protein-coding genes. However, a few unconventional start codons were also used by protein-coding genes of a few species, including TTG for the *nad1* locus in *S. tenuis potomacus*, *S. pizzinii*, *S. allegheniensis*, and *B. brachycaudus*. However, *C. forbesi* used start codon GTG for *nad1*. Another unconventional start codon included GTG for the *cox2* locus in *S. tenuis potomacus* and *S. pizzinii* that has not been observed in other amphipod mitogenomes. In addition, *S. allegheniensis* used start codon GTG for the *nad5* locus. Most of the protein-coding genes used TAA or TAG stop codons; however, there were genes that used an incomplete TA– or T– – stop codon. Previous studies have shown that these incomplete stop codons are frequently observed in other amphipods (Pons *et al*. 2014; Romanova *et al*. 2016; Li *et al*. 2019). These incomplete stop codons can be modified into complete stop codons by post-transcriptional polyadenylation (Ojala *et al*. 1981).

Variation in the length of protein-coding genes and overlap between some adjacent protein-coding genes were observed among the five new crangonyctid mitogenomes and when in comparison with other amphipod mitogenomes. The *nad5* gene of *C. forbesi* was 1,974 bp in length and was substantially larger than in other amphipods (1,665–1,719 bp). This length difference is because of additional nucleotides found at the 3' end of the *nad5* gene. Contrastingly, the *nad2* gene of *S. allegheniensis* was 895 bp in length and was shorter than *nad2* in other species (943–1,023 bp). In addition, the *nad6* gene of *B. brachycaudus* was 531 bp in length and was longer than in other species (486–507 bp), while the *cytb* gene was 1098 bp in length and was shorter than in other compared species (1117–1140 bp). The *atp6* gene of *S. tenuis potomacus*, *S. pizzinii*, and *S. allegheniensis*

overlapped with the adjacent *atp8* gene by 41 bp, whereas no overlap was observed in *B. brachycaudus* and *C. forbesi*. An overlap between the *atp8* and *atp6* genes has been reported in other *Stygobromus* species including *S. tenuis* and *S. indentatus* (Aunins *et al*. 2016).

All 22 tRNA genes were encoded in the mitogenomes of *S. tenuis potomacus*, *S. pizzinii*, *S. allegheniensis*, *B. brachycaudus*, and *C. forbesi*. A total of fourteen tRNA genes were encoded on the heavy strand, and a total of eight tRNA genes were encoded on the light strand in all species. However, *C. forbesi* possessed several transposed tRNAs when compared to other amphipods and the ancestral pancrustacean gene order. The length of the tRNAs ranged from 50–66 bp, and most of the tRNAs had ideal cloverleaf secondary structures. However, *trnS1* and *trnS2* of *S. tenuis potomacus*, *S. pizzinii*, *B. brachycaudus*, and *C. forbesi* were missing the D loop, while *trnQ* was missing the TΨC loop, which has been observed previously in other *Stygobromus* species (Aunins *et al*. 2016). In addition, tRNAs of *B. brachycaudus* and *S. allegheniensis* displayed additional unique differences. The tRNAs *trnC* and *trnR* of *B. brachycaudus* lacked the D loop and *trnQ* lacked the acceptor stem in addition to missing the TΨC loop, which was not observed in other species. In contrast, tRNAs *trnS1* and *trnS2* of *S. allegheniensis* contained the D loop.

Alignment of the large ribosomal RNA (*rrnL*) gene of all five crangonyctid amphipods revealed the presence of a long sequence (52 bp) overhang on the 5’ end of the *rrnL* gene of *C. forbesi*. In the crangonyctid amphipod mitogenomes, the *rrnL* gene was flanked by tRNAs *trnL1* and *trnV*; however, the *rrnL* gene of *C. forbesi* was immediately located upstream of the small ribosomal RNA (*rrnS*) and both loci were flanked by tRNAs

trnL1 and *trnI*. The length of the *rrnL* gene in the other crangonyctid species (1034 – 1037 bp) was similar to that of other amphipod species.

The small ribosomal RNA (*rrnS*) gene of all five amphipod species had a relatively conserved 3′ end. Analogous to the *rrnL* gene, a long sequence (47 bp) overhang was observed on the 5′ end of the *rrnS* gene in *C. forbesi*. The 3′ end of the *rrnS* gene in all crangonyctid mitogenomes was followed by a continuous stretch of thymines, which was identified as the 5' end of the non-coding control region. However, the *rrnS* gene in *C. forbesi* was immediately followed by tRNA *trnI*. The control region of *C. forbesi* contained a transposition and was located downstream of the *nad1* gene.

A comparison between the two *S. tenuis potomacus* mitogenomes revealed few differences. Their gene order and AT content (69.1%) were identical, and both had equal numbers of intergenic spacers. The unconventional start codon GTG for the *cox2* locus found in the *S. tenuis potomacus* (MN175621) mitogenome we sequenced was not observed in the *S. tenuis potomacus* sequenced previously (KU869712; Aunins *et al*. 2016). The conventional start codon ATC was used instead. In addition, start codons for three other protein-coding loci (*cox1*, *nad3*, and *nad5*) differed in their 3rd position between the two mitogenomes. The *cytb* locus of *S. tenuis potomacus* (KU869712) used the incomplete stop codon T– – while the conventional stop codon TAA was used in *S. tenuis potomacus* (MN175621). In addition, the stop codon for *nad3* locus differed in their 3rd position (TAA versus TAG) between the two mitogenomes. Another difference observed was in lengths of the non-coding control region and *nad3* locus found. The control region of *S. tenuis potomacus* (KU869712) was 773 bp in length and was substantially larger than in *S. tenuis potomacus* (MN175621, 556 bp). In addition, the *nad3* locus of *S. tenuis potomacus*

(KU869712) was 381 bp in length and larger than in *S. tenuis potomacus* (MN175621, 354 bp). Apart from these differences, the location, structure, and length of transfer and ribosomal RNAs were nearly identical between the two mitogenomes.

Bayesian phylogenetic inference of the 13 protein-coding gene alignment (Figure 1) yielded a topology congruent with previous studies of amphipod phylogenetic relationships (Aunins *et al*. 2016; Romanova *et al*. 2016; Yang *et al*. 2017). Members of Crangonyctidae (*Bactrurus*, *Crangonyx*, and *Stygobromus*) form a well-supported clade. However, the genus *Stygobromus* was not monophyletic, as *B. brachycaudus* and *C. forbesi* were nested within *Stygobromus*. *Crangonyx forbesi* was sister to all other crangonyctids, although support for this relationship was lower. The paraphyly of *Stygobromus* also was uncovered in a phylogenetic analysis of the *cox1* locus (Niemiller *et al*. 2018).

1.4 Conclusion

In this study, we have added the complete mitogenomes of four groundwater amphipods — *S. pizzinii*, *S. allegheniensis*, *S. tenuis potomacus*, and *B. brachycaudus* — as well as the complete mitogenome of a spring-dwelling amphipod *C. forbesi* to the growing list of publicly available amphipod mitogenomes. Although all five new mitogenomes exhibited the complete set of 37 genes commonly observed in other amphipods, *C. forbesi* displayed several unique gene arrangements, including the transposition of genes *trnH–nad4–nad4l* downstream after *nad6–cytb–trnS2*. This particular gene arrangement deviates significantly from the ancenstral pancrustacean gene order and has not been detected in other amphipod species to date. Phylogenetic analysis supports the monophyly of Crangonyctidae but suggest that the genus *Stygobromus* is not

a monophyletic group. In general, our contribution of these five additional crangonyctid mitogenomes will be highly valuable for inferring the phylogenetic relationships, biogeography, and trait evolution of amphipods and investigating mitogenome evolution.

1.5 Data Availability Statement

The data that support the findings of this study are openly available in NCBI GenBank at https://www.ncbi.nlm.nih.gov/nuccore/MN175619.1 (*Bactrurus brachycaudus*), https://www.ncbi.nlm.nih.gov/nuccore/MN175620.1 (*Stygobromus pizzinii*), https://www.ncbi.nlm.nih.gov/nuccore/MN175621.1 (*Stygobromus tenuis potomacus*), https://www.ncbi.nlm.nih.gov/nuccore/MN175622.1 (*Stygobromus allegheniensis*), and https://www.ncbi.nlm.nih.gov/nuccore/MN175623.1 (*Crangonyx forbesi*)

The sample voucher numbers, related meta-data, and raw sequencing data are openly available in NCBI SRA RunSelector at https://www.ncbi.nlm.nih.gov/Traces/study/?acc=PRJNA657640.

CHAPTER 2. COMPARATIVE MITOGENOMIC ANALYSIS OF SUBTERRANEAN AND SURFACE AMPHIPODS (CRUSTACEA: AMPHIPODA) WITH SPECIAL REFERENCE TO THE FAMILY CRANGONYCTIDAE

2.1 Introduction

Caves and other subterranean habitats, such as groundwater aquifers and superficial subterranean habitats (SSHs; Culver and Pipan 2009), represent some of the most challenging environments that exist on Earth. The primary characteristic of all subterranean habitats is the lack of light and associated photosynthesis (Culver and Pipan 2009; Soares and Niemiller 2020). Though some subterranean ecosystems are supported by chemoautotrophic production by microbial communities (Engel *et al*. 2004; Porter *et al*. 2009), chemoautotrophy rarely provides enough energy to support several trophic levels in most subterranean ecosystems (Poulson and White 1969; Culver and Pipan 2009). The primary source of energy input for many cave systems is the organic matter transferred from the surface hydrologically or by animals that frequently enter and exit caves (Simon and Benfield 2001; Culver and Pipan 2009), which drive the structure and dynamics of subterranean communities (Huppop 2000; Graening and Brown 2003; Huntsman *et al*. 2011). Although most subterranean ecosystems are largely thought to be energy-limited (Venarsky *et al*. 2014), food availability can be highly variable both among and within cave systems (Culver *et al*. 1995; Juan *et al*. 2010). Previous studies have shown that many subterranean organisms living in such energy-limited habitats have undergone several physiological and metabolic adaptations to sustain themselves during extended food

shortages (Hervant *et al*. 1997; Issartel *et al*. 2010). Among these troglomorphic traits, low metabolic rate is a key adaptation that occurs in both terrestrial and aquatic fauna of subterranean communities (Bishop *et al*. 2014; Nair *et al*. 2020).

Mitochondria are the primary sites of energy production in cells, generating ~95% of the adenosine triphosphate (ATP) required for everyday activities of life through oxidative phosphorylation (OXPHOS; Das 2006; Shen *et al*. 2010; Yang *et al*. 2014). The mitochondrial genome—mitogenome—encodes 13 essential proteins including two ATP synthases (*atp6* and *atp8*), three cytochrome oxidases (*cox1, cox2,* and *cox3*), seven NADPH reductases (*nad1*, *nad2*, *nad3*, *nad4*, *nad4l*, *nad5*, and *nad6*), and cytochrome *b* (*cytb*) subunits. All mitochondrial protein-coding genes (PCGs) play a vital role in the electron transport chain (Boore 1999; Burger *et al*. 2003; Xu *et al*. 2007). Due to the unique characteristics of mitochondria, including maternal inheritance, small genomic size, absence of introns, and their surplus availability in cells, the use of mitochondrial DNA (mtDNA) loci and mitogenomes has a long history in population genetics, phylogenetics, and molecular evolution studies (*e.g.*, Ballard and Pichaud 2014; Bourguignon *et al*. 2018; Zou *et al*. 2018). Previous studies have demonstrated a close association between mitochondrial loci and energy metabolism (Shen *et al*. 2009, 2010; da Fonseca *et al*. 2008; Zhang *et al*. 2013). Although considered to largely evolve under purifying selection, there is growing evidence that mitogenomes may undergo episodes of directional selection in response to shifts in physiological or environmental pressures (Botero-Castro *et al*. 2018; Sun *et al*. 2020) leading to improved metabolic performance under new environmental conditions (da Fonseca *et al*. 2008; Garvin *et al*. 2011; Welch *et al*. 2014). For example, previous studies that investigated varying selective pressures acting on mitochondrial

PCGs of insects and mammals have revealed significant positive selective constraints at several loci that have comparatively increased energy demands (Shen *et al*. 2010; Yang *et al*. 2014; Li *et al*. 2018). Similarly, other studies have shown the various adaptive mitochondrial responses of organisms surviving in extreme environments including the deep sea and Tibetan Plateau (Mu *et al*. 2018; Li *et al*. 2018; Sun *et al*. 2020). However, these adaptations can occur at different metabolic levels, not just mitochondrial metabolism (Beall 2007; Scott *et al*. 2011). Thus, variation in mitogenomes of species inhabiting different environments may reflect only a small portion of these adaptive metabolic changes. Despite this limitation, previous studies have detected signals of directional selection in the mitogenomes of organisms dwelling in contrasting habitats with varying energy demands (Peng *et al*. 2011; Yang *et al*. 2015; Li *et al*. 2016)**.**

Amphipods (Class Malacostraca: Order Amphipoda) are one of the most ecologically diverse crustacean groups including over 10,000 species (Arfianti *et al.* 2018; Horton *et al.* 2020), occurring in a diverse array of aquatic and even terrestrial environments globally, from aphotic groundwater aquifers and hadal depths to freshwater streams and lakes in temperate and tropical forests, among other habitats (Bousfield 1983; Barnard and Karaman 1991). Several studies have demonstrated the genetic basis of subterranean adaptation in several taxa, including dytiscid diving beetles (Hyde *et al.* 2018), cave dwelling-cyprinid fishes (Wu *et al*. 2010; Dowling *et al*. 2002), anchialine cave shrimps (Guo *et al*. 2018), and cave isopods (Protas *et al*. 2011). However, we still have a limited understanding of the mechanisms of subterranean adaptations in amphipods. Although physiological adaptations to challenging environments like cave and groundwater ecosystems have been well-studied in amphipods (*e.g.*, Hervant *et al*. 1997;

Nair *et al*. 2020), no studies to date have addressed the selective pressures and the molecular evolution mechanisms of mitochondrial energy metabolism loci in amphipods occupying caves and other subterranean habitats. Subterranean amphipods likely experience different evolutionary pressures on energy management due to lower levels of predation, lower food resources, and more stable environments compared to surface-dwelling taxa that generally experience higher levels of predation and energy resources (Qiu *et al*. 2012; Qu *et al*. 2013).

In this study, we compared the mitogenomes of surface and subterranean amphipods, including the 13 mitochondrial PCGs involved in the OXPHOS pathway to understand the potential molecular mechanisms of energy metabolism in this diverse crustacean group. Our aims were to test whether the mitochondrial PCGs showed evidence of adaptive evolution in subterranean environments in amphipods. We tested the hypothesis that the mitogenome of surface-adapted amphipods will be imprinted by mitogenomic adaptations to the energy demanding environment with greater signal of directional selection when compared to their subterranean counterparts. We compared base composition, codon usage, gene order rearrangement, conducted comparative mitogenomic and phylogenomic analyses, and examined evolutionary signals using publicly available amphipod mitogenomes. In particular, we focused on the amphipod family Crangonyctidae, a diverse family that comprises species inhabiting a variety of surface and subterranean habitats and for which several mitogenomes have been sequenced and annotated recently (*e.g.*, Aunins *et al*. 2016; Benito *et al*. 2021).

2.2 Materials and Methods

We generated new mitogenomes recently for the following crangonyctid species: *Stygobromus pizzinii*, *S. tenuis potomacus*, *Bactrurus brachycaudus*, *Stygobromus allegheniensis*, and *Crangonyx forbesi* (Benito *et al*. 2021). We then downloaded from GenBank the annotated mitogenomes of 30 additional amphipod taxa that occupy aquatic habitats, including groundwater and springs, and three isopods that served as outgroups for comparative analyses.

2.2.1 Mitogenome Analyses

Nucleotide composition, amino acid frequencies, and codon usage were calculated in PhyloSuite v1.1.15 (Zhang *et al*. 2020). The web-based program CREx (http://pacosy.informatik.uni-leipzig.de/crex, Bernt *et al*. 2007) was used to perform pair-wise comparison of the gene orders in the mitogenome to determine rearrangement events. CREx comparisons were based on common intervals, and it considers common rearrangement scenarios including transpositions, reversals, reverse transpositions, and tandem-duplication-random-losses (TDRLs). In addition, phylograms and gene orders were visualized in iTOL (https://itol.embl.de/, Letunic and Bork 2021) using files exported from PhyloSuite. AT and GC skew of entire mitogenomes and individual loci were calculated in PhyloSuite using the formulae: AT-skew = (A - T)/(A + T) and GC-skew = (G-C)/(G + C). Welch two sample t-tests were performed between the surface and subterranean amphipods for different mitogenomic features, including mitogenome length, AT content, AT and GC skew, and rRNA length using R (R Core Team 2021).

Visualization of AT-skew, GC-skew, AT-content, and amino acid frequencies were generated in R.

2.2.2 Phylogenetic Inference

The amino acid sequences of 13 PCGs of the five new mitogenomes (Benito *et al.* 2021), 30 previously published amphipod mitogenomes, and three isopod mitogenomes were aligned using MAFFT version 7 (Katoh and Standley 2013). Poorly aligned regions were eliminated using Gblocks version 0.91b (Castresana 2000). The best partitioning strategy and best-fit evolutionary models for each partition were inferred using PartitionFinder v2.1.1 (Lanfear *et al*. 2012). Phylogenetic relationships of the 35 amphipod mitogenomes and three isopod mitogenomes using the concatenated 13 PCG alignment were determined using Bayesian inference in MrBayes v3.2 (Ronquist *et al*. 2012). The analyses contained two parallel runs with four chains each and were conducted for 5,000,000 generations (sampling every 100 generations). After dropping the first 25% "burn in" trees to ensure stationarity after examination of log-likelihood values for each Bayesian run using Tracer v1.7 (Rambaut *et al*. 2018), the remaining 37,500 sampled trees were used to estimate the consensus tree and the associated Bayesian posterior probabilities. All analyses were conducted within PhyloSuite.

2.2.3 Positive Selection and Selection Pressure Analyses of Mitochondrial PCGs

We performed base-substitution analyses on entire mitogenomes as well as for each of the 13 PCGs individually to compare surface versus subterranean amphipod taxa. The non-synonymous to synonymous rate ratio (dN/dS or ω) for each PCG was estimated using

the free-ratio model using the CodeML program implemented in PAML v4.8a (Xu and Yang 2013). The ω values were estimated for surface and subterranean species separately and visualized in R for comparison. To estimate the probability of positively selected sites in each PCG across all amphipod species, we implemented site models (M1 and M2, M8a and M8), where ω was allowed to vary among sites (Yang 2007). To further investigate if some lineages and sites in particular lineages have undergone positive selection, we conducted maximum likelihood analyses on all PCG using the branch model and branch-site model in EasyCodeML v1.21, a visual tool for analysis of selection using CodeML (Gao *et al*. 2019). For branch models, the 'one-ratio' model (model 0), and 'two-ratios' model were implemented in the combined dataset of 13 PCG as well as on each PCG to identify if selective pressure differed between an amphipod lineage of interest (foreground branch) and other amphipod lineages (background branches). A likelihood ratio test (LRT) was conducted for each PCG to test whether ω was homogenous across all lineages. In the branch model, the null hypothesis assumes that the average ω values of branches of interest (ωF) is equal to that of other branches (ωB), whereas the alternative hypothesis assumes the opposite ωF≠ωB. If the alternative hypothesis is supported and ω > 1, the foreground lineage is under positive selection. The branch-site model allows ω to differ among codon sites in a foreground lineage when compared to background lineages. We implemented the branch-site model to identify sites on specific lineages regulated by positive selection. Selected sites were considered positively selected only if they passed a Bayes Empirical Bayes (BEB) analysis with a posterior probability of >95%.

We performed selection pressure analyses on the concatenated 13 PCGs dataset aligned using the codon mode as well as on each PCG with the Bayesian topology (see

Figure 4) as the guidance tree using several approaches available from the Datamonkey Webserver (Weaver *et al*. 2018). First, we implemented aBSREL (Adaptive Branch-Site Random Effects Likelihood), an improved version of the commonly used "branch-site" models, to test if positive selection has occurred on a proportion of branches (Smith *et al*. 2015). We implemented BUSTED (Branch-site Unrestricted Statistical Test for Episodic Diversification) to test for gene-wide (not site-specific) positive selection by querying if a gene has experienced positive selection in at least one site on at least one branch (Murrell *et al*. 2015). Finally, we implemented RELAX (Wertheim *et al*. 2015) to test whether the strength of selection has been relaxed or intensified along a specified set of test branches.

2.3 Results and Discussion

We compared the mitogenomes for 35 surface and groundwater amphipods, including recently sequenced mitogenomes of one spring-dwelling and six groundwater species in the family Crangonyctidae by Aunins *et al*. (2016) and Benito *et al*. (2021), to determine whether subterranean species show evidence of adaptive evolution in subterranean habitats. Our study examined whether features of mitogenomes (*e.g.* base composition, codon usage, gene order) differed in relation to dominant habitat (surface vs. subterranean) and inferred the evolutionary forces potentially shaping mitogenome evolution in amphipods, with an emphasis on crangonyctid species.

2.3.1 Mitogenome Length and Content

Mitogenome sizes ranged from 14,113 to 18,424 bp for all amphipods and 14,661 to 15,469 bp for crangonyctid amphipods (Table 1). Mean mitogenome size of surface

amphipods (15,770 ± 1206 bp; mean ± 1 standard deviation) was higher than that of the subterranean amphipods (14,716 ± 297 bp) (t-test: t = -3.31, df = 15.3, p-value = 0.005; Supplementary Figure S1). All crangonyctid amphipod mitogenomes possessed 13 PCGs, two rRNA genes, 22 tRNA genes, a control region, and intergenic spacers of varying number and lengths (Supplementary Figure S2, annotations of the genomes are presented in Supplementary Table S1) like other arthropods (Clary and Wolsetenholme 1985). The length of the crangonyctid mitogenomes was similar to lengths reported for other amphipod families including Allocrangonyctidae, Caprellidae, Eulimnogammaridae, Gammaridae, Hadziidae, Lysianassidae, Metacrangonyctidae, Micruropodidae, Pallaseidae, Pontogeneiidae, Talitridae. Variation in mitogenome length within Crangonyctidae appears to be related to length variation in the *nad5, rrnL*, and *rrnS* loci, which were all notably longer in the *C. forbesi* mitogenome.

2.3.2 Base Composition and AT- and GC-skews

Mitogenome AT% in all amphipods ranged from 62.2 to 76.9% (Table 1). Mean AT% of the subterranean amphipods (71.8 ± 3.6%) was higher than that of the surface amphipods (67.6 ± 3.4%) (t-test: t = 3.49, df = 31.2, p-value = 0.001). Mean AT% of all 13 PCG of the subterranean amphipods was significantly higher than that of the surface amphipods (Supplementary Figure S3a). Variation in AT% across crangonyctid amphipod taxa ranged 63.9–69.3%, with a mean of 67.9 ± 1.93% (Table 1). Across loci, AT% ranged from a minimum of 60.0% at the *cox1* locus and a maximum of 75.5% at the *nad4l* locus (Figure 1A). Variation in AT% across all PCGs combined ranged from 61.9% (*B. brachycaudus*) to 69.0% (*S. indentatus*). Genes encoded on the negative strand had a

slightly higher AT-content values than those on the positive strand. The *nad6* locus showed the greatest variation in AT-content across species. *Bactrurus brachycaudus* displayed the outlier lower AT% values for most of the PCG (Table 2, Figure 1A). Similarly, *Bactrurus brachycaudus* had the lowest AT content (63.9%) among the crangonyctid mitogenomes, while all other mitogenomes had comparatively typical AT content reported for other arthropods (Crozier and Crozier 1993; Dotson and Beard 2001). This could indicate that the evolution of the *B. brachycaudus* mitogenome has occurred under different evolutionary pressures (adaptive or non-adaptive) than other subterranean crangonyctids.

Mitogenome AT-skew in all amphipods ranged from -0.062 to -0.037. Mean AT-skew of the surface amphipods (0.001 ± 0.02) was positive and slightly higher than that of the subterranean amphipods (-0.004 ± 0.02) (t = -0.63101, df = 31.117, p-value = 0.5326). Mean AT-skew of four PCG (*cox1*, *cox2*, *nad2*, *nad3*) of surface amphipods was significantly higher than that of the subterranean amphipods, whereas the mean AT-skew of *nad4* of the subterranean amphipods was significantly higher than that of the surface amphipods (Supplementary Figure S3b). Among crangonyctid amphipods, mean AT-skew was 0.022 ± 0.02 (range 0.004 to 0.061), with all mitogenomes displaying positive skew. Mitogenome GC-skew ranged from -0.431-0.120. Mean GC-skew of the subterranean amphipods (-0.129 ± 0.15) was negative and higher than that of the surface amphipods (-0.236 ± 0.05) (t = 3.01, df = 24.3, p-value = 0.006). Mean GC-skew of seven PCG (*atp6*, *atp8*, *cox1*, *cox2*, *cox3, nad2*, *nad3*) of subterranean amphipods was significantly higher than that of the surface amphipods, whereas the mean GC-skew of *nad4* of the surface amphipods was significantly higher than that of the subterranean amphipods (Supplementary Figure S3c). Among crangonyctid amphipods, mean GC-skew was -0.264

± 0.01 (range -0.275 to -0.248) with all mitogenomes displaying negative skew (Table 1). Strand asymmetry is commonly observed in mitogenomes (Reyes *et al*. 1998; Wei *et al*. 2010), however, at times it can hinder phylogenetic reconstruction and yield false phylogenetic artefacts especially when unrelated taxa display inverted skews (Hassanin *et al*. 2005; Zhang *et al*. 2019). *Bactrurus brachycaudus* exhibited the lowest AT skew among the crangonyctid mitogenomes (0.004), while *S. tenuis* had the lowest GC skew (−0.275). Crangonyctid amphipod mitogenomes exhibited positive GC-skew values for genes encoded on the (-) strand and negative GC-skew for genes encoded on the (+) strand (Figure 1C), whereas all PCGs exhibited negative AT-skew values (Figure 1B). Except the six loci (*nad1, nad4, nad4L, nad5, rrnL,* and *rrnS*) which were encoded on the (-) strand, most PCG had negative GC skews. Such strand bias is typical for most mitochondrial genomes in metazoan (Ki *et al*. 2010; Krebes and Bastrop 2012). This is consistent with the hypothesis that strand asymmetry is caused by spontaneous deamination of bases in the leading strand during replication (Reyes *et al*. 1998). All other mitogenomes had comparatively typical AT and GC skew values like other amphipod species (Pons *et al*. 2014; Romanova *et al*. 2016). The only outlier to this pattern was the positive GC skew value of tRNAs encoded on the (+) strand of B. *brachycaudus* (0.012). In general, crangonyctid mitogenomes exhibited relatively consistent skews.

Higher AT-content was observed at first ($t = 3.80$, $df = 67.8$, p-value < 0.001), second ($t = 5.13$, $df = 67.9$, p-value < 0.001), and third codon positions ($t = 4.26$, $df = 60.7$, p-value < 0.001) of PCGs on both the (+) and (-) strands in subterranean amphipods compared to surface amphipods (Supplementary Figure S3a). Among crangonyctid amphipods, a contrasting pattern was observed in the nucleotide composition per codon

position (Figure 1A). AT-content was higher at third positions on both strands (74.5 ± 4.1% on (+) strand; 74.6 ± 4.4% on (-) strand) compared to first (58.5 ± 0.9% on (+) strand; 63.0 ± 2.7% on (-) strand) and second positions (62.4 ± 0.6% on (+) strand; 62.8 ± 0.9% on (-) strand). AT skew was near zero at the first codon position, whereas a T nucleotide-enrichment (about -0.4 AT skew value on average) in loci of both strands was observed at the second codon position. Intermediate negative AT skews was observed at the third codon position (Figure 1B). GC skew was positive for the first codon position, negative or close to zero for the second codon position and showed greater variation at third codon positions between loci on the positive and negative strands (Figure 1C).

2.3.3 Rearrangements of Mitochondrial Genome

Conserved gene blocks in crangonyctid mitogenomes include: (1) cox1-tRNA-*L2*-*cox2*-tRNA-*K,D-atp8-atp6-cox3-nad3*-tRNA-*A,S1,N,E,R,F-nad5;* (2) tRNA-*H-nad4-nad4l;* (3) *nad6-cytb*-tRNA-*S2;* (4) tRNA-*L1-rrnL*; (5) *rrnS*-tRNA-*I;* and (6) tRNA-*Y,Q* (Figure 2B). The gene orders in subterranean species (genera *Stygobromus* and *Bactrurus*) are identical except for the transposition of tRNA-*G,W*. However, a few unique gene order arrangements were observed in the spring-dwelling *C. forbesi*. The gene order of *C. forbesi* differs from the four subterranean species in the locations of the conserved gene blocks (tRNA-*H-nad4-nad4l* and *nad6-cytb*-tRNA-*S2* and tRNA-*L1-rrnL* and *rrnS*-tRNA-*I* and tRNA-*Y,Q*), seven tRNAs (*P,T,M,V,G,C, and W*), and two protein-coding loci: *nad1* and *nad2*. Compared to the conserved mitogenome gene orders of other crangonyctid mitogenomes, another unique feature in the rearranged *C. forbesi* mitogenome was the presence of at least two long (~50 and 70 bp) non-coding regions (Supplementary Table

S1). The locations of rRNA genes in all crangonyctid mitogenomes are mostly similar compared to the pancrustacean ground pattern except for *C. forbesi* where the rRNA genes had altered positions (Figure 2A and 2B). Rearrangements in the mitogenome is common especially when it involves only tRNA-coding genes (Matsumoto *et al*. 2009). In case of ribosomal RNA genes or PCGs, rearrangements occur much less frequently, and they are commonly referred to as major rearrangements, as they might potentially affect the differential regulation of replication and transcription of mitogenomes (Rawlings *et al*. 2001).

CREx analysis indicated that transpositions and TDRL may have been responsible for the evolution of mitogenomes in crangonyctid amphipods. Two transpositions of tRNA-*R,N,S1,E* and two steps of TDRL from the ancestral pan-crustacean pattern were needed to generate the gene order observed in *Stygobromus* species. In addition to the same two transpositions, one TDRL, and a transposition within a second TDRL from the ancestral pattern were required to generate the gene order in *Bactrurus*. However, four different transpositions (tRNA-*N,S1,* tRNA-*T,P,* tRNA-*W,C* and gene block tRNA-*H-nad4-nad4L*-tRNA-*P,T-nad6-cytb*-tRNA-*S2*) and three steps of TDRL from the ancestral pattern were needed to generate the gene order observed in *C. forbesi* (Supplementary Figure S4).

Similar to *C. forbesi*, other surface amphipods including *Gmelinoides fasciatus* (Micruropodidae) and *Onisimus nanseni* (Lysianassidae) exhibited a highly rearranged gene order. Other surface amphipods that exhibited a moderate to highly rearranged gene order include *Gondogeneia antarctica* (Pontogeneiidae), *Platorchestia parapacifica* and *P. japonica* (Talitridae), *Pallaseopsis kessleri* (Pallaseidae), and the two basal amphipods

Caprella scaura and *C. mutica* (Caprellidae) (Figure 2A). Interestingly, a subterranean amphipod *Pseudoniphargus daviui* (Allocrangonyctidae) also exhibited a moderate rearranged gene order. The stark contrast between the highly conserved gene order in most subterranean amphipods and the highly volatile gene order in many of the surface amphipods may support the hypothesis that evolution of mitogenomic architecture could be highly discontinuous. A prolonged period of inactivity in gene order and content could have been dictated by a rearrangement event resulting in a destabilized mitogenome, which is much more likely to undergo an exponentially accelerated rate of mitogenomic rearrangements (Zou *et al*. 2017). Thus, it is appealing to examine mitogenomes of surface amphipod families represented by just a single taxon in our dataset.

2.3.4 Codon Usage and Amino Acid Frequencies

In addition to the regular start codons (ATA and ATG) and uncommon start codons (ATT, ATC, TTG, and GTG), surface amphipods, particularly *Caprella scaura*, possessed one rare start codon CTG, whereas subterranean amphipods possessed three rare start codons including CTG, TTT, and AAT to initiate the mitochondrial PCGs (Supplementary Table S2). Codon usage analysis of the five crangonyctid amphipods mitogenomes identified the existence of all codon types typical for any invertebrate mitogenome. In addition to the regular start codons (ATA and ATG), uncommon start codons (ATT, ATC, TTG, and GTG) were also present to initiate the mitochondrial PCG. Such unusual start codons have been reported previously in other arthropods (Lessinger *et al*. 2000; Boore *et al*. 2005). A few PCG in the crangonyctid mitogenomes possessed truncated or incomplete stop codons (TA- and T--) that have been described in other crustaceans (Supplementary

Table S1). These are presumably completed after a post-transcriptional polyadenylation (Ojala *et al*. 1981; Castresana *et al*. 1998; Nardi *et al*. 2001). Among the crangonyctid mitogenomes, the most frequently used codons are TTA (Leu2; 5.64% to 8.49%) and TTT (Phe; 5.94% to 6.78%). Other frequently used codons include ATT (Ile; 4.92% to 6.85%) and ATA (Met; 4.13% to 5.34%) (Supplementary Table S3). These four codons are also among the most abundant in non-crangonyctid amphipods included in this study. This bias towards the AT-rich codons is quite typical for arthropods (Wilson *et al*. 2000). Among crangonyctid amphipod mitogenomes, relative synonymous codon usage (RSCU) values, which is the measure of the extent that synonymous codons depart from random usage, showed a high prevalence of A or T nucleotides at third codon positions (Figure 3). This trend was also observed in other subterranean and surface amphipods. This positive correlation between RSCU and AT content at third codon positions has been reported in mitochondrial genomes of the abalone and oyster (Ren *et al*. 2010; Xin *et al*. 2011).

In PCGs, the second copy of leucine (8.86–10.01%) and cysteine (0.95–1.17%) are the most and the least used amino acids, respectively. Amino acid frequency analysis of both surface and subterranean amphipods indicated that five amino acids (leucine, phenylalanine, isoleucine, methonine, and valine) account for more than half of the total amino acid composition and exhibited greater variation among species (Supplementary Figure S5; Supplementary Table S4).

2.3.5 Transfer RNA Genes

All 22 tRNA genes were identified in the mitogenomes of crangonyctid amphipods. However, the locations of tRNA genes were highly variable among these mitogenomes,

and they also displayed altered positions relative to the pancrustacean ground pattern (Figure 2; Supplementary Figure S3). The secondary structures of all mitogenome-encoded tRNAs belonging to crangonyctid amphipods were predicted and ranged in length from 50 to 66 bp. Most of the tRNAs displayed the regular clover-leaf structures, however, a few displayed aberrant structures. The tRNA-Ser1 (UCU) lacked the DHU arm in all crangonyctid species. Similarly, the tRNA-Ser2 (UGA) lacked the DHU arm in all crangonyctid species except *S. allegheniensis* where tRNA-Ser2 (UGA) possessed the DHU arm. The DHU arm was also missing in the tRNA-Cys and tRNA-Arg of *B. brachycaudus* and tRNA-Arg of *C. forbesi*. The tRNA-Gln lacked the TψC arm in all crangonyctid species except *C. forbesi* where tRNA-Gln possessed the TψC arm. In addition to lacking the TψC arm, tRNA-Gln of *B. brachycaudus* lacked the acceptor stem as well (Supplementary Figure S6). The presence of aberrant structures in tRNAs have been observed in several other crustaceans and invertebrates (Okimoto *et al*. 1990; Bauza-Ribot *et al*. 2009; Ito *et al*. 2010; Lin *et al*. 2012), which may be the result of replication slippage (Macey *et al*. 1997) or selection towards minimization of the mitogenome (Yamazaki *et al*. 1997).

2.3.6 Ribosomal RNA Genes

The length of *rrnL* genes in all amphipods ranged 604–1,137 bp and that of *rrnS* genes ranged 196–1,631 bp. *rrnL* length of the subterranean amphipods (1,055 ± 26 bp) was higher than that of the surface amphipods (971 ± 108 bp) (t-test: t = -2.94, df = 15.2, p-value = 0.001). *rrnS* length of the surface amphipods (694 ± 290 bp) was slightly higher than that of the subterranean amphipods (684 ± 16 bp) (t-test: t = 0.13, df = 14.1, p-value

= 0.896). The length of *rrnL* genes in crangonyctid amphipods ranged 1,034–1,090 bp and that of *rrnS* genes ranged 671–695 bp. The length of rRNA genes in crangonyctid amphipods was similar to that of other amphipod mitogenomes except *C. forbesi*, which had long overhangs (~50 bp and ~25 bp) on the 5' end of the *rrnL* and *rrnS* genes, respectively. AT content ranged 67.8–72.8% in the *rrnL* genes and 71.5–77.2% in the *rrnS* genes of crangonyctid species, respectively. GC-skew values for rRNA genes were positive (0.259 to 0.426) and comparable to that of PCGs encoded on the (-) strand (Supplementary Table S5).

2.3.7 Control Region and Intergenic Spacers

In the mitogenome of *S. pizzinii* the putative control region (CR) was identified as a 1,021 bp sequence between the *rrnS* gene and the *trnl-trnM-trnC-trnY-trnQ-nad2* gene cluster. Similarly, CR was observed in the other crangonyctid mitogenomes, including *S. tenuis* (556 bp), *S. allegheniensis* (991 bp), *B. brachycaudus* (531 bp), *S. indentatus* (535 bp), and *S. tenuis potomacus* (773 bp). The CR was similarly located between the *rrnS* and *nad2* genes in some of the other mitogenomes of non-crangonyctid amphipods, including *G. duebeni (*Krebes and Bastrop 2012), *O. nanseni* (Ki *et al*. 2010), *G. antarctica* (Shin *et al*. 2012), *P. daviui* (Bauzà-Ribot *et al*. 2012), and for the pancrustacean ground pattern. However, the adjacent tRNA genes were often different. In *G. fasciatus,* the CR region was located between the *rrnS* and *nad5* genes (Romanova *et al*. 2016). In contrast, a control region 843 bp was observed in *C. forbesi* which is located between the *nad1* and *trnM-trnV-nad2* gene cluster and separated by few intergenic spacers was identified as the CR (Supplementary Figure S1; Supplementary Table S1). The only other surface amphipod

that had a similar CR location to *C. forbesi* was *P. kessleri* with the CR located between *nad1* and *nad2* genes, although the adjacent tRNA genes were different (Romanova *et al*. 2016). Thus, the variable location of the CR in *C. forbesi* was in concordance with a few surface amphipods, while the subterranean amphipods mostly followed the universal pattern between *rrnS* and *nad2* genes.

The non-coding regions or intergenic spacers identified in the crangonyctid mitogenomes varied in number and length. The number of intergenic spacers ranged from 7 to 17 and their lengths ranged from 1 to 99 bp (mean 13.0 bp ± 18.6). Two crangonyctid mitogenomes (*S. allegheniensis* and *C. forbesi*) possessed the largest intergenic spacers (a total of 220 bp and 249 bp, respectively; Supplementary Table S1). Among the non-crangonyctid amphipods, *G. fasciatus* and *G. antarctica* possessed relatively large non-coding intergenic spacers (a total of 3,863 bp and 4,354 bp, respectively; Shin *et al*. 2012; Romanova *et al*. 2016).

2.3.8 Phylogenetic Inference

The phylogenetic analyses of the 13 concatenated PCG from 35 amphipod species using Bayesian Inference (BI) resulted in a well-supported phylogeny, with the crangonyctid species forming a well-supported monophyletic group (Figure 4). Within Crangonyctidae, *Stygobromus* species formed a monophyletic group sister to *Bactrurus* + *Crangonyx*; however, few crangonyctid taxa were included in our analysis. A previous study based on the *cox1* gene found that *Stygobromus* was not monophyletic, but several relationships had low support (Niemiller *et al*. 2018). Likewise, *Stygobromus* was recovered as polyphyletic in a multilocus concatenated phylogenetic analysis of the

Crangonyctidae by Copilas-Ciocianu *et al*. (2019). In addition, several well-supported clades were recovered within Crangonyctidae but relationships among these clades had low support. Other past studies have not supported monophyly of the widespread genera (*i.e.*, *Crangonyx*, *Stygobromus*, and *Synurella*) in the family based on either morphological (Koenemann and Holsinger 2001) or molecular data (Kornobis *et al*. 2011, 2012). A comprehensive phylogenomic study with robust taxonomic sampling is greatly needed to better elucidate evolutionary relationships and test biogeographic and ecological hypotheses regarding the origin and diversification of this diverse family of subterranean and surface-dwelling amphipods.

2.3.9 Selection in PCGs

Most of the energy required for active movement to escape predation and meet energy demands is supplied by the mitochondrial electron transport chain (Shen *et al*. 2009, 2010). Mitochondrial genes encode for all of the protein complexes related to oxidative phosphorylation except for succinate dehydrogenase (complex II) (Scheffler 1998; Carroll *et al*. 2009; McKenzie *et al*. 2009). Because of their importance in oxidative phosphorylation during cellular respiration, it is unsurprising that many studies have shown evidence of purifying (negative) selection in mitochondrial PCG (Meiklejohn *et al*. 2007; Hao *et al*. 2017; Sun *et al*. 2020). While we found strong evidence for purifying selection in amphipod mitochondrial PCG in our selection analyses, we also found signatures of positive selection in a few of the mitochondrial PCGs in the surface amphipods.

The one-ratio model (model 0) analyses conducted for all 13 PCG revealed that the ω values for each gene ranged from 0.021 to 0.130 and were significantly less than 1 (Table

3). Using a free-ratio model (model 1; Shen *et al*. 2009), we calculated the ω values for the 13 PCG for the terminal branches to estimate the strength of selection between different primary habitats (*i.e.*, subterranean vs. surface). The *cox2* gene significantly differed in ω values between the amphipods of the two habitat types ($p = 0.020$), with higher ω values for the surface amphipods. Similarly, *cox1* and *cox3* genes also exhibited a similar trend ($p = 0.095$ and $p = 0.057$, respectively) (Figure 5). This could be because the rate at which slightly deleterious mutations (ω) responsible for the mitochondrial gene evolution accumulates comparatively faster in *cox* gene family of the surface lineages. However, this result is quite contradictory to previous studies showing higher functional constraint and conserved pattern in the genes coding for *cox* proteins than in other mitochondrial genes (Meiklejohn *et al*. 2007; Zapelloni *et al*. 2021).

To test if specific branches have undergone variable selective pressures, especially those amphipod branches adapted to surface habitats, we employed the two-ratio branch model *(*model 2). We evaluated the selective pressures acting on surface amphipods (foreground, $\omega 1$) and subterranean amphipods (background, $\omega 0$). LRT tests showed that the two-ratio model fits were significantly better than the one-ratio model for two genes: *atp6* ($p = 0.018$) and *cytb* ($p = 0.002$) (Table 3), indicating a divergence in selective pressure between surface and subterranean amphipods. In addition, a similar trend was observed for the *cox2* gene ($p = 0.072$). We then tried the same two-ratio model to estimate selection acting on each subterranean and surface amphipod lineage to further examine the difference between them. When the ω values for the individual PCG were compared between each amphipod lineage and the other 34 amphipod taxa, several genes in surface amphipod mitogenomes were found to be undergoing positive selection ($\omega 1 > \omega 0$; Figure

6; Supplementary Table S6). This suggests that many surface amphipods have experienced directional selection in their mitochondrial genes due to high energy demands and was accordance to the previous studies of insect orders (Yang *et al*. 2014; Mitterboeck *et al*. 2017; Li *et al*. 2018). In contrast, several genes in subterranean amphipod mitogenomes were undergoing purifying selection ($\omega 1 < \omega 0$). Surprisingly, a few genes in subterranean taxa displayed positive selection ($\omega 1 > \omega 0$; Figure 6; Supplementary Table S6). This indicated that the subterranean amphipods have undergone stronger evolutionary constraints to remove deleterious mutations to maintain efficient energy metabolism (Shen *et al*. 2009). Overall, this proves that molecular evolution of mitochondrial genes is correlated to the changes in the energy requirements.

To test if individual gene codons were subject to positive selection, we implemented two pairs of site models (M1a vs. M2a and M8a vs. M8). The M8 model identified one positively selected site on the *atp8* gene (37 N; $p = 0$) and one positively selected site on the *nad5* gene (482 Q; $p = 0$). Similarly, The M2a model identified two positively selected sites (37 N & 31 S; $p = 0.0194$) on the *atp8* gene (Table 4).

Similar to the mitochondrial genes of flying grasshoppers that have evolved to adapt to increased energy demands to maintain the flight capacity (Li *et al*. 2018), the mitochondrial genes of surface amphipods may have evolved mechanisms to meet increased energy demands due to predation, dispersal, etc. Although surface amphipods appear to be evolving under selective pressures different from those of the subterranean taxa and their mitochondrial genes have accumulated more nonsynonymous than synonymous mutations compared to subterranean taxa, the branch model tests did not clearly support positive selection on these branches, and we cannot rule out the influence

of relaxed selection. Previous studies have demonstrated that positive selection will act on only a few sites for a short period of evolutionary time, and a signal of positive selection often is overwhelmed by continuous negative selection that sweeps across most sites in a gene sequence (Zhang *et al*. 2005).

In contrast to branch models where ω varies only among branches, branch-site models allow selection to vary both among amino acid sites and lineages. Thus, branch-site models are considered quite useful in distinguishing positive selection from relaxed or purifying selection (Zhang *et al*. 2005). Using the more stringent branch-site model, we detected positive selection in 14 branches and 12 genes with a total of 308 amino acid sites under positive selection. Among them, 80 amino acid sites in seven genes (*atp6*, *atp8*, *cox3*, *nad2*, *nad3*, *nad4*, and *nad5*) were identified on the subterranean branches, whereas 228 amino acid sites in 10 genes (*atp6*, *atp8*, *cox1*, *cox2*, *cytb*, *nad1*, *nad2*, *nad3*, *nad5*, and *nad6*) were identified on the surface branches. Nearly three times as many positively selected amino acid sites were detected on surface branches compared to subterranean branches. Most of the positively selected genes on surface branches were found in *C. forbesi* with 114 sites (Figure 7; Supplementary Table S7). In total, eight positive selected genes (*atp6*, *atp8*, *cox1*, *cox2*, *cytb*, *nad1*, *nad4*, and *nad5*) were identified by the branch-site model and by at least one other model on the surface branches, whereas only four positive selected genes (*atp6*, *atp8*, *cox3*, and *nad5*) were identified on the subterranean branches.

The identification of many positively selected amino acid sites suggests that episodic positive selection has acted on mitochondrial PCGs of surface amphipods. In addition, we also identified a few positively selected sites on subterranean branches

primarily in *B. brachycaudus* with 39 sites and *P. daviui* with 25 sites (Supplementary Table S7). *Bactrurus brachycaudus* is usually associated with springs and caves (Taylor and Niemiller 2016), whereas *P. daviui* associated with groundwater wells (Bauzà-Ribot *et al*. 2012).

2.3.10 Direction and Magnitude of Selection Pressures

Given the crucial role played by the mitochondrial genome in metabolic energy production (Hassanin *et al*. 2009), we hypothesized that the mitogenome of surface amphipods may show evidence of adaptation (directional selection) to life in surface habitats where energy demand is higher relative to subterranean habitats. We found support for directional selection in surface lineages based on three different selection analyses (RELAX, aBSREL, and BUSTED). In summary, all tests confirmed the existence of a moderate signal of positive or diversifying selection, as well as signal for significant relaxed purifying selection in the mitogenome of surface amphipods. This supports a previous study by Carlini and Fong (2017) who reported evidence for relaxation of functional evolutionary constraints (positive or diversifying selection) in the transcriptome of a subterranean amphipod *Gammarus minus*. The authors correlated the signal to the adaptation of this species to the subterranean environment.

We implemented aBSREL on the concatenated 13 PCG dataset comprising all 35 species as test branches and detected episodic diversifying selection in seven species: *P. daviui* ($p = 0$), *O. nanseni* ($p = 0.0008$), *G. fasciatus* ($p = 0.0298$), *G. fossarum* ($p = 0.045$), *B. jaraguensis* ($p = 0.0016$), *C. forbesi* ($p = 0$), and *B. brachycaudus* ($p = 0.0001$). We then used aBSREL to conduct independent tests for the crangonyctid species as the test branch

and the remaining species as reference branches. We detected evidence of episodic diversifying selection in *C. forbesi* ($p = 0$) and *B. brachycaudus* ($p = 0.0001$) (Table 5). Using BUSTED, which provides a gene-wide test for positive selection, we detected evidence of episodic diversifying selection in three of the surface species: *C. forbesi* ($p = 0.011$), *G. fasciatus* ($p = 0.033$), *G. antarctica* ($p = 0.009$), whereas evidence of gene-wide episodic diversifying selection was found in just one of the subterranean species, *P. daviui* ($p = 0.020$) (Table 5). Using RELAX, which tests whether the strength of selection has been relaxed or intensified along a specified set of test branches, we detected selection evidence of relaxed selection in *C. forbesi* ($p = 0$) and other surface species, including *O. nanseni, G. fasciatus, G. fossarum, G. antarctica,* and *P. kessleri* with a p value of 0. Contrastingly, evidence of intensification of selection was detected in subterranean species including *S. tenuis* ($p = 0$), *S. allegheniensis* ($p = 0.0025$), *S. indentatus* ($p = 0$), and *S. pizzinii* ($p = 0$). Surprisingly, a few of the surface species including *C. mutica* ($p = 0.015$), *E. cyaneus* ($p = 0$), and *P. japonica* ($p = 0$) exhibited intensification of selection and subterranean species including *P. daviui* ($p = 0$) and *M. dominicanus* ($p = 0.015$) exhibited relaxation of selection (Table 5).

In addition to the concatenated 13 PCG dataset, we also conducted selection analyses for each PCG to determine which genes might be evolving under unique selection pressures. We found evidence of directional selection in *atp8* of *C. forbesi* ($p = 0.026$) and *nad3* of *S. pizzinii* ($p = 0.041$) using aBSREL and *cox3* of *B. brachycaudus* ($p = 0.029$) using BUSTED. *Atp8* of the surface amphipod *C. forbesi* exhibited strong evidence of directional selection, which was quite surprising as *atp8* is a small gene sometimes missing from metazoan mitogenomes and normally evolves under highly relaxed selection

pressures (Gissi *et al*. 2008). Based on missing evidence for relaxation selection pressures in subterranean amphipods, we can confirm that *atp8* is indeed evolving under predominantly strong directional selection in surface amphipods which needs to be explored further. RELAX analyses uncovered five genes (*cox1, cox3, cytb, nad1,* and *nad3*) that exhibited relaxed selection and one gene (*atp6*) that exhibited intensification of selection in *C. forbesi*. Similarly, three genes (*cox3, nad5,* and *nad6*) in *B. brachycaudus* showed evidence of relaxed selection. Several genes in other subterranean species, including S. *tenuis, S. allegheniensis,* and *S. pizzinii*, exhibited varying levels of intensification of selection, whereas none exhibited relaxed selection (Table 6). Some of these outliers were expected, as *nad5* and *nad6* are known to evolve faster among the mitochondrial genes (Zhang *et al*. 2018). Also, evidence for relaxation of functional evolutionary constraints (positive or diversifying selection) has been reported in the *nad* family of subterranean *Gammarus* species adapted to the subterranean environment (Carlini and Fong 2017). Although this may explain outliers in the subterranean *B. brachycaudus* mitogenome, it remains unclear why *cox3* exhibited signatures of relaxed selection. This gene is generally one of the most conserved mitochondrial loci in animals (Oliveira *et al*. 2008; Xiao *et al*. 2011; Pons *et al*. 2014), and high levels of purifying selection has been observed in the *cox* family in other amphipod species (Sun *et al*. 2020). In *C. forbesi*, *atp6* showed signatures of positive selection, which contrasted most other PCGs in its mitogenome that exhibited relaxed selection. Overall, in accordance with the results obtained using the concatenated dataset, individual mitochondrial genes of subterranean amphipods mostly exhibited varying levels of purifying selection, whereas surface amphipods predominantly exhibited relaxed selection.

To provide further evidence of positive selection, we implemented the RELAX, aBSREL, and BUSTED algorithms on the branch, branch-site, and site models. Eight genes (*atp8*, *cox1*, *cox2*, *cytb*, *nad1*, *nad4*, *nad5*, and *nad6*) all involved in the OXPHOS pathway were under positive selection in surface branches by at least two methods. The genes *nad1*, *nad4*, *nad5*, and *nad6* encode the subunits of NADH dehydrogenase also called Complex I that initiates the oxidative phosphorylation process. Complex I is the largest and most complicated proton pump of the respiratory chain. Involved in electron transfer from NADH to ubiquinone to supply the proton motive force used for ATP synthesis (Wirth *et al*. 2016), Complex I plays a key role in cellular energy metabolism by pumping gradient of protons across the mitochondrial membrane producing more than one-third of mitochondrial energy (Dröse *et al*. 2011). This would clarify the reason behind detecting more evidence of positive selection in complex I than in other complexes. Genes *cox1* & *cox2* which act as regulators encodes the catalytic core of Cytochrome c oxidase also called as Complex IV. Complex IV is directly involved in electron transfer and proton translocation (Zhang *et al*. 2013). Gene *atp8* encodes a part of ATP synthase, also called Complex V, and plays a major role in the final assembly of ATPase (Zhang *et al*. 2013). In summary, our selection analyses revealed signals of positive selection in several mitochondrial genes of surface amphipods, which may be associated with increased energy demands in surface environments. In contrast, subterranean amphipods showed signatures of purifying selection, which may be related to maintaining efficient energy metabolism in subterranean habitats.

2.4 Conclusion

In this study, we compared mitogenome features including AT/GC-skew, codon usage, gene order, phylogenetic relationships, and selection pressures acting upon amphipods inhabiting surface and subterranean habitats. We described a novel mitochondrial gene order for *C. forbesi*. We identified a signal of directional selection in the protein-encoding genes of the OXPHOS pathway in the mitogenomes of surface amphipods and a signal of purifying selection in subterranean species, which is consistent with the hypothesis that the mitogenome of surface-adapted amphipods has evolved in response to a more energy demanding environment compared to subterranean species. Our comparative analyses of gene order, locations of non-coding regions, and base-substitution rates points to habitat as an important factor influencing the evolution of amphipod mitogenomes. However, the generation and study of mitogenomes from additional amphipod taxa, including other crangonyctid species, are needed to better elucidate phylogenetic relationships and the evolution of mitogenomes of amphipods. In addition, more evidence is needed to further validate our inferences, such as studying the effects of amino acid changes on three-dimensional protein structure and function. Nevertheless, our study provides a necessary foundation for the study of mitogenome evolution in amphipods and other crustaceans.

CHAPTER 3. PHYLOGENOMICS AND BIOGEOGRAPHY OF NORTH AMERICAN TRECHINE CAVE BEETLES (COLEOPTERA: CARABIDAE)

3.1 Introduction

Caves are home to unique and diverse communities, and they represent one of the most unforgiving environments on Earth. The primary characteristic of caves and subterranean habitats is the lack of light and associated photosynthesis leading to limited food resources (Culver and Pipan 2009; Soares and Niemiller 2020). Cave organisms tend to exhibit similar morphological, physiological, and behavioral traits, such as loss of eyes, reduction of pigmentation, development of nonvisual sensory organs, changes in metabolism, and longer lifespan (termed troglomorphy; Racovitza 1907; Vandel 1964; Culver *et al*. 1990; Culver and Pipan 2009). Caves and their fauna have been viewed as model systems for addressing fundamental questions in evolutionary biology, ecology, biogeography, and speciation (Poulson and White 1969; Juan *et al*. 2010; Mammola 2019). However, testing hypotheses related to speciation, evolution, and biogeography of cave fauna often has been hampered in the past due to difficulties associated with sampling subterranean habitats, convergence in morphology, and extinction of surface ancestors (Holsinger 2000; Porter 2007; Juan and Emerson 2010). Despite these challenges, new molecular approaches developed in recent years have provided an opportunity to better understand the evolutionary and biogeographic processes that facilitate adaptation and speciation and shape distributions in subterranean environments (Jeffery 2009; Juan *et al*. 2010; Liu *et al*. 2017; Torres-Dowdall *et al*. 2018). However, many questions remain unanswered or poorly investigated (Juan *et al*. 2010; Morvan *et al*. 2013; Mammola *et al*.

2019). For example, the relationships between many subterranean and surface lineages remain largely unknown and the timing and patterns of diversification of many cave organisms are understudied.

In North America, subterranean biodiversity is primarily associated with distinct karst regions (Culver *et al*. 2003; Hobbs *et al*. 2012; Niemiller *et al*. 2019), a type of landscape underlain by carbonate rocks that possess dense networks of subterranean drainage systems, sinkholes, springs, and caves (Ford and Williams 2007). Four karst regions – Interior Low Plateau, Appalachians, Edwards Plateau & Balcones Escarpment, and Ozarks – account for nearly 80% of the more than 1350 described troglobiotic (terrestrial) and stygobiotic (aquatic) diversity in the continental United States and Canada (Niemiller *et al*. 2019). Several phylogenetic studies in recent years have advanced our knowledge of the biogeography, evolution, and systematics of many subterranean organisms in North America, shedding light on distributional patterns, cryptic diversity, colonization history, and modes of speciation. However, most studies have focused largely on three primarily aquatic taxonomic groups that account for only ~4% of subterranean biodiversity in the United States and Canada (Niemiller *et al*. 2019): cavefishes (*e.g.*, Dillman *et al*. 2011; Garcia-Machado *et al*. 2011; Niemiller *et al*. 2012, 2013a,b; Strecker *et al*. 2012; Fumey *et al*. 2018; Hart *et al*. 2020), salamanders (*e.g.*, Chippindale *et al*. 2000; Hillis *et al*. 2001; Wiens *et al*. 2003; Niemiller *et al*. 2008, 2009; Bendik *et al*. 2013; Phillips *et al*. 2017; Devitt *et al*. 2019; Grant *et al*. 2022), and crayfishes (*e.g.*, Buhay and Crandall 2005, 2008, 2009; Buhay *et al*. 2007; Stern *et al*. 2017; Dooley *et al*. 2022). Recent North American phylogenetic studies involving troglobiotic invertebrates are comparatively few, but include spiders (*e.g.*, Hedin 1997a,b, 2015; Paquin and Hedin 2004;

Snowman *et al*. 2010; Hedin *et al*. 2018), harvestmen (Derkarabetian *et al*. 2010, 2022; Hedin and Thomas 2010), springtails (Katz *et al*. 2018), millipedes (Loria *et al*. 2011), and beetles (Gomez *et al*. 2016; Leray *et al*. 2019). Surprisingly underrepresented in phylogenetic studies are trechine cave beetles (Coleoptera, Carabidae, Trechini), which account for 18% of all described troglobiotic diversity in the United States and Canada (Niemiller *et al*. 2019).

Cave trechines are prominent in many terrestrial subterranean habitats in Asia, Europe, and North America, and they are an ideal model system to study the evolution and biogeography of subterranean life (Faille *et al*. 2010, 2014; Ribera *et al*. 2010; Rizzo *et al*. 2013; Chen *et al*. 2021). They are small, predatory ground beetles (3–8 mm long) almost all of which lack eyes, are flightless, and are depigmented with long, slender bodies, elongated appendages, and sensory setae ('aphaenopsian'; Barr 2004; Ober *et al*. 2022). North American cave trechine beetles include 162 taxa in six genera distributed primarily in the Appalachians (APP), Interior Low Plateau (ILP), and Ozarks (OZK) karst regions of central and eastern North America (Barr 2004; Ober *et al*. 2022). The genus *Pseudanophthalmus* is exceptionally diverse with 155 described species and subspecies occurring throughout the APP and ILP and is arranged into 26 morphologically-defined species groups (Barr 2004); however, there may be more than 80 additional undescribed species in this genus, including one species in the OZK (Peck 1998; Barr 2004; Ober *et al.* 2022). Unlike *Pseudanophthalmus*, species richness in the other cave trechine genera is low, including *Ameroduvalius* (one species in the ILP), *Darlingtonea* (one species in the ILP), *Neaphaenops* (1 species in the ILP), *Nelsonites* (two species in the ILP), and *Xenotrechus* (two species in the eastern OZK). These five genera are thought to be closely

related to *Pseudanophthalmus* (Valentine 1952; Barr 1972, 1980, 1981, 1985b; Maddison *et al*. 2019), but phylogenetic relationships among them remain unclear. Moreover, all genera but *Xenotrechus* co-occur with *Pseudanophthalmus*. All but 19 described cave trechines in North America are at an elevated risk of extinction (NatureServe 2022), in part because most species have exceptionally small ranges <10,000 km^2 (*i.e.*, short-range endemism sensu Harvey 2002), and many are known from just one or a few cave systems (Culver *et al*. 2000, 2003; Barr 2004; Niemiller and Zigler 2013; Niemiller *et al*. 2017; Malabad *et al*. 2021; Harden *et al*. 2022).

Several hypotheses have been proposed related to the diversification and biogeography of temperate North American cave fauna (Barr and Holsinger 1985; Holsinger 2000), including cave trechines (recently reviewed in Ober *et al*. 2022). These hypotheses largely group into two contrasting (but not mutually exclusive) scenarios. Speciation in cave organisms is traditionally thought to occur at the surface-cave ecotone when subterranean populations diverge from related surface populations (Barr 1985a; Barr and Holsinger 1985; Holsinger 2000). Due to limited dispersal ability or significant barriers to dispersal, multiple, closely related subterranean species are the product of several independent subterranean colonization events from one or more surface ancestors followed by isolation and speciation without significant subterranean dispersal (*i.e.*, the multiple origins hypothesis). Speciation also could occur in subterranean habitats whereby a small number of subterranean colonization events by one or more surface ancestors followed by isolation and diversification but also with subsequent subterranean dispersal and diversification (*i.e.*, few origins hypothesis). Several recent studies have suggested that this mode of speciation may be more common than previously thought. The discovery of

monophyletic groups comprised of many subterranean lineages has been inferred as strong evidence for the role of subterranean speciation after a single or a few colonization events by surface ancestors (Holsinger 2000; Faille *et al*. 2010; Juan *et al*. 2010; Ribera *et al*. 2010). Subterranean speciation is generally thought to be the product of limited dispersal through subterranean corridors followed by isolation causing vicariance (Barr and Holsinger 1985; Holsinger 2000; Ribera *et al*. 2010). However, the evolutionary and biogeographic mechanisms factors that facilitate subterranean speciation are not well known or well-studied, as most investigations of speciation in cave organisms have focused on the morphological and evolutionary changes that accompany invasion and colonization from the surface (Holsinger 2000; Juan *et al*. 2010).

Determining the biogeographic and evolutionary history of a group of organisms is difficult when related sister lineages are either extinct or remain unsampled. For many groups of subterranean organisms, this is especially problematic as surface relatives are extinct, making distinguishing between single versus multiple colonization scenarios and distinguishing between modes of speciation (*i.e.*, surface-subterranean ecotone versus subterranean vicariance) impossible using phylogenetic evidence alone. Monophyly of many subterranean taxa may reflect multiple, independent subterranean colonization events by a single surface ancestor with little to no subsequent subterranean dispersal and speciation or a single subterranean colonization event with substantial subterranean dispersal and speciation. Except for *P. sylvaticus*, which is known from deep soil of a high elevation spruce forest in West Virginia (Barr 1967b), no obvious surface-dwelling sister group is known in eastern North America for North American cave trechines. *Pseudanophthalmus* and the other four troglobiotic genera belong to the *Trechoblemus*

series in the *Trechus* assemblage (Jeannel 1926, 1927, 1928, 1930, 1949). *Trechoblemus* is predominately a Eurasian genus of surface-dwelling and winged trechine beetles with a single species *T. westcotti* Barr,1972 known from the Willamette Valley of Oregon (Barr 1972), which is closest known surface relative to North American cave trechines (Ober *et al*. 2022). Although there are few studies that discuss the genetic diversity and phylogeny of cave-dwelling carabid beetles in North America (Gómez *et al*. 2016; Boyd *et al*. 2020), a comprehensive study investigating the phylogeny, divergence time, and biogeography of cave carabid beetles is not yet available. The evolutionary and colonization history of subterranean habitats of eastern North American trechine beetles is complex with many questions remaining unanswered. More research is required to understand the diversification, origin, molecular systematics, and biogeographic patterns of cave trechines in eastern North America.

The aim of this study is to generate the first molecular phylogenetic framework for the study of the origin, diversification, and biogeography of cave trechines in eastern North America using ultraconserved elements (UCEs) phylogenomics. Cave beetle distributions and speciation likely have been influenced by both intrinsic (*e.g.*, dispersal ability, body size, degree of subterranean adaptation) and extrinsic (*e.g.*, vicariant events and habitat connectivity) factors (Juan *et al*. 2010; Porter 2007). Intrinsic factors may be more relevant at smaller spatial scales, such as within and among cave systems, and occurring over ecological timescales (Porter 2007), whereas extrinsic factors may be more relevant at larger spatial scales, such as within karst regions, and longer evolutionary timescales. Our objectives were to: (1) reconstruct the phylogenetic relationships of cave trechines using UCEs to examine monophyly and relationships of morphological species groups; (2)

conduct divergence time analyses to determine relative timing of diversification particularly with respect to predictions of the climate-relict hypothesis; and (3) conduct biogeographic analyses to reconstruct colonization history and dispersal.

3.2 Materials and Methods

3.2.1 Study Area

Caves and subterranean fauna in the central and eastern United States are primarily associated with three major karst biogeographic regions (Culver and Hobbs 2002; Culver *et al*. 2003; Niemiller *et al*. 2019): the Appalachian Ridge and Valley, also called Appalachians (APP), the Interior Low Plateau (ILP), and the Ozarks (OZK) (Barr and Holsinger 1985; Culver *et al*. 2000; Figure 1). Rock strata in the ILP and OZK are mostly horizontally-bedded, whereas the rock layers in the APP were significantly faulted and folded due to past tectonic events associated with the uplift of the Appalachian Mountains. The ILP possesses the greatest number and density of caves, while the OZK is the largest in terms of surface area. The APP is ca. 37,000 km^2 extending from southeastern New York to northeastern Alabama and includes a series of parallel sandstone ridges with intervening carbonate valleys. The ILP is ca. 61,000 km^2 covering a large region west of the Cumberland Plateau from southern Indiana and Illinois southward through central Kentucky, central Tennessee, and northern Alabama (Ober *et al*. 2022). The OZK is ca. 129,500 km^2 covering southern Missouri and northern Arkansas and features flat to gently fold cherty sandstone and cherty dolomite (McKenney and Jacobson 1996). The APP and ILP karst regions are proximate to one another, and they come in contact near the junction of the borders of Tennessee, Alabama, and Georgia. Both karst regions contain prominent

karst exposures near the surface and support numerous caves (Culver *et al*. 2003; Weary and Doctor 2014). The ILP (400+ taxa) and APP (320+ taxa) possess the greatest troglobiotic species richness in the United States while OZK ranks fourth with 115+ species (Culver *et al*. 2003; Hobbs 2012; Niemiller *et al*. 2019).

3.2.2 Specimen Sampling

We collected specimens of 45 cave trechine taxa from caves in the APP, ILP, and OZK, including 41 species of *Pseudanophthalmus* from 20 of the 26 species groups defined by Barr (2004), as well as *Ameroduvalius jeanneli*, *Darlingtonea kentuckensis*, *Neaphaenops tellkampfi*, *and Nelsonites jonesi* (Supplemental Table 1; Figure 1). We were not able to obtain specimens of *Nelsonites walteri* nor both species of *Xenotrechus*. Beetles were collected from terrestrial and riparian cave habitats, such as mud banks, the splash zones of active drips, near streams and rimstone pools, underneath rocks and coarse woody debris, or within cobble and gravel. Two species of the genus *Trechus — T. obtusus* and *T. humbolti —* and *Trechoblemus westcotti* were included as outgroups (Supplemental Table 1).

3.2.3 DNA Extraction, UCE Library Preparation, and Sequencing

To generate a reduced representation genomic dataset of Ultraconserved Element (UCE) loci of cave trechines, we employed UCE phylogenomics (Faircloth *et al*. 2012; Branstetter *et al*. 2017), an approach that combines the targeted enrichment of thousands of nuclear UCE loci with multiplexed next-generation sequencing. Enrichment was performed using a published bait set that targets loci shared across all Coleoptera

('Coleoptera 1.1Kv1'; Faircloth 2017). It includes 13,674 unique baits targeting 1,172 UCE loci. Whole genomic DNA was extracted from muscle, meso-metathorax, and leg of the cave trechine specimens following the DNA extraction method from Maddison *et al*. 1999. The extracted DNA samples were shipped to Arbor Biosciences (Ann Arbor, Michigan, USA) for downstream processing which included library preparation, target enrichment, and Illumina high throughput sequencing.

3.2.4 Bioinformatics Processing

The raw demultiplexed fastq reads was cleaned and processed using the software package Phyluce v1.7.1 (Faircloth 2016) and associated programs. We used Illumiprocessor v2.0 (Faircloth 2013), which is a parallel wrapper of Trimmomatic v0.40 (Bolger *et al*. 2014) to clean and trim raw reads. All programs hereafter beginning with 'phyluce' are python programs part of the Phyluce package After trimming, we generated summary stats on the trimmed reads using the Phyluce script '*phyluce_assembly_get_fastq_lengths*'. We assembled the cleaned/trimmed reads using '*phyluce_assembly_assemblo_abyss*' with the AbySS assembler v2.3.5 (Simpson *et al*. 2009) using a kmer value of 60 on the CaveBio lab workstation. Next, we generated summary statistics (counts/lengths) of the assembled contigs using '*phyluce_assembly_get_fasta_lengths*'. To identify UCE loci, we used '*phyluce_assembly_match_contigs_to_probes*' that incorporates Lastz v1.0 (Harris 2007) to match the Coleoptera v1 probe set sequences to contig sequences with a minimum coverage of 75% and minimum identity of 75% and created a relational database of hits. We used '*phyluce_assembly_get_match_counts*' to create an initial database of loci counts

per taxon. Next, we used '*phyluce_assembly_get_fastas_from_match_counts*' to get a count of the number of UCE loci captured for each taxon. We then divided each UCE loci into a separate fasta file using '*phyluce_assembly_explode_get_fastas_file*', as we later wanted to construct individual gene tree phylogenies, and then aligned the sequences and trimmed the edges using '*phyluce_align_seqcap_align*' which implements MAFFT v7.130b (Katoh and Standley 2013). To remove poorly aligned regions, we trimmed internal and external regions of alignments using '*phyluce_align_get_gblocks_trimmed_alignments_from_untrimmed*' that incorporates Gblocks (Talavera & Castresana 2007), with reduced stringency parameters (b1:0.5, b2:0.5, b3:12, b4:7). Next, we generated summary statistics on the alignments using '*phyluce_align_get_align_summary_data*'. Lastly, we used '*phyluce_align_get_only_loci_with_min_taxa*' and '*phyluce_align_concatenate_alignments*' to filter and create two concatenated data sets or matrix by selecting aligned loci that contains at least 50% and 75% of the total taxa for phylogenetic inference.

3.2.5 UCE Phylogenomics

We performed two different types of phylogenomic analyses: (1) concatenated analyses using RaxML-HPC2 Workflow on XSEDE v8.2.12 (Stamatakis 2014) on the CIPRES Science Gateway v3.3 (https://www.phylo.org/) and MrBayes v3.2 (Ronquist *et al*. 2012) module of PhyloSuite v1.1.15 (Zhang *et al*. 2019) (2) species tree analyses using ASTRAL-II v5.7.8 (Mirarab and Warnow 2015) and SVDQuartets (Chifman and Kubatko 2014) implemented in PAUP* v4.0a169 (Swofford 1998). For all subsequent analyses, we used the 75% complete data-matrix.

For concatenated analyses, we defined each UCE locus as its own character set and then used PartitionFinder v2.1.1 (Lanfear *et al*. 2017) implementing the 'greedy' search algorithm (Lanfear *et al*. 2012) to select for the best partitioning strategy for the data under the General Time Reversible + Gamma (GTRGAMMA) site rate substitution model using the AICc metric (Burnham and Anderson 2002). We then conducted 20 maximum-likelihood (ML) searches in RaxML-HPC2 Workflow on XSEDE (Towns *et al*. 2014). We also performed non-parametric bootstrap replicates under GTRGAMMA using the autoMRE option to optimize the number of bootstrap replicates for this large dataset. We reconciled the bootstrap replicates with the best fitting ML tree. To confirm the reliability of the tree topology, the concatenated dataset was also processed through MrBayes for Bayesian phylogenetic reconstruction. We used the same partitioning strategy described above and estimated the most appropriate site rate substitution model for MrBayes using PartitionFinder v2.1.1. We conducted two independent runs of one cold chain and three heated chains (default settings) for 5,000,000 Markov chain Monte Carlo (MCMC) generations sampling every 100 generations in MrBayes. After dropping the first 25% 'burn-in' trees to ensure stationarity and examining the log-likelihood values for each Bayesian run using Tracer v1.7 (Rambaut *et al*. 2018), the remaining 37,500 sampled trees were used to estimate the consensus tree and the associated Bayesian posterior probabilities. A midpoint rooted ML and Bayesian trees with support values were generated using FigTree v1.4.4 (Rambaut 2010). To evaluate if allowing missing taxa on total alignment length affected the topology of the tree, the above-mentioned analyses were also conducted on the 50% complete data-matrix.

For species tree analyses, we first reconstructed the ML gene tree estimated for each of the UCE loci in IQTree v1.6.8 (Nguyen *et al*. 2015) module of PhyloSuite under the General Time Reversible + Gamma (GTRGAMMA) site rate substitution model. We conducted 1,000 iterations as well as 5,000 non-parametric ultrafast bootstrap replicates. We reconciled the bootstrap replicates with the best fitting ML tree of each locus. Second, these gene trees were input to ASTRAL-II to create a multispecies coalescent species tree and assessed support with 150 bootstrap replicates creating support values akin to posterior probabilities of nodes. Finally, we used another species tree method to look for congruence between methods. We created a multispecies coalescent species tree using SVDQuartets in PAUP* where we evaluated 100,000 random quartets using the Quartet FM (QFM) algorithm (Reaz *et al*. 2014). We assessed support with 200 bootstrap replicates. A midpoint rooted ASTRAL and SVDQuartets species trees with support values were generated using FigTree v.1.4.4.

3.2.6 Estimation of Divergence Times

To estimate the relative age of divergence of the cave trechine lineages, we used the Bayesian relaxed phylogenetic approach implemented in BEAST2 v2.6.7 (Drummond and Rambaut 2007), which allows for variation in substitution rates among branches (Drummond *et al*. 2006). We implemented a GTR+G model of DNA substitution with four rate categories for all partitions of the 75% complete data-matrix. We chose the uncorrelated lognormal relaxed molecular clock model to estimate the substitution rates and the Yule process of speciation as the tree prior. Because of the absence of robust closely related fossil records for trechine beetles, we used a secondary calibration method to date

the trees. Thus, we set the 'ucldMean' prior a lognormal distribution with average equal to a per-branch rate of 0.0012 substitutions/site/MY and a standard deviation of 0.059. All other parameters including the 'ucldStdev' prior was left with default settings. This rate is obtained from Faille *et al*. (2013) for the subterranean Trechini based on the colonization of the Alps in Europe (0.0010 and 0.0013 substitutions/site/MY for the nuclear small ribosomal unit 18S rRNA and large ribosomal unit 28S rRNA, respectively).

We performed two independent runs for 25,000,000 generations sampling every 500 generations in BEAST2. For both runs, we assessed convergence, likelihood, stationarity and verified effective sample size (ESS) values of each parameter using Tracer v1.7. After discarding a burn-in of 10%, results of both runs were combined with LogCombiner v2.6.7 and the consensus tree was compiled with TreeAnnotator v2.6.7 (Drummond and Rambaut 2007). A midpoint rooted tree with geological time scale depicting the node ages, 95% highest posterior density interval (HPD) bars, and posterior probabilities was generated using the R strap package (R Development Core Team, 2014) using a custom R script.

3.2.7 Ancestral Range Estimation

We estimated ancestral ranges using stochastic likelihood-based models of geographic range evolution implemented in the R package 'BioGeoBEARS 0.2.1' (Matzke 2013). We executed and compared the standard dispersal-extinction-cladogenesis (DEC) model with the (DEC+J) model (Ree *et al*. 2005; Ree and Smith 2008), which includes an additional parameter j that allows for founder-event speciation by jump dispersal (Matzke 2014). The 'j' parameter allows for a daughter lineage to immediately occupy via long-

distance dispersal a new area that is different from the parental lineage. We used the maximum clade credibility time-calibrated tree from the concatenated BEAST2 analysis as the input tree. After pruning the outgroup species, we assigned each species in the tree the biogeographic major karst regions (represented by single letter code) spanned by the represented lineage: Appalachian Ridge and Valley (A), the Interior Low Plateau (I), and the Ozark Highlands (O). In addition, to have a closer look at the effect of different karst subregions on the represented lineage, we conducted an additional analysis by assigning each species in the tree to biogeographic karst subregions: Ridge and Valley (R), Wills Valley (V), Greenbrier Karst (G), Inner Bluegrass (I), Western Pennyroyal (W), Highland Rim (H), Pine Mountain (P), Cumberland Plateau (C), Nashville Basin (N), Outer Bluegrass (O), Sequatchie Valley (S), and Ozarks (Z). Each species was coded as being present or absent in each of these areas and the maximum number of areas occupied by a single species was set to 4. The models were compared to each other using two different methods: (1) the likelihood of each model was compared with Akaike Information Criterion (AIC); and (2) the nested models were compared with each other using a chi-squared test to determine if the model with 'j' parameter was preferred. Biogeographic stochastic mapping (Dupin *et al*. 2017) using the best-fitting biogeographic model (DEC+J) was plotted on the cave trechine chronogram, and the number of dispersal events among different karst regions was assessed.

3.3 Results

Information on specimen vouchers, DNA quantities, raw Illumina reads before and after quality filtering and trimming, and SRA accession numbers can be found in

Supplemental Table 1. After filtering and trimming, our aligned UCE contigs included 3,240,300 characters, of which 2,885,448 were nucleotides and 354,852 (11.0%) were missing data. Mean locus length was 224 nucleotides (range: 42–798 bp). Each locus contained an average of 12 informative characters (range: 0–95). We analyzed both 50% and 75% coverage data matrices. The 75% UCE data matrix comprised 16,794 base pairs (bp) and 68 UCE loci for 48 specimens (45 cave trechine beetles and three outgroup taxa). The 50% UCE data matrix comprised 65,376 base pairs (bp) and 274 loci for the 48 specimens.

Raw Illumina sequence reads are available at the NCBI Short-Read Archive (BioProject PRJNA894729; https://www.ncbi.nlm.nih.gov/bioproject/PRJNA894729) while aligned data matrices and tree files are available at Dryad (doi:10.5061/dryad.sj3tx967w).

3.3.1 Concatenated Phylogenetic Analyses

Both ML and Bayesian inference (BI) trees inferred from the 75% concatenated UCE data matrix resolved similar topologies with high support for most nodes (Figures 2 & 3). Likewise, ML and Bayesian inferred from the 50% concatenated UCE data matrix (Supplemental Figures 1 & 2) were highly congruent with trees inferred from the 75% concatenated UCE data matrix; therefore, we focus our discussion based on the results of concatenated (and species tree analyses in the next section) of the 75% UCE data matrix. Of the 15 *Pseudanophthalmus* species groups as defined by Barr (2004) for which more than one taxon was sampled, 12 were monophyletic. Four species groups were paraphyletic, including the *cumberlandus*, *grandis*, *jonesi*, and *simplex* species groups.

Moreover, the genus *Pseudanophthalmus* was not recovered as monophyletic as all of the other cave trechine genera included in the analyses (*i.e.*, *Ameroduvialis*, *Darlingtonea*, *Neaphaenops*, and *Nelsonites*) were nested within *Pseudanophthalmus*.

Four primary clades were recovered in both ML and BI analyses (Clades A, B, C, and D in Figures 2 & 3, Supplemental Figures 1 & 2). *Pseudanophthalmus* formed two main clades that largely corresponds with the karst biogeographic regions: Appalachian Ridge and Valley and Interior Low Plateau. The genera *Ameroduvalius*, *Darlingtonea*, *Neaphaenops*, and *Nelsonites* were nested within the Interior Low Plateau *Pseudanophthalmus* clade. We found strong support for a clade with the geographically isolated undescribed *Pseudanophthalmus* species sister to the *P.petrunkevitchi* species group of the APP (Clade A). Clade B includes seven *Pseudanophthalmus* species groups (*engelhardti*, *hubrichti*, *tennesseensis, hirsutus, jonesi, alabamae,* and *hypolithos*) along with one species from the *grandis* species group (*P. virginicus*). The *P. hubbardi* and *pusio* species groups along with the other species from the *grandis* species group (*P. grandis*), all from the APP, form Clade C. Finally, Clade D is comprised of taxa from the ILP, including nine *Pseudanophthalmus* species groups (*tenuis, intermedius, robustus, menetriesi, pubescens, horni, barri, simplex,* and *cumberlandus*) as well as the genera *Ameroduvalius*, *Darlingtonea*, *Neaphaenops*, and *Nelsonites*. The weakly supported sister group relationships of the *P. barberi* and *P. tenuis* clade within Clade D was the only difference between the BI and the ML phylogenetic trees.

3.3.2 Species Tree Analyses

The species tree inferred from the 75% UCE data matrix using ASTRAL-II differed slightly from concatenated ML and BI trees. Differences were not strongly supported, however, and generally involved the placement of *P. thomasi*, *P. fridigus*, *N. telkampfi viator*, *P. troglodytes*, and *D. kentuckensis* within primary clades (Figure 4). In general, support for several deeper nodes in the ASTRAL species tree was relatively low. The species tree reconstructed using SVDQuartets (Supplemental Figure 3) was similar in topology to the concatenated ML and BI trees (Figures 2 & 3). However, several deeper nodes in the SVDQuartets species tree were weakly supported.

3.3.3 Divergence Time Estimation

Divergence time estimation in BEAST2 using a molecular clock rate recovered a crown age for North American trechine beetles of 11.5 Mya (95% highest posterior density [HPD] interval: 9.8–13.2 Mya) in the middle Miocene at the beginning of the Tortonian (Figure 5 and Supplemental Figure 4) forming the main ARV and ILP clades. Additional diversification of primary clades with the ARV and ILP clades occurred shortly thereafter at 9.5 Mya (95% HPD: 7.7–11.4 Mya) and 10.7 Mya (95% HPD: 9.1–12.4 Mya), respectively. Diversification within the primary clades occurred primarily from the late Miocene into the Pliocene with estimated species divergences all slightly predated the Pliocene (*i.e.*, 5 Mya). The divergence events estimated to have occurred in the Pleistocene were limited primarily to species pairs within species groups (Figure 5 & Supplemental Figure 4). Diversification during the Pliocene and Pleistocene were almost exclusively

within the same geographical area and involved species of the same morphological species group.

3.3.4 Ancestral Range Reconstruction

The DEC+J model provided a significantly better fit for the phylogeny resolved using BEAST compared to other karst region models (Table 2). The earliest divergence in cave trechine beetles was estimated to occur ~11.5 Mya in a widespread common ancestor most likely occurring in Appalachian Ridge and Valley with dispersal into the Interior Low Plateau between 9.0–12.5 Mya (Figure 5). At least three additional colonization events from the Appalachians Ridge and Valley into the Interior Low Plateau are supported and are estimated to have occurred 2.0–5.5 Mya. The model also supports a single dispersal event from the Appalachian Ridge and Valley into the Ozark Highlands (Figure 6).

With respect to karst subregions, the DEC + J model again was the best model (AIC = 151.9; Supplemental Table 1). This model also supported a common ancestor in the Appalachian Ridge and Valley 11.5 Mya (95% HPD: 9.5–13.5 Mya) during the middle Miocene (Figure 6). The *petrunkevitchi* species group was the first group to colonize the Appalachian Ridge and valley karst and its sister lineage later dispersed into the Ozark Highlands. The main radiation in the Interior Low Plateau appears to have occurred from east to west first into the Cumberland Plateau then into more western ILP karst subregions, including the Highland Rim, Western Pennyroyal, Nashville Basin, Inner Bluegrass, and Outer Bluegrass. The general pattern of diversification within the Ridge and Valley clade suggests dispersal from north to south with occasional dispersal into adjacent ILP karst subregions, including the Cumberland Plateau, Pine Mountain, and Sequatchie Valley.

3.4 Discussion

This study provided the first phylogenomic framework for the North American cave trechine beetles improving upon previous studies with expansive taxon and data sampling, resulting in a robust hypothesis of evolutionary relationships. Our results based on phylogenetic analyses of 68 genomic UCE loci shed new light on the systematics of carabid beetles and highlight potential areas for future taxonomic research. The study also presents a comprehensive analysis of biogeographic history for these unique beetles for the first time.

3.4.1 Carabidae UCEs

Concatenated analyses in this study yielded more well-resolved trees than coalescent-based species tree methods (*i.e.*, ASTRAL and SVDQuartets). Lower support values for some deeper branches in coalescent-based species tree approaches may be due to gene-tree discordance caused by incomplete lineage sorting (Maddison 1997; Edwards 2009), lack of phylogenetic signal among gene trees, and missing data (Thomson *et al*. 2008; Edwards 2009; Gatesy and Springer 2014; Springer and Gatesy 2014, 2016; Xi *et al*. 2015; Edwards *et al*. 2016; Meiklejohn *et al*. 2016; Moyle *et al*. 2016). Previous studies have demonstrated that species tree methods including ASTRAL and SVDQuartets can be prone to errors in gene tree estimation (Edwards *et al*. 2016; Hosner *et al*. 2016; Meiklejohn *et al*. 2016; Springer and Gatesy 2016). Hosner *et al*. (2016) showed that missing data can also cause significant errors in species tree estimation, especially when including taxa with only partially captured UCE loci. The length of a given UCE locus obtained via targeted capture methods may vary greatly among samples (Hosner *et al*. 2016; Streicher *et al*.

2016). In our dataset, the 50% and 75% complete dataset contained considerably fewer UCEs than other published beetle UCE datasets (Baca *et al*. 2017; Van Dam *et al*. 2017; Gustafson *et al*. 2020; Bradford *et al*. 2022; Sota *et al*. 2022) suggesting the potential for the missing data to affect our coalescent analyses. Other several studies have shown that there is a positive trade-off in constructing a larger data matrix by allowing the inclusion of loci with missing taxa when conducting concatenated analyses (Hosner *et al*. 2016; Streicher *et al*. 2016). Therefore, the amount of missing data in our analyses may possibly account for the differences in resolution between our concatenated and coalescent analyses. No genomes for carabid beetles were publicly available for use during the development of the Coleoptera bait set (Faircloth 2017), which may account for the lower recovery of UCEs in our study.

3.4.2 Evolutionary History and Biogeography of North American Cave Trechine Beetles

Distinguishing between the two hypotheses (multiple vs. few origins) and the continuum of colonization, dispersal, and speciation scenarios in between is not trivial and often requires other sources of data, such as geological, climatical, or paleontological evidence, to reconstruct phylogeographic histories of cave organisms. Evidence in support of either scenario in North American cave trechines from past studies is limited. Barr (2004) proposed that varying levels of troglomorphy observed among North American cave trechine beetles were indicative of multiple episodes of cave colonization, with the assumption that degree of troglomorphy reflects time since cave colonization by a lineage. However, this assumption remains to be explicitly tested in North American cave trechines

and most other cave taxa. Many previous authors have favored the multiple origins hypothesis whereby once populations became restricted to subterranean habitats, subterranean dispersal was extremely limited except in areas with more expansive karst and interconnected subterranean habitats (Krekeler 1959; Barr 1967, 1985a; Barr and Holsinger 1985). Moreover, many caves in the APP and ILP contain two or more (up to six in the Mammoth Cave System) species of cave trechine beetles belonging to different species groups, supporting multiple cave colonization events (Niemiller *et al*. 2021; Ober *et al*. 2022).

As an extension of the multiple origins hypothesis, Barr (1981, 1985, 2004) incorporated a temporal context for cave colonization whereby ancestral surface species occurring in or near karst areas adopted a deep soil existence during glacial periods of the Pleistocene or earlier followed by colonization events into caves during the interglacial periods of the Pleistocene in response to warming and drying of surface habitats, *i.e.*, the climate-relict hypothesis (Jeannel 1943; Holsinger 1988, 2000; Ashmole 1993). Jeannel (1926, 1927, 1928, 1930) and Barr (1967) also hypothesized climate change after the last glacial maximum during the Pleistocene as a major driver of cave colonization and diversification in cave trechine beetles. However, given the large radiation (>150 taxa) of cave trechine species in North America, Ober *et al*. (2022) did not believe such a recent timeframe for cave colonization and diversification plausible; rather the authors hypothesized multiple colonization events occurring over several interglacial periods occurring as early as the late Pliocene when many caves in the APP and ILP were forming (Poulson and White 1969; Clark 2001; Shofner *et al*. 2001) to account for cave trechine diversity observed today. Moreover, the presence of two or more cave trechine species

occurring in a single cave or cave system (*e.g.*, the Mammoth Cave assemblage contains six cave trechine species; Niemiller *et al*. 2021) suggests multiple cave colonization events occurred over multiple interglacial periods.

Although taxonomic sampling is incomplete, our analyses support a multiple origin scenario that predates the Pleistocene. Diversification of North American cave trechines began in the middle Miocene with several species groups present by the end of the Miocene and further diversification into the Pliocene, rejecting the Pleistocene climate-relict hypothesis (Jeannel 1943; Barr 1967; Holsinger 1988, 2000; Ashmole 1993) as the primary driver of diversification. Under the climate-relict hypothesis, we might expect simultaneous independent cave colonization events by a single widespread species or multiple closely related species reflected as a polytomy on the inferred species tree and burst of accumulation of lineages on a lineage-through-time plot. In North American cave trechines, we instead see an increase of diversification several million years earlier in the late Miocene into the early Pliocene (Figure 5). Our results are consistent with other recent studies of cave-dwelling beetles. Faille *et al*. (2011) estimated that a lineage of European trechines colonized caves about 10 Mya, while Ribera *et al*. (2010) estimated that major lineages of Western Mediterranean cave beetles in the family Leiodidae diverged about 30 Mya. Among other North American cave beetle taxa, diversification events of cave carabid beetles of the genus *Rhadine* (tribe Platynini) in the Edwards Plateau and Balcones Escarpment karst region of Texas occurred within the past 4–5 million years (Gómez *et al*. 2016) coinciding with a period of cave development in the Balcones Escarpment (Ward 2006; White *et al*. 2009). Leray *et al*. (2019) examined diversification of the *hirtus*-group of the small carrion beetle genus *Ptomaphagus* (family Leiodidae), which consists of 19

cave and soil-dwelling species in the central and southeastern United States and co-occurs with trechine cave beetles in many cave systems in the ILP and southern APP. Two main periods of diversification in troglobiotic *Ptomaphagus* were identified: 1) seven geographically distinct lineages diverged across ILP and southern APP 6–8.5 Mya; and 2) a lineage in the southern Cumberland Plateau of Alabama and Tennessee diversified into 12 species over the last 6 million years. Estimated dates of diversification in *Ptomphagus* are quite similar to those in cave trechines. Although significant diversification predates the Pleistocene Epoch, Pleistocene glacial cycles likely have had important impacts on the evolutionary history and biogeography of North American cave trechines.

The strong palaeogeographical signal in the distribution of the cave trechine species is likely to be related to their ecological habitats. There are likely to have been multiple independent colonizations, with each lineage having different degrees of morphological adaptation to the subterranean environment. The strongest factor driving the subterranean colonizations may have been the aridification of the climate since the late Miocene (Krijgsman *et al*. 2000; Micheels *et al*. 2009). The multiple origins of the subterranean populations or species within the cave trechine group confirmed both by our phylogenetic results is in contrast to hypotheses proposed for other radiations of subterranean beetles, for example in the Pyrenees (Faille *et al*. 2010, 2013; Ribera *et al*. 2010; Rizzo *et al*. 2013; Cieslak *et al*. 2014), where the entire lineages of species are found exclusively within the deep subterranean environment without morphological variation in some troglomorphic characters (Jeannel 1924, 1928; Salgado *et al*. 2008). This is likely due to the unique combination of widespread ancestral epigean species with multiple colonizations of the

subterranean environment, giving rise to troglobiotic species with very limited geographical ranges that have persisted unaltered over long evolutionary periods.

The hypothesized limited dispersal ability of North American cave trechine beetles (Barr 1967a). Barr 1981; Barr 1985a) and results of this study suggest that troglomorphy evolved in this group multiple times through convergent evolution. In cave trechine evolution, a single cave colonization event would result in a molecular signature of shared loss-of-function (LOF) mutations particularly in loci involved in the regression of eyes and pigmentation among cave trechine lineages (assuming lack of strong selection on particular LOF mutations). In contrast, shared LOF mutations in eye and pigmentation loci are not expected among the cave trechine lineages under a multiple independent cave colonization scenario. However, genomic analyses will be required to test these hypotheses and determine if identical LOF mutations occur in geographically separated cave trechine lineages (single cave invasion subsequently followed by long-distance dispersal) or different LOF mutations occur in geographically distinct lineages (multiple cave invasions subsequently followed by isolation and vicariance). Previous studies supporting multiple independent cave colonization hypothesis include subterranean diving water beetles with distinct mutations in pigmentation and eye opsin genes (Leys *et al*. 2005; Tierney *et al*. 2015) and also in the eye rhodopsin locus of geographically distinct lineages of amblyopsid cavefishes in eastern North America (Niemiller *et al*. 2013).

A single cave colonization scenario requires cave trechines to be able to disperse several kilometers across non-karst terrain that now separates the distributions of many species and species groups. Although dispersal remains to be rigorously examined in cave trechines, it is thought their dispersal ability is quite limited. All cave trechine species are

small and wingless, and all but one species has been observed in surface habitats (Ober *et al*. 2022). Carbonate strata in APP is patchy and discontinuous, and caves are generally smaller and more isolated within this fragmented karst (Hack 1969; Barr 1967a; Barr 1981; Culver 1982; Barr 1985a). Limestone valleys are separated by sandstone ridges in the faulted and folded strata limiting dispersal of cave trechines (Barr 1985). Consequently, most *Pseudanophthalmus* species in the APP are frequently limited to a single or few isolated cave systems (Barr 1965, 1981, 2004; Malabad *et al*. 2021). For example, 22 of the 64 described *Pseudanophthalmus* species from the APP are single-cave endemics. While evidence suggests a multiple cave colonization scenario is more likely with limited subterranean dispersal, long-distance dispersal cannot be completely ruled out, as long-distance dispersal has been hypothesized for troglobiotic leiodids in central Pyrenees (Rizzo *et al*. 2013).

In contrast to the APP, only 12 of 84 species in the ILP are endemic to single caves and species generally have larger distributions (Ober *et al*. 2022). For example, *Darlingtonea kentuckensis* is distributed over an area of ca. 3,728 km^2 along the western escarpment of the Cumberland Plateau in southeastern Kentucky and adjacent northern Tennessee. Karst in the ILP is characterized by expansive exposures of highly soluble, horizontal-bedded limestones with numerous sinuous branch-like caves systems with four major subregions (Barr 1967): (1) the Bluegrass Region in Kentucky; (2) the western escarpment of the Cumberland Plateau from northeastern Kentucky to northern Alabama; (3) the Central Basin in Tennessee; and (4) the Pennyroyal Plateau from southern Indiana near Bloomington south extending westward into Kentucky and north-central Tennessee. Among these subregions, escarpments of the Pennyroyal Plateau, Central Basin, and the

western escarpment of the Cumberland Plateau possess numerous caves that may facilitate dispersal of terrestrial troglobionts, including cave trechines (Barr 1985a). Cave systems in the ILP are more highly connected than in the APP apart from the Bluegrass Region in northern Kentucky near the Ohio River, which is comprised of smaller and more isolated caves that may limit subterranean dispersal in this region (Barr 1967). Cave trechine species are more likely to occur in sympatry in the ILP than in the APP, which may reflect different species colonizing subterranean habitats during different time periods and subsequently dispersing through the highly connected karst of the ILP (Barr 1967).

Although there is evidence for subterranean dispersal in the ILP, vicariance appears to have had a significant role in the diversification and shaping the distributions of many species and species groups in not only the ILP but also the APP. Both hydrological and geological barriers separate both species and species groups in the ILP. For example, within the *P. tenuis* species group *P. barberi* in northern Kentucky is separated from the other five species in the species group by the Ohio River (Barr 2004; Ober *et al*. 2022). Likewise, two species of the *P. barri* group, *P. barri* and *P. troglodytes*, occur on opposite sides of the Ohio River (Barr 2004). The Ohio River, which formed 0.8 Mya thousand years ago and has down cut via erosion through the cave-bearing strata (Gray 1991; Teller and Goldthwait 1991), appears to be a significant barrier to dispersal for not only troglobionts but also stygobionts (Niemiller *et al*. 2013). Similarly, the Cumberland River separates the two species of *Nelsonites* with *N. jonesi* to the north and *N. walteri* to the south of the river (Barr 1985a). In contrast, smaller tributaries may not be substantial barriers for all cave trechines, as some species, such as *Neaphaenops tellkampfii* along the Green River in Kentucky and *P. ciliaris* and *P. loganensis* along the Red River in Kentucky and

Tennessee, occur on both sides of smaller river and streams (Barr and Peck 1965; Ober *et al*. 2022). Barr and Peck (1965) hypothesized that flooding may wash beetles out of caves and transporting them downstream to other cave systems on either bank. Barr (1985a) offered support for this hypothesis by examining the relationship between the meander frequencies of rivers and the frequency of cave trechine species occurring on opposite sides of a river. He concluded that the higher the river meander frequency, the more often troglobionts occurred on both sides of the river, suggestive of passive dispersal via flooding.

Barr and Holsinger (1985) hypothesized that dispersal and gene flow between populations cave-dwelling species, particularly troglobionts, can be reduced and ultimately isolated by erosion (*i.e.*, downcutting) of surface water courses into and through cave-bearing strata during the Pliocene and Pleistocene. This vicariance-by-erosion model (sensu Leray *et al*. 2019) may be a particularly attractive hypothesis for explaining diversification and distribution patterns among taxa within species groups along the Cumberland Plateau and other prominent escarpments of the ILP and also in the APP whereby incisional history and hydrological drainage reorganizations may have influenced the evolutionary history of troglobionts. For example, Leray *et al*. (2019) propose that the vicariance by erosion model best explained diversification and distributional patterns of *Ptomaphagus* cave fungus beetles, which cooccur with *Pseudanophthalmus*, in the highly dissected southern Cumberland Plateau. The earliest *Ptomaphagus* lineages to divergence in the southern Cumberland Plateau are found in the most peripheral and isolated escarpments, whereas lineages that diverged later are found toward the central region of the southern Cumberland Plateau, a pattern consistent with the predictions of the model.

The biogeography of cave trechine beetles in North America highlights the complex evolution of the eastern North America cave biodiversity hotspots. While our understanding of the evolutionary history and biogeography of this diverse assemblage of cave beetles is incomplete due to incomplete taxonomic sampling, we briefly summarize the biogeography and evolutionary history of eastern Northern American cave trechines based on evidence to date. The surface ancestor(s) to cave trechines appears to have colonized the eastern North America karst regions from an east to west manner during the middle Miocene and Pliocene. The ancestral surface origin of the cave trechine beetles is uncertain but appears to have been in the southern Appalachians. Based on the concatenated phylogeny, the earliest ancestor was likely distributed in the APP and dispersed into the ILP about 11.5 Mya. A surface ancestor likely dispersed into the Ozarks Highlands from the APP between 4.1–7.2 Mya. The APP appears to have served both as a cradle for in situ diversification and as bridge linking the southern Appalachians and ILP, enabling the dispersal and subsequent diversification of these cave beetles. In the APP, there was a burst of diversification in the early Pliocene 7.0 Mya as well as in the Pleistocene 1.2 – 3.7 Mya. After colonization into the ILP, there was further diversification around 10.7 Mya followed by a burst of diversification in the late Miocene 8.0 Mya as well as in the Pliocene into the Pleistocene 1.2 – 5.7 Mya. Based on the dating of radioactive cave sediments, the oldest caves along the western escarpment of the Cumberland Plateau in the ILP are estimated to be 3.5–5.7 million years old (Sasowsky *et al*. 1995; Anthony and Granger 2004, 2007); however, cave development may have begun much earlier (White 2009).

3.4.3 Systematics of North American Cave Trechine Beetles

Large (7 mm), slender, morphologically similar species, *Neaphaenops tellkampfii*, *Darlingtonea kentuckensis*, *Ameroduvalius jeanneli*, and *Nelsonites walteri* have been thought to be closely related to *Pseudanophthalmus* species (Valentine 1952; Barr 1972, 1980, 1981, 1985b; Maddison *et al*. 2019). However, the phylogenetic relationships among these five genera have been unclear due to limited sampling of species within the genus *Pseudanophthalmus*. We found that *Pseudanophthalmus* as currently recognized is paraphyletic with respect to the four other genera (*Neaphanops*, *Darlingtonea*, *Nelsonites*, and *Ameroduvalius*) suggesting that these genera are derived from *Pseudanophthalmus*. The relatively widespread distribution of *Neaphaenops tellkampfii* includes the Pennyroyal Plateau in Kentucky and the distributions of the other three genera includes the western escarpment of the Cumberland Plateau in the ILP of eastern Kentucky and north-central Tennessee. The distributions of all four genera overlap with *Pseudanophthalmus*, and they often co-occur within the same cave systems (Ober *et al*. 2022). Contrastingly, *Xenotrechus*, which is known from a few caves in southeastern Missouri west of the Mississippi River (Barr and Krekeler 1967) is believed to be distantly related to *Pseudanophthalmus* and trechine genera. We were unable to include any *Xenotrechus* species in our study. Although the phylogenetic placement of genus *Xenotrechus* is unknown, *Xenotrechus* is hypothesized to be related to the eastern European genera *Chaetoduvalius* and *Geotrechus* lacking any close relatives in North America (Barr and Krekeler 1967). At present, we assume the exclusion of *Xenotrechus* has minimal impacts on our biogeographic interpretations.

While the phylogeny of cave trechines in eastern North America is incomplete, Barr's (2004) species group classification arrangement provides a framework for how several species of cave trechines may be related. Barr categorized species into species groups based on shared morphological characters, such as the shape of a groove at the apex of the elytra and features of the male genitalia, and also their distributions in karst regions. Barr (1981, 1985a) hypothesized that closely related species morphologically generally co-occur in the same geographical area. Our results showed that apart from the *cumberlandus*, *grandis*, *jonesi*, and *simplex* species groups, all other species groups proposed by Barr (2004) formed monophyletic groups. For the species groups that were not recovered as monophyletic, morphological similarity may reflect morphological convergence and cryptic speciation, which is frequently reported in phylogenetic studies of cave organisms (Wiens *et al*. 2003; Derkarabetian *et al*. 2010; Niemiller *et al*. 2012; Maddison *et al*. 2019). However, additional and more comprehensive taxonomic sampling is warranted.

3.5 Conclusion

This study represents the first effort to establish a time-calibrated phylogenomic framework for cave trechine beetles in North America, elucidating a rich and intriguing history of evolution. Our results conflicted with previous generic and many species-group taxonomic classification hypotheses based on morphology. In particular, the genera *Neaphanops*, *Darlingtonea*, *Nelsonites*, and *Ameroduvalius* were nested within *Pseudanopthalmus* and some species groups were not recovered as monophyletic. The surface ancestor of cave trechines likely began colonizing caves in the Appalachians Ridge and Valley in the middle Miocene around 11.5 Mya. The evolution of *Pseudonaphthalmus* is characterized by rapid early radiation followed by a series of dispersal events into the

Ozark Highlands and Interior Low Plateau, with many of the major clades attaining their present-day geographic distributions by the early Miocene and with multiple additional episodes of cave colonization and diversification occurring throughout the Pliocene and Pleistocene. Additional research is needed to better characterize the levels of diversity, speciation, and the origins of cave trechines and to understand their phylogeographic patterns in eastern North America. In summary, molecular systematics and biogeography of these unique cave beetles offer a model for other comparative evolutionary and ecological studies of troglobionts to further our understanding of factors driving speciation and biogeographic patterns.

GENERAL CONCLUSIONS

CHAPTER 1

The objective of this chapter was to assemble and characterize the complete mitogenomes of four species of groundwater amphipods, including *Stygobromus pizzinii*, *Stygobromus allegheniensis*, *Stygobromus tenuis potomacus,* and *Bactrurus brachycaudus*, as well as a surface spring-dwelling species, *Crangonyx forbesi*. I used the NOVOPlasty assembler to de-novo assemble the mitogenomes from the quality assessed and trimmed raw sequence reads. Webservers including MITOS, MiTFi, and NCBI ORFfinder were used to annotate and confirm the start and stop positions of the protein-coding genes, transfer RNAs, ribosomal RNAs, and control region. I also provided a comparative analysis of structure and gene order of the newly sequenced mitogenomes and also provided the first intraspecific comparison of two different population of *S. tenuis potomacus*. Relationships among the newly sequenced mitogenomes and 18 previously published amphipod mitogenomes were inferred using Bayesian inference in MrBayes module implemented in PhyloSuite.

All mitogenomes contained 13 protein-coding genes, 22 tRNA genes, one small ribosomal RNA (*rrnS*) gene, and one large ribosomal RNA (*rrnL*) gene, and a non-coding control region representative of the Kingdom Animalia. Both *C. forbesi* and *S. allegheniensis* had more intergenic spacers than *B. brachycaudus*, *S. pizzinii*, or *S. tenuis potomacus*. Gene order in *S. tenuis potomacus*, *S. pizzinii*, *S. allegheniensis*, and *B. brachycaudus* was almost identical to the ancestral pancrustacean gene order except for the transposition of a few tRNA genes. However, *C. forbesi* exhibited quite distinctive gene

rearrangements. ATT, ATC, ATG, and ATA were the most frequently used start codons for most protein-coding genes. However, a few unconventional start codons were also used by protein-coding genes of a few species. Similarly, most of the protein-coding genes used TAA or TAG stop codons; however, few genes used an incomplete TA– or T– – stop codon. Variation in the length of protein-coding genes and overlap between some adjacent protein-coding genes were observed among the five new crangonyctid mitogenomes. Most of the tRNAs had ideal cloverleaf secondary structures. However, tRNAs of *B. brachycaudus* and *S. allegheniensis* displayed unique differences. *C. forbesi* mitogenome revealed the presence of a long sequence overhang on the 5' end of the *rrnL* and *rrnS* genes and displayed an unique transposition of the genes. Comparison between the two *S. tenuis potomacus* mitogenomes revealed few differences in their start codons, stop codons, and lengths of the non-coding control region and *nad3* locus.

Bayesian phylogeny revealed members of Crangonyctidae (*Bactrurus*, *Crangonyx*, and *Stygobromus*) formed a well-supported clade. However, the genus *Stygobromus* was not monophyletic, as *B. brachycaudus* and *C. forbesi* were nested within *Stygobromus*. *Crangonyx forbesi* was found as sister to all other crangonyctids, although support for this relationship was lower. Our newly sequenced five crangonyctid mitogenomes are valuable for inferring the phylogenetic relationships, biogeography, and trait evolution of amphipods and investigating mitogenome evolution.

CHAPTER 2

The objective of this chapter was to test whether the mitochondrial PCGs showed evidence of adaptive evolution in subterranean environments in amphipods and how that

differ from surface counterparts. I compared base composition, codon usage, gene order rearrangement, conducted comparative mitogenomic and phylogenomic analyses, and examined evolutionary signals of 35 amphipod mitogenomes representing 13 families, with an emphasis on Crangonyctidae. I calculated nucleotide composition, amino acid frequencies, and codon usage using PhyloSuite. I used the web-based program CREx to perform pair-wise comparison of the gene orders in the mitogenome to determine rearrangement events and visualized using iTOL. Phylogenetic relationship of the 35 amphipod mitogenomes and three isopod mitogenomes using the concatenated 13 PCG alignment was determined using Bayesian inference in MrBayes module implemented in PhyloSuite. I performed base-substitution analyses (non-synonymous to synonymous rate ratio; dN/dS or ω) on entire mitogenomes as well as for each of the 13 PCGs individually to compare surface versus subterranean amphipod taxa using the free-ratio, one-ratio, two-ratios branch model, site model, and branch-site model using the EasyCodeML program implemented in PAML. In addition, I performed selection pressure analyses on the concatenated 13 PCGs dataset as well as on each PCG using several approaches (aBSREL, BUSTED, RELAX) available in the Datamonkey webserver.

Mitogenome sizes ranged from 14,113 to 18,424 bp for all amphipods. Mean mitogenome size of surface amphipods was significantly higher than that of the subterranean amphipods. Mitogenome AT% in all amphipods ranged from 62.2 to 76.9%. Mean AT% of the subterranean amphipods was significantly higher than that of the surface amphipods. Mitogenome AT-skew in all amphipods ranged from -0.062 to -0.037. Mean AT-skew of the surface amphipods was positive and slightly higher than that of the subterranean amphipods. Mitogenome GC-skew in all amphipods ranged from -0.431-

0.120. Mean GC-skew of the subterranean amphipods was negative and significantly higher than that of the surface amphipods. Among crangonyctid amphipods, the spring-dwelling *Crangonyx forbesi* exhibited a unique gene order, a long *nad5* locus, longer *rrnL* and *rrnS* loci, and unconventional start codons. This was further confirmed by CREx analysis. Similar to *C. forbesi*, other surface amphipods including *Gmelinoides fasciatus*, *Onisimus nanseni*, *Gondogeneia antarctica*, *Platorchestia parapacifica*, *P. japonica*, *Pallaseopsis kessleri*, *Caprella scaura*, and *C. mutica* exhibited a high to moderate rearranged gene order. Interestingly, a subterranean amphipod *Pseudoniphargus daviui* also exhibited a moderate rearranged gene order. Among crangonyctid amphipod mitogenomes, relative synonymous codon usage (RSCU) values showed a high prevalence of A or T nucleotides at third codon positions. In PCGs, the second copy of leucine and cysteine are the most and the least used amino acids, respectively. The length of *rrnL* genes in all amphipods ranged 604–1,137 bp and that of *rrnS* genes ranged 196–1,631 bp. The length of *rrnL* of the subterranean amphipods was significantly higher than that of the surface amphipods, whereas the length of *rrnS* of the surface amphipods was slightly higher (but not significant) than that of the subterranean amphipods. In all crangonyctid mitogenomes except *C. forbesi*, the putative control region (CR) was identified between the *rrnS* and *nad2* gene cluster. The same pattern was observed in mitogenomes of non-crangonyctid amphipods including *G. duebeni*, *O. nanseni*, *G. antarctica*, *P. daviui*, and for the pancrustacean ground pattern. Contrastingly, in *C. forbesi*, the CR was observed between the *nad1* and *nad2* gene cluster and separated by few intergenic spacers. Similar to chapter 1, the phylogenetic analyses of the 13 concatenated PCG from 35 amphipod species using Bayesian Inference (BI) resulted in a well-supported phylogeny, with the

crangonyctid species forming a well-supported monophyletic group. Within Crangonyctidae, *Stygobromus* species formed a monophyletic group sister to *Bactrurus* + *Crangonyx*.

With selection pressure analyses, the one-ratio model (model 0) conducted for all 13 PCG revealed that the ω values for each gene ranged from 0.021 to 0.130 and were significantly less than 1. With the free-ratio model (model 1), the *cox2* gene significantly differed in ω values between the amphipods of the two habitat types with higher ω values for the surface amphipods. Similarly, *cox1* and *cox3* genes also exhibited a similar trend (but not significant). LRT tests showed that the two-ratio model (model 2) fits were significantly better than the one-ratio model for two genes: *atp6* and *cytb*, indicating a divergence in selective pressure between surface and subterranean amphipods. With the two-ratio branch model, several genes in surface amphipod mitogenomes were found to be undergoing positive selection ($\omega 1 > \omega 0$). In contrast, several genes in subterranean amphipod mitogenomes were undergoing purifying selection ($\omega 1 < \omega 0$). Surprisingly, a few genes in subterranean taxa displayed positive selection ($\omega 1 > \omega 0$). The M8 site model identified one positively selected site on the *atp8* gene (37 N) and one positively selected site on the *nad5* gene (482 Q). Similarly, the M2a site model identified two positively selected sites (37 N & 31 S) on the *atp8* gene. Using the more stringent branch-site model, we detected positive selection in 14 branches and 12 genes with a total of 308 amino acid sites under positive selection. Among them, 80 amino acid sites in seven were identified on the subterranean branches, whereas 228 amino acid sites in 10 genes were identified on the surface branches. Using aBSREL on the concatenated 13 PCG dataset comprising all 35 species, we detected episodic diversifying selection in seven species. Using BUSTED,

we detected evidence of gene-wide episodic diversifying selection in three of the surface species, whereas such evidence was found in just one of the subterranean species. Using RELAX, we detected evidence of relaxed selection in in six of the surface species. Contrastingly, evidence of intensification of selection was detected in four of subterranean species. We found evidence of directional selection in *atp8* of *C. forbesi* and *nad3* of *S. pizzinii* using aBSREL and *cox3* of *B. brachycaudus* using BUSTED. In addition, RELAX analyses uncovered five genes (*cox1, cox3, cytb, nad1,* and *nad3*) that exhibited relaxed selection and one gene (*atp6*) that exhibited intensification of selection in *C. forbesi*. Several genes in subterranean species including S. *tenuis, S. allegheniensis,* and *S. pizzinii* exhibited varying levels of intensification of selection, whereas none exhibited relaxed selection. Therefore, evidence of directional selection was detected in several protein-encoding genes of the OXPHOS pathway in the mitogenomes of surface amphipods, while a signal of purifying selection was more prominent in subterranean species.

Overall, gene order, locations of non-coding regions, and base-substitution rates points to habitat as an important factor influencing the evolution of amphipod mitogenomes. This study provides a necessary foundation for the study of mitogenome evolution in amphipods and other crustaceans.

CHAPTER 3

The objective of this chapter was to generate the first molecular phylogenetic framework for the study of the origin, diversification, and biogeography of cave trechines in eastern North America using ultraconserved elements (UCEs) phylogenomics. Monophyly of several subterranean lineages often has been viewed as evidence for a single

colonization and a strong role for diversification occurring underground. However, the same phylogenetic pattern can result in multiple colonizations by a widespread surface ancestor but the ancestor subsequently went extinct. I used a multilocus dataset (68 UCE loci) from 45 cave trechine species distributed primarily in the Appalachians valley and ridge (APP), Interior Low Plateau (ILP), and Ozarks (OZK) karst regions to develop a robust phylogenetic framework. We performed two different types of phylogenomic analyses on the 75% complete data-matrix: (1) concatenated analyses using RaxML on the CIPRES Science and MrBayes module of PhyloSuite (2) species tree analyses using ASTRAL-II and SVDQuartets implemented in PAUP. To estimate the relative age of origin and divergence of the cave trechine lineages, we used the Bayesian relaxed phylogenetic approach implemented in BEAST2 on the concatenated 75% complete data-matrix. We estimated ancestral geographic ranges and patterns of dispersal using stochastic likelihood-based models (DEC vs. DEC+J) of geographic range evolution implemented in the R package BioGeoBEARS.

Both ML and Bayesian inference (BI) trees inferred from the 75% concatenated UCE data matrix resolved similar topologies with high support for most nodes. Among the 16 *Pseudanophthalmus* species groups, 12 were monophyletic. Four species groups were paraphyletic, including the *cumberlandus*, *grandis*, *jonesi*, and *simplex* species groups. The genus *Pseudanophthalmus* was not recovered as monophyletic as all of the other cave trechine genera (*Ameroduvialis*, *Darlingtonea*, *Neaphaenops*, and *Nelsonites*) were nested within *Pseudanophthalmus*. The species tree inferred from the 75% UCE data matrix using ASTRAL-II differed slightly from concatenated ML and BI trees and involved the placement of *P. thomasi*, *P. fridigus*, *N. telkampfi viator*, *P. troglodytes*, and *D.*

kentuckensis within primary clades. The species tree reconstructed using SVDQuartets was similar in topology to the concatenated ML and BI trees. However, several deeper nodes in both ASTRAL SVDQuartets species trees were weakly supported.

Divergence time estimation in BEAST2 using a molecular clock rate recovered a crown age of 11.5 Mya in the middle Miocene forming the main ARV and ILP clades. Additional diversification of primary clades within the ARV and ILP clades occurred thereafter at 9.5 Mya and 10.7 Mya, respectively. Diversification within the primary clades occurred primarily from the late Miocene into the Pliocene with estimated species divergences all slightly predated the Pliocene (*i.e.*, 5 Mya). The surface ancestor(s) to cave trechines appears to have colonized the eastern North America karst regions from an east to west manner during the middle Miocene and Pliocene. The ancestral surface origin is uncertain but appears to have been in the southern Appalachians. A surface ancestor likely dispersed into the Ozarks Highlands from the APP between 4.1–7.2 Mya. The APP appears to have served both as a cradle for in situ diversification and as bridge linking the southern Appalachians and ILP, enabling the dispersal and subsequent diversification of these cave beetles. In the APP, there was a burst of diversification in the early Pliocene 7.0 Mya as well as in the Pleistocene 1.2 – 3.7 Mya. After colonization into the ILP, there was further diversification around 10.7 Mya followed by a burst of diversification in the late Miocene 8.0 Mya as well as in the Pliocene into the Pleistocene 1.2 – 5.7 Mya. These results support a multiple colonization scenario by a now extinct surface ancestor and a more limited role for subterranean speciation via dispersal.

This study represents the first effort to establish a time-calibrated phylogenomic framework for cave trechine beetles in North America, elucidating a rich and intriguing

history of evolution. The molecular systematics and biogeography of these unique cave beetles offer a model for other comparative evolutionary and ecological studies of troglobionts to further our understanding of factors driving speciation and biogeographic patterns.

REFERENCES

1. Aley, T. J. (1976). Hydrology and surface management. In: The National Cave Management Symposium Proceedings. Albuquerque, New Mexico, p. 44-45.
2. Arfianti, T., Wilson, S., & Costello, M. J. (2018). Progress in the discovery of amphipod crustaceans. PeerJ, 6, e5187.
3. Badino, G. (2004). Cave temperatures and global climatic change. International Journal of Speleology, 33(1), 10.
4. Barnard, J. L., & Karaman, G. S. (1991) The families and genera of marine gammaridean Amphipoda (except marine gammaroids). Records of the Australian Museum, Supplement, 13(1/2): 1–866.
5. Barr, T. C. (2004) A Classification and Checklist of the Genus *Pseudanophthalmus Jeannel* (Coleoptera: Carabidae: Trechinae). Virginia Museum of Natural History Special Publication 11.
6. Bousfield, E. L. (1983). An updated phyletic classification and palaeohistory of the Amphipoda. Crustacean Issues, 1: pp. 257-277
7. Cameron, S. L., Yoshizawa, K., Mizukoshi, A., Whiting, M. F., & Johnson, K. P. (2011). Mitochondrial genome deletions and minicircles are common in lice (Insecta: Phthiraptera). Bmc Genomics, 12(1), 1-15.
8. Castro, L. R., Austin, A. D., & Dowton, M. (2002). Contrasting rates of mitochondrial molecular evolution in parasitic Diptera and Hymenoptera. Molecular biology and evolution, 19(7), 1100-1113.
9. Chen, M., Guo, W., Huang, S., Luo, X., Tian, M., & Liu, W. (2021). Morphological

adaptation of cave-dwelling ground beetles in china revealed by geometric morphometry (Coleoptera, Carabidae, Trechini). Insects, 12(11), 1002.

10. Chevaldonné, P., & Lejeusne, C. (2003). Regional warming-induced species shift in north-west Mediterranean marine caves. Ecology Letters, 6(4), 371-379.
11. Christman, M. C., Culver, D. C., Madden, M. K., & White, D. (2005). Patterns of endemism of the eastern North American cave fauna. Journal of Biogeography, 32(8), 1441-1452.
12. Cooper, S. J. B., Hinze, S., Leys, R., Watts, C. H. S., and Humphreys, W. F. (2002). Islands under the desert: molecular systematics and evolutionary origins of stygobitic water beetles (Coleoptera: Dytiscidae) from central Western Australia. Invertebrate Systematics 16, 589–598.
13. Culver, D. C. (1986). Cave faunas. In: Soulé M. (Ed.), Conservation biology. Sinauer Associates, Inc., Sunderland, p. 427-443.
14. Culver, D. C., Kane, T. C., Fong, D. W. (1995) Adaptation and natural selection in caves: the evolution of *Gammarus minus*. Harvard University Press, Cambridge, MA.
15. Culver, D. C., & Pipan, T. (2009) Biology of caves and other subterranean habitats. Oxford University Press, Oxford, UK.
16. Culver, D. C., & Sket, B. (2002). Biological monitoring in caves. Acta Carsologica, 31(1).
17. Culver, D. C., Master, L. L., Christman, M. C., & Hobbs III, H. H. (2000). Obligate cave fauna of the 48 contiguous United States. Conservation Biology, 14(2), 386-401.
18. Dayrat, B. (2005). Towards integrative taxonomy. Biological Journal of the Linnean Society 85, 407–415.

19. Deharveng, L. (2000). The cave fauna of Southeast Asia. Origin, evolution and ecology. Ecosystems of the world Vol 30 Subterranean ecosystems, 603-632.
20. Derkarabetian, S., Paquin, P., Reddell, J., & Hedin, M. (2022). Conservation genomics of federally endangered Texella harvester species (Arachnida, Opiliones, Phalangodidae) from cave and karst habitats of central Texas. Conservation Genetics, 23(2), 401-416.
21. Dooley, K. E., Niemiller, K. D. K., Sturm, N., & Niemiller, M. L. (2022). Rediscovery and phylogenetic analysis of the Shelta Cave Crayfish (*Orconectes sheltae* Cooper & Cooper, 1997), a decapod (Decapoda, Cambaridae) endemic to Shelta Cave in northern Alabama, USA. Subterranean Biology, 43, 11-31.
22. Dowton, M., Cameron, S. L., Dowavic, J. I., Austin, A. D., & Whiting, M. F. (2009). Characterization of 67 mitochondrial tRNA gene rearrangements in the Hymenoptera suggests that mitochondrial tRNA gene position is selectively neutral. Molecular Biology and Evolution, 26(7), 1607-1617.
23. Eberhard, S. M., Halse, S. A., & Humphreys, W. F. (2005). Stygofauna in the Pilbara region, north-west Western Australia: a review. Journal of the Royal Society of Western Australia, 88, 167.
24. Elliott, W. R. (2000). Conservation of the North American cave and karst biota. In: Wilkens H., Culver D.C. & Humphreys W.F. (Eds.), Ecosystems of the World 30: Subterranean ecosystems. Elsevier, Amsterdam, p.665-689.
25. Elliott, W. R. (1992). Fire ants invade Texas caves. American Caves, 5(1), 13.
26. Encinares, J. M. A., & Lit Jr, I. L. (2014). Evaluation of leaf litter baits for sampling insects in Bulalon Cave, Burdeos, Polillo Island, Quezon Province, Philippines.

Philippine Entomologist, 28(1), 76-89.

27. Faille, A., Andújar, C., Fadrique, F., Ribera, I., (2014) Late Miocene origin of an Ibero-Maghrebian clade of ground beetles with multiple colonizations of the subterranean environment. Journal of Biogeography, 41, 1979–1990.
28. Faille, A., Ribera, I., Deharveng, L., Bourdeau, C., Garnery, L., Quéinnec, E., Deuve, T., (2010). A molecular phylogeny shows the single origin of the Pyrenean subterranean Trechini ground beetles (Coleoptera: Carabidae). Molecular Phylogenetics and Evolution, 54, 97–106.
29. Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic biology, 61(5), 717-726.
30. Gao, Z., Wynne, J. J., & Zhang, F. (2018). Two new species of cave-adapted pseudoscorpions (Pseudoscorpiones: Neobisiidae, Chthoniidae) from Guangxi, China. The Journal of Arachnology, 46(2), 345-354.
31. Gibert, J., & Deharveng, L. (2002). Subterranean Ecosystems: A Truncated Functional Biodiversity: This article emphasizes the truncated nature of subterranean biodiversity at both the bottom (no primary producers) and the top (very few strict predators) of food webs and discusses the implications of this truncation both from functional and evolutionary perspectives. BioScience, 52(6), 473-481.
32. Gilgado, J. D., Enghoff, H., Tinaut, A., & Ortuno, V. M. (2015). Hidden biodiversity in the Iberian Mesovoid Shallow Substratum (MSS): New and poorly known species of the millipede genus *Archipolydesmus* Attems, 1898 (Diplopoda, Polydesmidae).

Zoologischer Anzeiger-A Journal of Comparative Zoology, 258, 13-38.

33. Graening, G. O., & Brown, A. V. (2003). Ecosystem dynamics and pollution effects in an Ozark cave stream. Journal of the American Water Resources Association, 39(6), 1497-1507.

34. Grant, E. H. C., Mulder, K. P., Brand, A. B., Chambers, D. B., Wynn, A. H., Capshaw, G., Niemiller, M. L., Phillips, J. G., Jacobs, J. F., Kuchta, S. R., & Bell, R. C. (2022). Speciation with gene flow in a narrow endemic West Virginia cave salamander (*Gyrinophilus subterraneus*). Conservation Genetics, 1-18.

35. Hart, P. B., Niemiller, M. L., Burress, E. D., Armbruster, J. W., Ludt, W. B., & Chakrabarty, P. (2020). Cave-adapted evolution in the North American amblyopsid fishes inferred using phylogenomics and geometric morphometrics. Evolution, 74(5), 936-949.

36. Hedin, M., Derkarabetian, S., Blair, J., & Paquin, P. (2018). Sequence capture phylogenomics of eyeless Cicurina spiders from Texas caves, with emphasis on US federally-endangered species from Bexar County (Araneae, Hahniidae). ZooKeys, (769), 49.

37. Hillis, D. M., and Wiens, J. J. (2000). Molecules versus morphology in systematics: conflicts, artifacts, and misconceptions. In 'Phylogenetic Analysis of Morphological Data'. (Ed. J. J. Wiens.) pp. 1–19. (Smithsonian Institution Press: Washington, DC, USA.)

38. Horton, T., Thurston, M. H., Vlierboom, R., Gutteridge, Z., Pebody, C. A., Gates, A. R., & Bett, B. J. (2020). Are abyssal scavenging amphipod assemblages linked to climate cycles? Progress in Oceanography, 184, 102318.

39. Howarth, F. G. (1983). Ecology of cave arthropods. Annual review of entomology, 28(1), 365-389.

40. Howarth, F. G., & Stone, F. D. (1982). Conservation of Hawaii's speleological resources. In Proceedings of the third international symposium on volcanospeleology, Bend, Oregon (pp. 124-126).

41. Howarth, F. G., James, S. A., McDowell, W., Preston, D. J., & Imada, C. T. (2007). Identification of roots in lava tube caves using molecular techniques: implications for conservation of cave arthropod faunas. Journal of Insect Conservation, 11(3), 251-261.

42. Howarth, F. G., James, S. A., McDowell, W., Preston, D. J., & Imada, C. T. (2007). Identification of roots in lava tube caves using molecular techniques: implications for conservation of cave arthropod faunas. Journal of Insect Conservation, 11(3), 251-261.

43. Juan, C., Guzik, M. T., Jaume, D., & Cooper, S. J. (2010). Evolution in caves: Darwin's 'wrecks of ancient life'in the molecular era. Molecular Ecology, 19(18), 3865-3880.

44. Katz, A. D., Taylor, S. J., & Davis, M. A. (2018). At the confluence of vicariance and dispersal: Phylogeography of cavernicolous springtails (Collembola: Arrhopalitidae, Tomoceridae) codistributed across a geologically complex karst landscape in Illinois and Missouri. Ecology and Evolution, 8(20), 10306-10325.

45. Laurenti, J. N. (1768) Specimen Medicum, Exhibens Synopsin Reptilium Emendatam Cum Experimentis Circa Venenaet Antidota Reptilium Austriacorum. Wien. 214pp.

46. Lee, M. S. Y. (2004). The molecularisation of taxonomy. Invertebrate Systematics 18,

1–6.

47. Lefébure, T., Douady, C. J., and Gibert, J. (2006). Relationship between morphological taxonomy and molecular divergence within Crustacea: proposal of a molecular threshold to help species delimitation. Molecular Phylogenetics and Evolution 40, 435–447.
48. Leray, V. L., Caravas, J., Friedrich, M., & Zigler, K. S. (2019). Mitochondrial sequence data indicate "Vicariance by Erosion" as a mechanism of species diversification in North American *Ptomaphagus* (Coleoptera, Leiodidae, Cholevinae) cave beetles. Subterranean Biology, 29, 35-57.
49. Lisowski, E. A. (1983). Distribution, habitat, and behavior of the Kentucky cave shrimp Palaemonias ganteri Hay. Journal of Crustacean Biology, 3(1), 88-92.
50. Loria, S., Zigler, K., & Lewis, J. (2011). Molecular phylogeography of the troglobiotic millipede *Tetracion* Hoffman, 1956 (Diplopoda, Callipodida, Abacionidae). International Journal of Myriapodology, 5, 35-48.
51. Ma, C., Yang, P., Jiang, F., Chapuis, M. P., Shali, Y., Sword, G. A., & Kang, L. E. (2012). Mitochondrial genomes reveal the global phylogeography and dispersal routes of the migratory locust. Molecular Ecology, 21(17), 4344-4358.
52. Mammola, S., Goodacre, S. L., & Isaia, M. (2018). Climate change may drive cave spiders to extinction. Ecography, 41(1), 233-243.
53. Mitchell, R. W. (1970). Total number and density estimates of some species of cavernicoles inhabiting Fern Cave, Texas. In Annales de Spéléologie (Vol. 25, pp. 73-90).
54. Munasinghe, D. H. N., Burridge, C. P., and Austin, C. M. (2004). The systematics of

freshwater crayfish of the genus *Cherax* Erichson (Decapoda: Parastacidae) in eastern Australia re-examined using nucleotide sequences from 12S rRNA and 16S rRNA genes. Invertebrate Systematics 18, 215–225.

55. Murphy, N. P., and Austin, C. M. (2002). A preliminary study of 16S rRNA sequence variation in Australian Macrobrachium shrimps (Palaemonidae: Decapoda) reveals inconsistencies in their current classification. Invertebrate Systematics 16, 697–701
56. Nelson, L. A., Lambkin, C. L., Batterham, P., Wallman, J. F., Dowton, M., Whiting, M. F., Yeates, D. K., & Cameron, S. L. (2012). Beyond barcoding: A mitochondrial genomics approach to molecular phylogenetics and diagnostics of blowflies (Diptera: Calliphoridae). Gene, 511(2), 131-142.
57. Niemiller, M. L., & Zigler, K. S. (2013). Patterns of cave biodiversity and endemism in the Appalachians and Interior Plateau of Tennessee, USA. PLoS One, 8(5), e64177.
58. Niemiller, M. L., Zigler, K. S., Ober, K. A., Carter, E. T., Engel, A. S., Moni, G., Philips, T. K., & Stephen, C. D. (2017). Rediscovery and conservation status of six short-range endemic *Pseudanophthalmus* cave beetles (Carabidae: Trechini). Insect Conservation and Diversity, 10(6), 495-501.
59. Notenboom, J., Plénet, S., & Turquin, M. J. (1994). Groundwater contamination and its impact on groundwater animals and ecosystems. Groundwater ecology, 477-504.
60. Ober, K. A., Niemiller, M. L., & Philips, T. K. (2022) Cave trechine (Coleoptera: Carabidae) diversity and biogeography in North America. In: Cave Life – Drivers of Diversity and Diversification (Wynne JJ, ed.), John Hopkins Press.
61. Olson, R. (2005). The ecological effects of Lock and Dam No. 6 in Mammoth Cave National Park. In People, places, and parks: Proceedings of the 2005 George Wright

society conference on parks, protected areas, and cultural sites (pp. 294-299). Hancock, Michigan: The George Wright Society.

62. Page, T. J., Choy, S. C., and Hughes, J. M. (2005b). The taxonomic feedback loop: symbiosis of morphology & molecules. Biology Letters 1, 139–142.
63. Proudlove, G., and Wood, P. J. (2003). The blind leading the blind: cryptic subterranean species and DNA taxonomy. Trends in Ecology & Evolution 18, 272–273.
64. Pulido-Bosch, A., Martin-Rosales, W., López-Chicano, M., Rodriguez-Navarro, C. M., & Vallejos, A. (1997). Human impact in a tourist karstic cave (Aracena, Spain). Environmental geology, 31(3), 142-149.
65. Reddell, J. R. (1994). The cave fauna of Texas with special reference to the western Edwards Plateau. In The Caves and Karst of Texas. National Speleological Society.
66. Reeves, W. K. (1999). Exotic species in North American caves. In Proceedings of the 1999 National Cave and Karst Management Symposium. Southeastern Cave Conservancy (pp. 164-166).
67. Ribera, I., Fresneda, J., Bucur, R., Izquierdo, A., Vogler, A.P., Salgado, J.M., Cieslak, A., (2010). Ancient origin of a Western Mediterranean radiation of subterranean beetles. BMC Evolutionary Biology, 10, 29.
68. Rizzo, V., Comas, J., Fadrique, F., Fresneda, J., Ribera, I., (2013). Early Pliocene range expansion of a clade of subterranean Pyrenean beetles. Journal of Biogeography, 40, 1861–1873.
69. Ryu, T., Seridi, L., & Ravasi, T. (2012). The evolution of ultraconserved elements with different phylogenetic origins. BMC evolutionary biology, 12(1), 1-11.

70. Salvato, P., Simonato, M., Battisti, A., & Negrisolo, E. (2008). The complete mitochondrial genome of the bag-shelter moth *Ochrogaster lunifer* (Lepidoptera, Notodontidae). BMC genomics, 9(1), 1-15.

71. Shao, R., Dowton, M., Murrell, A., & Barker, S. C. (2003). Rates of gene rearrangement and nucleotide substitution are correlated in the mitochondrial genomes of insects. Molecular Biology and Evolution, 20(10), 1612-1619.

72. Shao, R., Kirkness, E. F., & Barker, S. C. (2009). The single mitochondrial chromosome typical of animals has evolved into 18 minichromosomes in the human body louse, *Pediculus humanus*. Genome research, 19(5), 904-912.

73. Souza Silva, M., Martins, R. P., & Ferreira, R. L. (2015). Cave conservation priority index to adopt a rapid protection strategy: a case study in Brazilian Atlantic rain forest. Environmental management, 55(2), 279-295.

74. Stephen, S., Pheasant, M., Makunin, I. V., & Mattick, J. S. (2008). Large-scale appearance of ultraconserved elements in tetrapod genomes and slowdown of the molecular clock. Molecular biology and evolution, 25(2), 402-408.

75. Sugai, L. S. M., Ochoa-Quintero, J. M., Costa-Pereira, R., & Roque, F. O. (2015). Beyond aboveground. Biodiversity and Conservation, 24(8), 2109-2112.

76. Taylor, S. J., Krejca, J. K., Smith, J. E., Block, V. R., & Hutto, F. R. (2003). Investigation of the potential for red imported fire ant (*Solenopsis invicta*) impacts on rare karst invertebrates at Fort Hood, Texas: A field study. Illinois Natural History Survey.

77. Trajano, E. (2000). Cave Faunas in the Atlantic Tropical Rain Forest: Composition, Ecology, and Conservation 1. Biotropica, 32(4b), 882-893.

78. Ubick, D., & Briggs, T. S. (2002). The harvestman family Phalangodidae 4. A review of the genus *Banksula* (Opiliones, Laniatores). The Journal of Arachnology, 30(2), 435-451.

79. Whitten, T. (2009). Applying ecology for cave management in China and neighboring countries. Journal of Applied Ecology, 46(3), 520-523.

80. Wynne, J. J., Bernard, E. C., Howarth, F. G., Sommer, S., Soto-Adames, F. N., Taiti, S., Mockford, E. L., Horrocks, M., Pakarati, L., & Pakarati-Hotus, V. (2014). Disturbance relicts in a rapidly changing world: the Rapa Nui (Easter Island) factor. BioScience, 64(8), 711-718.

81. Wynne, J. J., Sommer, S., Howarth, F. G., Dickson, B. G., & Voyles, K. D. (2018). Capturing arthropod diversity in complex cave systems. Diversity and Distributions, 24(10), 1478-1491.

82. Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

83. Aunins, A. W., Nelms, D. L., Hobson, C. S., & King, T. L. (2016). Comparative mitogenomic analyses of three North American stygobiont amphipods of the genus Stygobromus (Crustacea: Amphipoda), Mitochondrial DNA Part B, 1:1, 560-563.

84. Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, M., Middendorf, M., Stadler, P. F. (2013). MITOS: improved de novo metazoan mitochondrial genome annotation. Molecular Phylogenetics and Evolution, 69:313–319.

85. Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics (Oxford, England), 30(15), 2114–2120.

86. Boore, J. L., Lavrov, D. V., Brown, W. M. (1998). Gene translocation links insects and crustaceans. Nature, 392:667–668.

87. Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular biology and evolution, 17(4), 540-552.

88. De Melo Rodovalho, C., Lyra, M. L., Ferro, M., & Bacci Jr, M. (2014). The mitochondrial genome of the leaf-cutter ant *Atta laevigata*: a mitogenome with a large number of intergenic spacers. PLoS One, 9(5), e97117.

89. Dierckxsens, N., Mardulyn, P., & Smits, G. (2017). NOVOPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic acids research, 45(4), e18-e18.

90. Holsinger, J. R. (1967). Systematics, speciation, and distribution of the subterranean amphipod genus Stygonectes (Gammaridae). Smithsonian Contributions to Zoology 259: 1-176.

91. Holsinger, J. R. (1977). "A Review of the Systematics of the Holarctic Amphipod Family Crangonyctidae." Crustaceana. Supplement, no. 4, pp. 244–281.

92. Holsinger, J. R. (1978). "Systematics of the subterranean amphipod genus *Stygobromus* (Crangonyctidae) : Part II. Species of the eastern United States." Smithsonian Contributions to Zoology. 1–144.

93. Hubricht, L., & Mackin, J. (1940). Descriptions of Nine New Species of Fresh-Water Amphipod Crustaceans with Notes and New Localties for Other Species. The American Midland Naturalist, 23(1), 187-218.

94. Jühling, F., Pütz, J., Bernt, M., Donath, A., Middendorf, M., Florentz, C., & Stadler, P. F. (2012). Improved systematic tRNA gene annotation allows new insights into the

evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements. Nucleic acids research, 40(7), 2833–2845.

95. Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution, 30, 772-780.

96. Kilpert, F., & Podsiadlowski, L. (2006). The complete mitochondrial genome of the common sea slater, *Ligia oceanica* (Crustacea, Isopoda) bears a novel gene order and unusual control region features. BMC genomics, 7, 241.

97. Krebes, L., & Bastrop, R. (2012). The mitogenome of *Gammarus duebeni* (Crustacea Amphipoda): A new gene order and non-neutral sequence evolution of tandem repeats in the control region. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics, 7(2), 201-211.

98. Konemann, S., & Holsinger, J. R. (2002). Systematics of the North American subterranean amphipod genus *Bactrurus* (Crangonyctidae).

99. Lanfear, R., Calcott, B., Ho, S. Y., & Guindon, S. (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular biology and evolution, 29(6), 1695-1701.

100. Niemiller, M. L., Porter, M. L., Keany, J., Gilbert, H., Fong, D. W., Culver, D. C., Hobson, C. S., Kendall, K. D., Davis, M. A. and Taylor, S. J., (2018). Evaluation of eDNA for groundwater invertebrate detection and monitoring: a case study with endangered *Stygobromus* (Amphipoda: Crangonyctidae). Conservation Genetics Resources, 10(2), pp.247-257.

101. Ojala, D., Montoya, J., Attardi, G. (1981). tRNA punctuation model of RNA processing in human mitochondria. Nature, 290:470–474.

102. Pons, J., Bauzà-Ribot, M. M., Jaume, D., Juan, C. (2014). Next-generation sequencing, phylogenetic signal and comparative mitogenomic analyses in *Metacrangonyctidae* (Amphipoda: Crustacea). BMC Genomics, 15:566.

103. Romanova, E. V., Aleoshin, V. V., Kamaltynov, R. M., Mikhailov, K. V., Logacheva, M. D., Sirotinina, E. A., Gornov, A.Y., Anikin, A.S. and Sherbakov, D.Y., (2016). Evolution of mitochondrial genomes in Baikalian amphipods. BMC genomics, 17(14), 1016.

104. Romanova, E. V., Bukin, Y. S., Mikhailov, K. V., Logacheva, M. D., Aleoshin, V. V., & Sherbakov, D. Y. (2020). Hidden cases of tRNA gene duplication and remolding in mitochondrial genomes of amphipods. Molecular phylogenetics and evolution, 144, 106710.

105. Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology, 61(3), 539-542.

106. Shoemaker, C. R. (1938). A new species of fresh-water amphipod of the genus Synpleonia, with remarks on related genera. Proceedings of the Biological Society of Washington 51: 137-142.

107. Stokkan, M., Jurado-Rivera, J., Carlos, J., Damià, J., Pons, J. (2015). Mitochondrial genome rearrangements at low taxonomic levels: three distinct mitogenome orders in the genus *Pseudoniphargus*. Mitochondrial DNA, 27:3579–3589.

108. Wheeler, D. L., Church, D. M., Federhen, S., Lash, A. E., Madden, T. L., Pontius, J. U., Schuler, G. D., Schriml, L. M., Sequeira, E., Tatusova, T. A., & Wagner, L. (2003). Database resources of the National Center for Biotechnology. Nucleic acids research, 31(1), 28–33.

109. Yang, H. M., Song, J. H., Kim, M. S., & Min, G. S. (2017). The complete mitochondrial genomes of two talitrid amphipods, *Platorchestia japonica* and P. *parapacifica* (Crustacea, Amphipoda). Mitochondrial DNA Part B, 2(2), 757-758.

110. Yang, J. S., & Yang, W. J. (2008). The complete mitochondrial genome sequence of the hydrothermal vent galatheid crab *Shinkaia crosnieri* (Crustacea: Decapoda: Anomura): a novel arrangement and incomplete tRNA suite. BMC Genomics, 9(1), 257.

111. Zhang, D., Gao, F., Jakovlić, I., Zou, H., Zhang, J., Li, W. X., & Wang, G. T. (2020). PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Molecular Ecology Resources, 20(1): p. 348–355.

112. Zhang, J. (1997). "Systematics of the Freshwater Amphipod Genus *Crangonyx* (Crangonyctidae) in North America", Doctor of Philosophy (PhD), dissertation, Biological Sciences, Old Dominion University, DOI: 10.25777/s8tw-2q09.

113. Arfianti, T., Wilson, S., & Costello, M. J. (2018). Progress in the discovery of amphipod crustaceans. PeerJ, 6, e5187.

114. Aunins, A. W., Nelms, D. L., Hobson, C. S., & King, T. L. (2016). Comparative mitogenomic analyses of three North American stygobiont amphipods of the genus *Stygobromus* (Crustacea: Amphipoda). Mitochondrial DNA Part B, 1(1), 560-563.

115. Ballard, J. W. O., & Pichaud, N. (2014). Mitochondrial DNA: more than an evolutionary bystander. Functional Ecology, 28(1), 218-231

116. Barnard, J. L. & Karaman, G. S. (1991) The families and genera of marine gammaridean Amphipoda (except marine gammaroids). Records of the Australian Museum, Supplement, 13(1/2): 1–866.

117. Bauzà-Ribot, M. M., Juan, C., Nardi, F., Oromí, P., Pons, J., & Jaume, D. (2012). Mitogenomic phylogenetic analysis supports continental-scale vicariance in subterranean thalassoid crustaceans. Current Biology, 22(21), 2069-2074.

118. Bauzà-Ribot, M. M., Jaume, D., Juan, C., & Pons, J. (2009). The complete mitochondrial genome of the subterranean crustacean *Metacrangonyx longipes* (Amphipoda): A unique gene order and extremely short control region: Full-Length Research Paper. Mitochondrial DNA, 20(4), 88-99.

119. Beall, C. M. (2007). Two routes to functional adaptation: Tibetan and Andean high-altitude natives. Proceedings of the National Academy of Sciences, 104(suppl 1), 8655-8660.

120. Benito, J. B., Porter, M. L., & Niemiller, M. L. (2021). The mitochondrial genomes of five spring and groundwater amphipods of the family Crangonyctidae (Crustacea: Amphipoda) from eastern North America. Mitochondrial DNA Part B, 6(6), 1662-1667.

121. Bernt, M., Merkle, D., Ramsch, K., Fritzsch, G., Perseke, M., Bernhard, D., Schlegel, M., Stadler, P. F. & Middendorf, M. (2007). CREx: inferring genomic rearrangements based on common intervals. Bioinformatics, 23(21), 2957-2958.

122. Bishop, R., Humphreys, W. F., & Longley, G. (2014). Epigean and hypogean *Palaemonetes* sp. (Decapoda, Palaemonidae) from Edwards Aquifer: An examination of trophic structure and metabolism. Subterranean Biology, 14, 79-102.

123. Boore, J. L. (1999). Animal mitochondrial genomes. Nucleic Acids Research, 27(8), 1767-1780.

124. Boore, J. L., Macey, J. R., & Medina, M. (2005). Sequencing and comparing whole mitochondrial genomes of animals. Methods in Enzymology, 395, 311-348.

125. Botero-Castro, F., Tilak, M. K., Justy, F., Catzeflis, F., Delsuc, F., & Douzery, E. J. (2018). In cold blood: compositional bias and positive selection drive the high evolutionary rate of vampire bats mitochondrial genomes. Genome Biology and Evolution, 10(9), 2218-2239.

126. Bourguignon, T., Tang, Q., Ho, S. Y., Juna, F., Wang, Z., Arab, D. A., Cameron, S. L., Walker, J., Rentz, D., Evans, T. A. & Lo, N. (2018). Transoceanic dispersal and plate tectonics shaped global cockroach distributions: evidence from mitochondrial phylogenomics. Molecular Biology and Evolution, 35(4), 970-983.

127. Bousfield, E. L. (1983). An updated phyletic classification and palaeohistory of the Amphipoda. Crustacean Issues, 1: pp. 257-277

128. Burger, G., Gray, M. W., & Lang, B. F. (2003). Mitochondrial genomes: anything goes. Trends in Genetics, 19(12), 709-716.

129. Carlini, D. B., & Fong, D. W. (2017). The transcriptomes of cave and surface populations of *Gammarus minus* (Crustacea: Amphipoda) provide evidence for positive selection on cave downregulated transcripts. PLoS One, 12(10), e0186173.

130. Carroll, J., Fearnley, I. M., Wang, Q., & Walker, J. E. (2009). Measurement of the molecular masses of hydrophilic and hydrophobic subunits of ATP synthase and complex I in a single experiment. Analytical biochemistry, 395(2), 249-255.

131. Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular Biology and Evolution, 17(4), 540-552

132. Castresana, J., Feldmaier-Fuchs, G., & Pääbo, S. (1998). Codon reassignment and amino acid composition in hemichordate mitochondria. Proceedings of the National Academy of Sciences, 95(7), 3703-3707.

133. Clary, D. O., & Wolstenholme, D. R. (1985). The mitochondrial DNA molecule of *Drosophila yakuba*: nucleotide sequence, gene organization, and genetic code. Journal of Molecular Evolution, 22(3), 252-271.

134. Crozier, R. H., & Crozier, Y. C. (1993). The mitochondrial genome of the honeybee Apis mellifera: complete sequence and genome organization. Genetics, 133(1), 97-117.

135. Culver, D. C., Kane, T. C., & Fong, D. W. (1995) Adaptation and natural selection in caves: the evolution of *Gammarus minus*. Harvard University Press, Cambridge, 223 pp.

136. Culver, D. C., & Pipan, T. 2009. The biology of caves and other subterranean habitats. Oxford: Oxford University Press.

137. Culver, D. C., Holsinger, J. R., Christman, M. C., & Pipan, T. (2010). Morphological differences among eyeless amphipods in the genus *Stygobromus* dwelling in different subterranean habitats. Journal of Crustacean Biology, 30(1), 68-74.

138. Da Fonseca, R. R., Johnson, W. E., O'Brien, S. J., Ramos, M. J., & Antunes, A. (2008). The adaptive evolution of the mammalian mitochondrial genome. BMC Genomics, 9(1), 1-22.

139. Das, J. (2006) The role of mitochondrial respiration in physiological and evolutionary adaptation. Bioessays, 28(9), 890-901.

140. Dotson, E. M., & Beard, C. 3. (2001). Sequence and organization of the mitochondrial genome of the Chagas disease vector, *Triatoma dimidiata*. Insect Molecular Biology, 10(3), 205-215.

141. Dowling, T. E., Martasian, D. P., & Jeffery, W. R. (2002). Evidence for multiple genetic forms with similar eyeless phenotypes in the blind cavefish, *Astyanax mexicanus*. Molecular biology and evolution, 19(4), 446-455.

142. Dröse, S., Krack, S., Sokolova, L., Zwicker, K., Barth, H. D., Morgner, N., Heide, H., Steger, M., Nübel, E., Zickermann, V. & Kerscher, S. (2011). Functional dissection of the proton pumping modules of mitochondrial complex I. PLoS Biology, 9(8), e1001128.

143. Engel, A. S., Porter, M. L., Stern, L. A., Quinlan, S., & Bennett, P. C. (2004). Bacterial diversity and ecosystem function of filamentous microbial mats from aphotic (cave) sulfidic springs dominated by chemolithoautotrophic “Epsilonproteobacteria”. FEMS Microbiology Ecology, 51(1), 31-53.

144. Gao, F., Chen, C., Arab, D. A., Du, Z., He, Y., & Ho, S. Y. (2019). EasyCodeML: A visual tool for analysis of selection using CodeML. Ecology and Evolution, 9(7), 3891-3898.

145. Garvin, M. R., Bielawski, J. P., & Gharrett, A. J. (2011). Positive Darwinian selection in the piston that powers proton pumps in complex I of the mitochondria of Pacific salmon. PLoS One, 6(9), e24127.

146. Gissi, C., Iannelli, F., & Pesole, G. (2008). Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. Heredity, 101(4), 301-320.

147. Graening, G. O., and Brown, A. V. (2003), Ecosystem dynamics and pollution effects in an Ozark cave stream 1. JAWRA Journal of the American Water Resources Association, 39(6), 1497-1507.

148. Grossman, L. I., Wildman, D. E., Schmidt, T. R., & Goodman, M. (2004). Accelerated evolution of the electron transport chain in anthropoid primates. Trends in Genetics, 20(11), 578-585.

149. Guo, H., Yang, H., Tao, Y., Tang, D., Wu, Q., Wang, Z., & Tang, B. (2018). Mitochondrial OXPHOS genes provides insights into genetics basis of hypoxia adaptation in anchialine cave shrimps. Genes & Genomics, 40(11), 1169-1180.

150. Hao, Y. J., Zou, Y. L., Ding, Y. R., Xu, W. Y., Yan, Z. T., Li, X. D., Fu, W.B., Li, T.J. & Chen, B. (2017). Complete mitochondrial genomes of *Anopheles stephensi* and *An. dirus* and comparative evolutionary mitochondriomics of 50 mosquitoes. Scientific Reports, 7(1), 1-13.

151. Hassanin, A., Leger, N. E. L. L. Y., & Deutsch, J. (2005). Evidence for multiple reversals of asymmetric mutational constraints during the evolution of the mitochondrial genome of Metazoa, and consequences for phylogenetic inferences. Systematic Biology, 54(2), 277-298.

152. Hassanin, A., Ropiquet, A., Couloux, A., & Cruaud, C. (2009). Evolution of the mitochondrial genome in mammals living at high altitude: new insights from a study of the tribe Caprini (Bovidae, Antilopinae). Journal of Molecular Evolution, 68(4), 293-310.

153. Hervant, F., Mathieu, J., Barré, H., Simon, K., & Pinon, C. (1997). Comparative study on the behavioral, ventilatory, and respiratory responses of hypogean and epigean crustaceans to long-term starvation and subsequent feeding. Comparative Biochemistry and Physiology Part A: Physiology, 118(4), 1277-1283.

154. Holsinger, J. R. (1978). Systematics of the subterranean amphipod genus *Stygobromus* (Crangonyctidae): Part II. Species of the eastern United States. Smithsonian Contributions to Zoology, 266: 1-144.

155. Horton, T., Thurston, M. H., Vlierboom, R., Gutteridge, Z., Pebody, C. A., Gates, A. R., & Bett, B. J. (2020). Are abyssal scavenging amphipod assemblages linked to climate cycles? Progress in Oceanography, 184, 102318.

156. Hubricht, L., & Mackin, J. G. (1940). Descriptions of nine new species of fresh-water amphipod crustaceans with notes and new localties for other species. The American Midland Naturalist, 23(1), 187-218.

157. Huntsman, B. M., Venarsky, M. P., Benstead, J. P., & Huryn, A. D. (2011). Effects of organic matter availability on the life history and production of a top vertebrate predator (Plethodontidae: Gyrinophilus palleucus) in two cave streams. Freshwater Biology, 56(9), 1746-1760.

158. Huppop, K. (2000). How do cave animals cope with the food scarcity in caves? Ecosystems of the World, 159-188.

159. Hyde, J., Cooper, S. J., Munguia, P., Humphreys, W. F., & Austin, A. D. (2018). The first complete mitochondrial genomes of subterranean dytiscid diving beetles (*Limbodessus* and *Paroster*) from calcrete aquifers of Western Australia. Australian Journal of Zoology, 65(5), 283-291.

160. Issartel, J., Voituron, Y., Guillaume, O., Clobert, J., & Hervant, F. (2010). Selection of physiological and metabolic adaptations to food deprivation in the Pyrenean newt Calotriton asper during cave colonisation. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 155(1), 77-83.

161. Ito, A., Aoki, M. N., Yokobori, S. I., & Wada, H. (2010). The complete mitochondrial genome of *Caprella scaura* (Crustacea, Amphipoda, Caprellidea), with emphasis on the unique gene order pattern and duplicated control region. Mitochondrial DNA, 21(5), 183-190.

162. Juan, C., Guzik, M. T., Jaume, D., & Cooper, S. J. (2010). Evolution in caves: Darwin's 'wrecks of ancient life'in the molecular era. Molecular Ecology, 19(18), 3865-3880.

163. Jühling, F., Pütz, J., Bernt, M., Donath, A., Middendorf, M., Florentz, C., & Stadler, P. F. (2012). Improved systematic tRNA gene annotation allows new insights into the evolution of mitochondrial tRNA structures and into the mechanisms of mitochondrial genome rearrangements. Nucleic Acids Research, 40(7), 2833–2845.

164. Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular biology and evolution, 30(4), 772-780.

165. Ki, J. S., Hop, H., Kim, S. J., Kim, I. C., Park, H. G., & Lee, J. S. (2010). Complete mitochondrial genome sequence of the Arctic gammarid, *Onisimus nanseni* (Crustacea; Amphipoda): Novel gene structures and unusual control region features. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics, 5(2), 105-115.

166. Kilpert, F., & Podsiadlowski, L. (2010). The mitochondrial genome of the Japanese skeleton shrimp Caprella mutica (Amphipoda: Caprellidea) reveals a unique gene order and shared apomorphic translocations with Gammaridea. Mitochondrial DNA, 21(3-4), 77-86.

167. Koenemann, S., & Holsinger, J. R. (2001). Systematics of the North American subterranean amphipod genus *Bactrurus* (Crangonyctidae). Beaufortia, 51(1), 1-56.

168. Kornobis, E., Pálsson, S., Sidorov, D. A., Holsinger, J. R., & Kristjánsson, B. K. (2011). Molecular taxonomy and phylogenetic affinities of two groundwater amphipods, *Crangonyx islandicus* and *Crymostygius thingvallensis*, endemic to Iceland. Molecular phylogenetics and evolution, 58(3), 527-539.

169. Kornobis, E., Pálsson, S. N. Æ. B. J. Ö. R. N., & Svavarsson, J. (2012). Classification of *Crangonyx islandicus* (Amphipoda, Crangonyctidae) based on morphological characters and comparison with molecular phylogenies. Zootaxa, 3233(1), 52-66.

170. Krebes, L., & Bastrop, R. (2012). The mitogenome of *Gammarus duebeni* (Crustacea Amphipoda): A new gene order and non-neutral sequence evolution of tandem repeats in the control region. Comparative Biochemistry and Physiology Part D: Genomics and Proteomics, 7(2), 201-211.

171. Lanfear, R., Calcott, B., Ho, S. Y., & Guindon, S. (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular Biology and Evolution, 29(6), 1695-1701.

172. Lessinger, A. C., Martins Junqueira, A. C., Lemos, T. A., Kemper, E. L., Da Silva, F. R., Vettore, A. L., Arruda, P. & L. Azeredo-Espin, A. M. (2000). The mitochondrial genome of the primary screwworm fly *Cochliomyia hominivorax* (Diptera: Calliphoridae). Insect Molecular Biology, 9(5), 521-529.

173. Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Research, 49(W1), W293-W296.

174. Li, M., Jin, L., Ma, J., Tian, S., Li, R., & Li, X. (2016). Detecting mitochondrial signatures of selection in wild Tibetan pigs and domesticated pigs. Mitochondrial DNA Part A, 27(1), 747-752.

175. Li, X. D., Jiang, G. F., Yan, L. Y., Li, R., Mu, Y., & Deng, W. A. (2018). Positive selection drove the adaptation of mitochondrial genes to the demands of flight and high-altitude environments in grasshoppers. Frontiers in Genetics, 9, 605.

176. Lin, F. J., Liu, Y., Sha, Z., Tsang, L. M., Chu, K. H., Chan, T. Y., Liu, R. & Cui, Z. (2012). Evolution and phylogeny of the mud shrimps (Crustacea: Decapoda) revealed from complete mitochondrial genomes. BMC Genomics, 13(1), 1-12.

177. Macey, J. R., Larson, A., Ananjeva, N. B., & Papenfuss, T. J. (1997). Replication slippage may cause parallel evolution in the secondary structures of mitochondrial transfer RNAs. Molecular Biology and Evolution, 14(1), 30-39.

178. Macher, J. N., Leese, F., Weigand, A. M., & Rozenberg, A. (2017). The complete mitochondrial genome of a cryptic amphipod species from the Gammarus fossarum complex. Mitochondrial DNA Part B, 2(1), 17-18.

179. Matsumoto, Y., Yanase, T., Tsuda, T., & Noda, H. (2009). Species-specific mitochondrial gene rearrangements in biting midges and vector species identification. Medical and Veterinary Entomology, 23(1), 47-55.

180. McKenzie, M., Lazarou, M., and Ryan, M. T. (2009). Analysis of respiratory chain complex assembly with radiolabeled nuclear- and mitochondrial-encoded subunits. Methods in Enzymology, 456, 321–339.

181. Meiklejohn, C. D., Montooth, K. L., & Rand, D. M. (2007). Positive and negative selection on the mitochondrial genome. Trends in Genetics, 23(6), 259-263.

182. Mishmar, D., Ruiz-Pesini, E., Golik, P., Macaulay, V., Clark, A. G., Hosseini, S., Brandon, M., Easley, K., Chen, E., Brown, M. D. & Sukernik, R. I. (2003). Natural selection shaped regional mtDNA variation in humans. Proceedings of the National Academy of Sciences, 100(1), 171-176.

183. Mitterboeck, T. F., & Adamowicz, S. J. (2013). Flight loss linked to faster molecular evolution in insects. Proceedings of the Royal Society B: Biological Sciences, 280(1767), 20131128.

184. Mitterboeck, T. F., Liu, S., Adamowicz, S. J., Fu, J., Zhang, R., Song, W., Meusemann, K. & Zhou, X. (2017). Positive and relaxed selection associated with flight evolution and loss in insect transcriptomes. GigaScience, 6(10), gix073.

185. Mu, W., Liu, J., & Zhang, H. (2018). Complete mitochondrial genome of *Benthodytes marianensis* (Holothuroidea: Elasipodida: Psychropotidae): Insight into deep sea adaptation in the sea cucumber. PloS One, 13(11), e0208051.

186. Murrell, B., Weaver, S., Smith, M. D., Wertheim, J. O., Murrell, S., Aylward, A., Eren, K., Pollner, T., Martin, D. P., Smith, D. M. & Scheffler, K. (2015). Gene-wide identification of episodic selection. Molecular Biology and Evolution, 32(5), 1365-1371.

187. Nair, P., Huertas, M., & Nowlin, W. H. (2020). Metabolic responses to long-term food deprivation in subterranean and surface amphipods. Subterranean Biology, 33, pp. 1–15.

188. Nardi, F., Carapelli, A., Fanciulli, P. P., Dallai, R., & Frati, F. (2001). The complete mitochondrial DNA sequence of the basal hexapod *Tetrodontophora bielanensis*: evidence for heteroplasmy and tRNA translocations. Molecular Biology and Evolution, 18(7), 1293-1304.

189. Niemiller, M. L., Porter, M. L., Keany, J., Gilbert, H., Fong, D. W., Culver, D. C., Hobson, C. S., Kendall, K. D., Davis, M. A. & Taylor, S. J. (2018). Evaluation of eDNA for groundwater invertebrate detection and monitoring: a case study with endangered *Stygobromus* (Amphipoda: Crangonyctidae). Conservation Genetics Resources, 10(2), 247-257.

190. Ojala, D., Montoya, J., & Attardi, G. (1981). tRNA punctuation model of RNA processing in human mitochondria. Nature, 290(5806), 470-474.

191. Okimoto, R., & Wolstenholme, D. R. (1990). A set of tRNAs that lack either the T psi C arm or the dihydrouridine arm: towards a minimal tRNA adaptor. The EMBO journal, 9(10), 3405-3411.

192. Oliveira, D. C., Raychoudhury, R., Lavrov, D. V., & Werren, J. H. (2008). Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). Molecular Biology and Evolution, 25(10), 2167-2180.

193. Peng, Y., Yang, Z., Zhang, H., Cui, C., Qi, X., Luo, X., Tao, X., Wu, T., Chen, H., Shi, H. & Su, B. (2011). Genetic variations in Tibetan populations and high-altitude adaptation at the Himalayas. Molecular Biology and Evolution, 28(2), 1075-1081.

194. Perseke, M., Fritzsch, G., Ramsch, K., Bernt, M., Merkle, D., Middendorf, M., Bernhard, D., Stadler, P. F. and Schlegel, M. (2008). Evolution of mitochondrial gene orders in echinoderms. Molecular Phylogenetics and Evolution, 47(2), 855-864.

195. Pons, J., Bauzà-Ribot, M. M., Jaume, D., & Juan, C. (2014). Next-generation sequencing, phylogenetic signal and comparative mitogenomic analyses in Metacrangonyctidae (Amphipoda: Crustacea). BMC Genomics, 15(1), 1-16.

196. Porter, M. L., Engel, A. S., Kane, T. C., & Kinkle, B. K. (2009). Productivity-diversity relationships from chemolithoautotrophically based sulfidic karst systems. International Journal of Speleology, 38(1), 4.

197. Poulson, T. L., & White, W. B. (1969). The Cave Environment: Limestone caves provide unique natural laboratories for studying biological and geological processes. Science, 165(3897), 971-981.

198. Protas, M. E., Trontelj, P., & Patel, N. H. (2011). Genetic basis of eye and pigment loss in the cave crustacean, *Asellus aquaticus*. Proceedings of the National Academy of Sciences, 108(14), 5702-5707.

199. Qiu, Q., Zhang, G., Ma, T., Qian, W., Wang, J., Ye, Z., Cao, C., Hu, Q., Kim, J., Larkin, D. M. & Auvil, L. (2012). The yak genome and adaptation to life at high altitude. Nature Genetics, 44(8), 946-949.

200. Qu, Y., Zhao, H., Han, N., Zhou, G., Song, G., Gao, B., Tian, S., Zhang, J., Zhang, R., Meng, X. & Zhang, Y. (2013). Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan plateau. Nature Communications, 4(1), 1-9.

201. R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

202. Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic Biology, 67(5), 901.

203. Rawlings, T. A., Collins, T. M., & Bieler, R. (2001). A major mitochondrial gene rearrangement among closely related species. Molecular Biology and Evolution, 18(8), 1604-1609.

204. Ren, J., Liu, X., Jiang, F., Guo, X., & Liu, B. (2010). Unusual conservation of mitochondrial gene order in *Crassostrea* oysters: evidence for recent speciation in Asia. BMC Evolutionary Biology, 10(1), 1-14.

205. Ren, J., Shen, X., Jiang, F., & Liu, B. (2010). The mitochondrial genomes of two scallops, *Argopecten irradians* and *Chlamys farreri* (Mollusca: Bivalvia): the most

highly rearranged gene order in the family Pectinidae. Journal of Molecular Evolution, 70(1), 57-68.

206. Reyes, A., Gissi, C., Pesole, G., & Saccone, C. (1998). Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. Molecular Biology and Evolution, 15(8), 957-966.

207. Rivarola-Duarte, L., Otto, C., Jühling, F., Schreiber, S., Bedulina, D., Jakob, L., Gurkov, A., Axenov-Gribanov, D., Sahyoun, A. H., Lucassen, M. & Hackermüller, J. (2014). A first glimpse at the genome of the Baikalian amphipod *Eulimnogammarus verrucosus*. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution, 322(3), 177-189.

208. Romanova, E. V., Aleoshin, V. V., Kamaltynov, R. M., Mikhailov, K. V., Logacheva, M. D., Sirotinina, E. A., Gornov, A. Y., Anikin, A. S. & Sherbakov, D. Y. (2016). Evolution of mitochondrial genomes in Baikalian amphipods. BMC Genomics, 17(14), 291-306.

209. Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology, 61, 539-542.

210. Ruiz-Pesini, E., Mishmar, D., Brandon, M., Procaccio, V., & Wallace, D. C. (2004). Effects of purifying and adaptive selection on regional variation in human mtDNA. Science, 303(5655), 223-226.

211. Scheffler, I. E. (1998). Molecular genetics of succinate: quinone oxidoreductase in eukaryotes. Progress in Nucleic Acid Research and Molecular Biology, 60, 267-315.

212. Scott, G. R., Schulte, P. M., Egginton, S., Scott, A. L., Richards, J. G., & Milsom, W. K. (2011). Molecular evolution of cytochrome c oxidase underlies high-altitude adaptation in the bar-headed goose. Molecular Biology and Evolution, 28(1), 351-363.

213. Shen, Y. Y., Liang, L., Zhu, Z. H., Zhou, W. P., Irwin, D. M., & Zhang, Y. P. (2010). Adaptive evolution of energy metabolism genes and the origin of flight in bats. Proceedings of the National Academy of Sciences, 107(19), 8666-8671.

214. Shen, Y. Y., Shi, P., Sun, Y. B., & Zhang, Y. P. (2009). Relaxation of selective constraints on avian mitochondrial DNA following the degeneration of flight ability. Genome Research, 19(10), 1760-1765.

215. Shin, S. C., Cho, J., Lee, J. K., Ahn, D. H., Lee, H., & Park, H. (2012). Complete mitochondrial genome of the Antarctic amphipod *Gondogeneia antarctica* (Crustacea, amphipod). Mitochondrial DNA, 23(1), 25-27.

216. Simon, K. S., & Benfield, E. F. (2001). Leaf and wood breakdown in cave streams. Journal of the North American Benthological Society, 20(4), 550-563.

217. Smith, M. D., Wertheim, J. O., Weaver, S., Murrell, B., Scheffler, K., & Kosakovsky Pond, S. L. (2015). Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. Molecular biology and evolution, 32(5), 1342-1353.

218. Soares, D., & Niemiller, M. L. (2020). Extreme adaptation in caves. The Anatomical Record, 303(1), 15-23.

219. Sun, S., Wu, Y., Ge, X., Jakovlić, I., Zhu, J., Mahboob, S., Al-Ghanim, K. A., Al-Misned, F. & Fu, H. (2020). Disentangling the interplay of positive and negative

selection forces that shaped mitochondrial genomes of *Gammarus pisinnus* and *Gammarus lacustris*. Royal Society open science, 7(1), 190669.

220. Taylor, S. J., & Niemiller, M. L. (2016). Biogeography and conservation assessment of *Bactrurus* groundwater amphipods (Crangonyctidae) in the central and eastern United States. Subterranean Biology, 17, 1-29.

221. Venarsky, M. P., Huntsman, B. M., Huryn, A. D., Benstead, J. P., & Kuhajda, B. R. (2014). Quantitative food web analysis supports the energy-limitation hypothesis in cave stream ecosystems. Oecologia, 176(3), 859-869.

222. Vermulst, M., Wanagat, J., Kujoth, G. C., Bielas, J. H., Rabinovitch, P. S., Prolla, T. A., & Loeb, L. A. (2008). DNA deletions and clonal mutations drive premature aging in mitochondrial mutator mice. Nature Genetics, 40(4), 392-394.

223. Weaver, S., Shank, S. D., Spielman, S. J., Li, M., Muse, S. V., & Kosakovsky Pond, S. L. (2018). Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. Molecular biology and evolution, 35(3), 773-777.

224. Wei, S. J., Shi, M., Chen, X. X., Sharkey, M. J., van Achterberg, C., Ye, G. Y., & He, J. H. (2010). New views on strand asymmetry in insect mitochondrial genomes. PLoS One, 5(9), e12708.

225. Welch, A. J., Bedoya-Reina, O. C., Carretero-Paulet, L., Miller, W., Rode, K. D., & Lindqvist, C. (2014). Polar bears exhibit genome-wide signatures of bioenergetic adaptation to life in the arctic environment. Genome Biology and Evolution, 6(2), 433-450.

226. Wertheim, J. O., Murrell, B., Smith, M. D., Kosakovsky Pond, S. L., & Scheffler, K. (2015). RELAX: detecting relaxed selection in a phylogenetic framework. Molecular biology and evolution, 32(3), 820-832.

227. Wilson, K., Cahill, V., Ballment, E., & Benzie, J. (2000). The complete sequence of the mitochondrial genome of the crustacean *Penaeus monodon*: are malacostracan crustaceans more closely related to insects than to branchiopods? Molecular Biology and Evolution, 17(6), 863-874.

228. Wirth, C., Brandt, U., Hunte, C., & Zickermann, V. (2016). Structure and function of mitochondrial complex I. Biochimica et Biophysica Acta (BBA)-Bioenergetics, 1857(7), 902-914.

229. Wu, X., Wang, L., Chen, S., Zan, R., Xiao, H., & Zhang, Y. P. (2010). The complete mitochondrial genomes of two species from *Sinocyclocheilus* (Cypriniformes: Cyprinidae) and a phylogenetic analysis within Cyprininae. Molecular Biology Reports, 37(5), 2163-2171.

230. Xiao, J. H., Jia, J. G., Murphy, R. W., & Huang, D. W. (2011). Rapid evolution of the mitochondrial genome in chalcidoid wasps (Hymenoptera: Chalcidoidea) driven by parasitic lifestyles. PLoS One, 6(11), e26645.

231. Xin, Y., Ren, J., & Liu, X. (2011). Mitogenome of the small abalone *Haliotis diversicolor* Reeve and phylogenetic analysis within Gastropoda. Marine Genomics, 4(4), 253-262.

232. Xu, B., & Yang, Z. (2013). PAMLX: a graphical user interface for PAML. Molecular Biology and Evolution, 30(12), 2723-2724.

233. Xu, S., Luosang, J., Hua, S., He, J., Ciren, A., Wang, W., Tong, X., Liang, Y., Wang, J., & Zheng, X. (2007). High altitude adaptation and phylogenetic analysis of Tibetan horse based on the mitochondrial genome. Journal of Genetics and Genomics, 34(8), 720-729.

234. Yamazaki, N., Ueshima, R., Terrett, J. A., Yokobori, S. I., Kaifu, M., Segawa, R., Kobayashi, T., Numachi, K. I., Ueda, T., Nishikawa, K., & Watanabe, K. (1997). Evolution of pulmonate gastropod mitochondrial genomes: comparisons of gene organizations of Euhadra, Cepaea and Albinaria and implications of unusual tRNA secondary structures. Genetics, 145(3), 749-758.

235. Yang, L., Wang, Y., Zhang, Z., & He, S. (2015). Comprehensive transcriptome analysis reveals accelerated genic evolution in a Tibet fish, *Gymnodiptychus pachycheilus*. Genome biology and evolution, 7(1), 251-261.

236. Yang, Y., Xu, S., Xu, J., Guo, Y., & Yang, G. (2014). Adaptive evolution of mitochondrial energy metabolism genes associated with increased energy demand in flying insects. PloS One, 9(6), e99120.

237. Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution, 24(8), 1586-1591.

238. Yang, H. M., Song, J. H., Kim, M. S., & Min, G. S. (2017). The complete mitochondrial genomes of two talitrid amphipods, Platorchestia japonica and P. parapacifica (Crustacea, Amphipoda). Mitochondrial DNA Part B, 2(2), 757-758.

239. Zapelloni, F., Jurado-Rivera, J. A., Jaume, D., Juan, C., & Pons, J. (2021). Comparative Mitogenomics in Hyalella (Amphipoda: Crustacea). Genes, 12(2), 292.

240. Zhang, D., F. Gao, I. Jakovlić, H. Zou, J. Zhang, W.X. Li, and G.T. Wang (2020). PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Molecular Ecology Resources, 20(1), 348–355

241. Zhang, D., Li, W. X., Zou, H., Wu, S. G., Li, M., Jakovlić, I., Zhang, J., Chen, R., & Wang, G. T. (2018). Mitochondrial genomes of two diplectanids (Platyhelminthes: Monogenea) expose paraphyly of the order Dactylogyridea and extensive tRNA gene rearrangements. Parasites & Vectors, 11(1), 1-13.

242. Zhang, D., Zou, H., Hua, C. J., Li, W. X., Mahboob, S., Al-Ghanim, K. A., Al-Misned, F., Jakovlić, I., & Wang, G. T. (2019). Mitochondrial architecture rearrangements produce asymmetrical nonadaptive mutational pressures that subvert the phylogenetic reconstruction in Isopoda. Genome Biology and Evolution, 11(7), 1797-1812.

243. Zhang, H., Luo, Q., Sun, J., Liu, F., Wu, G., Yu, J., & Wang, W. (2013). Mitochondrial genome sequences of *Artemia tibetiana* and *Artemia urmiana*: assessing molecular changes for high plateau adaptation. Science China Life Sciences, 56(5), 440-452.

244. Zhang, J., Nielsen, R., & Yang, Z. (2005). Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Molecular Biology and Evolution, 22(12), 2472-2479.

245. Zhang, Z. Y., Chen, B., Zhao, D. J., & Kang, L. (2013). Functional modulation of mitochondrial cytochrome c oxidase underlies adaptation to high-altitude hypoxia in a Tibetan migratory locust. Proceedings of the Royal Society B: Biological Sciences, 280(1756), 20122758.

246. Zou, H., Jakovlić, I., Zhang, D., Chen, R., Mahboob, S., Al-Ghanim, K. A., Al-Misned, F., Li, W. X., & Wang, G. T. (2018). The complete mitochondrial genome of *Cymothoa indica* has a highly rearranged gene order and clusters at the very base of the Isopoda clade. PLoS one, 13(9), e0203089.

247. Zou, H., Jakovlić, I., Chen, R., Zhang, D., Zhang, J., Li, W. X., & Wang, G. T. (2017). The complete mitochondrial genome of parasitic nematode *Camallanus cotti*: extreme discontinuity in the rate of mitogenomic architecture evolution within the Chromadorea class. BMC Genomics, 18(1), 1-17.

248. Anthony, D. M., & Granger, D. E. (2004). A Late Tertiary Origin for Multilevel Caves Along the Western Escarpment of the Cumberland Plateau, Tennessee and Kentucky, Established by Cosmogenic super (26) Al and super (10) Be. Journal of Cave and Karst Studies, 66(2), 46–55.

249. Anthony, D. M., & Granger, D. E. (2007). A new chronology for the age of Appalachian erosional surfaces determined by cosmogenic nuclides in cave sediments. Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group, 32(6), 874-887.

250. Ashmole, N. P. (1993) Colonization of the underground environment in volcanic islands. Mémoires de Biospéologie, 20, 1–11.

251. Baca, S. M., Alexander, A., Gustafson, G. T., & Short, A. E. (2017). Ultraconserved elements show utility in phylogenetic inference of Adephaga (Coleoptera) and suggest paraphyly of 'Hydradephaga'. Systematic Entomology, 42(4), 786–795.

252. Barr, T. C. (1965). The Pseudanophthalmus of the Appalachian Valley (Coleoptera: Carabidae). American Midland Naturalist, 41-72.

253. Barr, T. C. (1967a) Observations on the ecology of caves. American Naturalist 101:475–491.

254. Barr, T. C. (1967b) A new Pseudanophthalmus from an epigean environment in West Virginia (Coleoptera: Carabidae). Psyche, 74, 166–172.

255. Barr, T. C. (1972) Trechoblemus in North America, with a key to North American genera of Trechinae (Coleoptera: Carabidae). Psyche 78:140–149.

256. Barr, T. C. (1980) New species groups of Pseudanophthalmus from the Central Basin of Tennessee (Coleoptera: Carabidae: Trechinae). Brimleyana, 3, 85–96.

257. Barr, T. C. (1981) Pseudanophthalmus from Appalachian caves (Coleoptera: Carabidae): the engelhardti complex. Brimleyana, 5, 37–94.

258. Barr, T. C. (1985a) Pattern and process in speciation of trechine beetles in eastern North America (Coleoptera: Carabidae: Trechinae). In Taxonomy, Phylogeny, and Biogeography of Beetles and Ants, edited by G.E. Ball, 350–407. Dordrecht, The Netherlands.

259. Barr, T. C. (1985b) New trechine beetles (Coleoptera: Carabidae) from the Appalachian region. Brimleyana, 11, 119 –132.

260. Barr, T. C. (2004) A Classification and Checklist of the Genus Pseudanophthalmus Jeannel (Coleoptera: Carabidae: Trechinae). Virginia Museum of Natural History Special Publication 11.

261. Barr, T. C., & Peck, S. B. (1965). Occurrence of a troglobitic Pseudanophthalmus outside a cave (Coleoptera: Carabidae). American Midland Naturalist, 73-74.

262. Barr, T. C., and Holsinger, J. (1985) Speciation in cave faunas. Annual Review of Ecology, Evolution, and Systematics 16:313–337.

263. Barr, T. C., and Krekeler, C. H. (1967). Xenotrechus, a new genus of cave trechines from Missouri (Coleoptera: Carabidae). Annals of the Entomological Society of America, 60(6), 1322-1325.

264. Bendik, N. F., Meik, J. M., Gluesenkamp, A. G., Roelke, C. E., & Chippindale, P. T. (2013). Biogeography, phylogeny, and morphological evolution of central Texas cave and spring salamanders. BMC Evolutionary Biology, 13(1), 1-18.

265. Bolger, A. M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30, 2114–2120.

266. Boyd, O. F., Philips, T. K., Johnson, J. R., & Nixon, J. J. (2020) Geographically structured genetic diversity in the cave beetle Darlingtonea kentuckensis Valentine, 1952 (Coleoptera, Carabidae, Trechini, Trechina). Subterranean Biology, 34, 1-23.

267. Bradford, T. M., Ruta, R., Cooper, S. J., Libonatti, M. L., & Watts, C. H. (2022). Evolutionary history of the Australasian Scirtinae (Scirtidae; Coleoptera) inferred from ultraconserved elements. Invertebrate Systematics, 36(4), 291–305.

268. Branstetter, M. G., Longino, J. T., Ward, P. S., & Faircloth, B. C. (2017) Enriching the ant tree of life: Enhanced UCE bait set for genome-scale phylogenetics of ants and other Hymenoptera. Methods in Ecology and Evolution, 8(6), 768–776.

269. Buhay, J. E., & Crandall, K. A. (2008). Taxonomic revision of cave crayfishes in the genus Orconectes, subgenus Orconectes (Decapoda: Cambaridae) along the Cumberland Plateau, including a description of a new species, Orconectes barri. Journal of Crustacean Biology, 28(1), 57-67.

270. Buhay, J. E., & Crandall, K. A. (2009). Taxonomic revision of cave crayfish in the genus Cambarus, subgenus Aviticambarus (Decapoda: Cambaridae) with descriptions

of two new species, C. speleocoopi and C. laconensis, endemic to Alabama, USA. Journal of Crustacean Biology, 29(1), 121-134.

271. Buhay, J. E., Moni, G., Mann, N., Crandall, K. A. (2007) Molecular taxonomy in the dark: evolutionary history, phylogeography, and diversity of cave crayfish in the subgenus Aviticambarus, genus Cambarus. Molecular Phylogenetics and Evolution, 42, 435–448.

272. Buhay, J., Crandall, K. (2005) Subterranean phylogeography of freshwater crayfishes shows extensive gene flow and surprisingly large population sizes. Molecular Ecology, 14, 4259–4273.

273. Burnham, K. P., & Anderson, D. R. (2002) Model selection and multimodel inference: a practical information-theoretic approach. Springer New York, New York, pp 267–351.

274. Chen, M., Guo, W., Huang, S., Luo, X., Tian, M., & Liu, W. (2021). Morphological adaptation of cave-dwelling ground beetles in china revealed by geometric morphometry (Coleoptera, Carabidae, Trechini). Insects, 12(11), 1002.

275. Chifman, J., & Kubatko, L. (2014) Quartet inference from SNP data under the coalescent model. Bioinformatics, 30(23), 3317–3324.

276. Chippindale, P. T., Price, A. H., Wiens, J. J., & Hillis, D. M. (2000). Phylogenetic relationships and systematic revision of central Texas hemidactyliine plethodontid salamanders. Herpetological monographs, 1-80.

277. Cieslak, A., Fresneda, J., & Ribera, I. (2014). Life-history specialization was not an evolutionary dead-end in Pyrenean cave beetles. Proceedings of the Royal Society B: Biological Sciences, 281(1781), 20132978.

278. Clark, S. H. (2001). Birth of the mountains: The geologic story of the Southern Appalachian Mountains. US Government Printing Office.

279. Culver, D. C, Master, L. L, Christman, M. C., Hobbs, H. H. (2000) Obligate cave fauna of the 48 contiguous United States. Conservation Biology, 14, 386–401.

280. Culver, D. C. (1982). Cave Life: Evolution and Ecology. Harvard University Press.

281. Culver, D. C., & Pipan, T. 2009. The biology of caves and other subterranean habitats. Oxford: Oxford University Press.

282. Culver, D. C., Christman, M. C., Elliott, W. R., Hobbs, H. H., Reddell, J. R. (2003) The North American obligate cave fauna: regional patterns. Biodiversity & Conservation, 12, 441–468.

283. Culver, D. C., Hobbs III, H.H. (2002) Patterns of species richness in the Florida stygobitic fauna. In: Martin, J.B., Wicks, C.M., Sasowsky, I.D. (Eds.), Hydrogeology and Biology of Post-Paleozoic Carbonate Aquifers. Special Publication 7. Karst Waters Institute, Charlestown, West Virginia, pp. 60–63.

284. Culver, D. C., Kane, T. C., Fong, D. W., Jones, R., & Taylor, M. A. (1990) Morphology of cave organisms: Is it adaptive?. Mémoires de biospéologie, 17(44), 13–26.

285. Derkarabetian, S., Paquin, P., Reddell, J., & Hedin, M. (2022). Conservation genomics of federally endangered Texella harvester species (Arachnida, Opiliones, Phalangodidae) from cave and karst habitats of central Texas. Conservation Genetics, 23(2), 401-416.

286. Derkarabetian, S., Steinmann, D. B., & Hedin, M. (2010) Repeated and time-correlated morphological convergence in cave-dwelling harvestmen (Opiliones, Laniatores) from montane western North America. PLoS One, 5(5), e10388.

287. Devitt, T. J., Wright, A. M., Cannatella, D. C., & Hillis, D. M. (2019). Species delimitation in endangered groundwater salamanders: Implications for aquifer management and biodiversity conservation. Proceedings of the National Academy of Sciences, 116(7), 2624-2633.

288. Dillman, C. B., Bergstrom, D. E., Noltie, D. B., Holtsford, T. P., & Mayden, R. L. (2011). Regressive progression, progressive regression or neither? Phylogeny and evolution of the Percopsiformes (Teleostei, Paracanthopterygii). Zoologica Scripta, 40(1), 45-60.

289. Dooley, K. E., Niemiller, K. D. K., Sturm, N., & Niemiller, M. L. (2022). Rediscovery and phylogenetic analysis of the Shelta Cave Crayfish (Orconectes sheltae Cooper & Cooper, 1997), a decapod (Decapoda, Cambaridae) endemic to Shelta Cave in northern Alabama, USA. Subterranean Biology, 43, 11-31.

290. Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology, 7(1), 1–8.

291. Drummond, A. J., Ho, S. Y. W., Phillips, M. J., & Rambaut, A. (2006). Relaxed phylogenetics and dating with confidence. PLoS Biology, 4(5), e88.

292. Dupin, J., Matzke, N. J., Särkine, T., Knapp, S., Olmstead, R. G., Bohs, L., & Smith, S. D. (2017). Bayesian estimation of the global biogeographical history of the Solanaceae. Journal of Biogeography, 44(4), 887–899.

293. Edwards, S. V. (2009) Is a new and general theory of molecular systematics emerging? Evolution, 63, 1–19.

294. Edwards, S. V., Xi, Z., Janke, A., Faircloth, B. C., McCormack, J. E., Glenn, T. C., Zhong, B., Wu, S., Lemmon, E.M., Lemmon, A.R. and Leaché, A.D. (2016). Implementing and testing the multispecies coalescent model: a valuable paradigm for phylogenomics. Molecular phylogenetics and evolution, 94, 447-462.

295. Faille, A., Andújar, C., Fadrique, F., Ribera, I., (2014) Late Miocene origin of an Ibero-Maghrebian clade of ground beetles with multiple colonizations of the subterranean environment. Journal of Biogeography, 41, 1979–1990.

296. Faille, A., Casale, A., & Ribera, I. (2011) Phylogenetic relationships of western Mediterranean subterranean Trechini ground beetles (Coleoptera: Carabidae). Zoologica Scripta, 40(3), 282–295.

297. Faille, A., Casale, A., Balke, M., & Ribera, I. (2013) A molecular phylogeny of Alpine subterranean Trechini (Coleoptera: Carabidae). BMC Evolutionary Biology, 13(1), 1–16.

298. Faille, A., Ribera, I., Deharveng, L., Bourdeau, C., Garnery, L., Quéinnec, E., Deuve, T., (2010). A molecular phylogeny shows the single origin of the Pyrenean subterranean Trechini ground beetles (Coleoptera: Carabidae). Molecular Phylogenetics and Evolution, 54, 97–106.

299. Faircloth, B. C. (2013). Illumiprocessor: A Trimmomatic Wrapper for Parallel Adapter and Quality Trimming.

300. Faircloth, B. C. (2016) PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics, 32, 786-788.

301. Faircloth, B. C. (2017). Identifying conserved genomic elements and designing universal bait sets to enrich them. Methods in Ecology and Evolution, 8(9), 1103–1112.

302. Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Systematic biology, 61(5), 717-726.

303. Ford, D., & Williams, P. D. (2007). Karst hydrogeology and geomorphology. John Wiley & Sons.

304. Fumey, J., Hinaux, H., Noirot, C., Thermes, C., Rétaux, S., & Casane, D. (2018). Evidence for late Pleistocene origin of Astyanax mexicanus cavefish. BMC evolutionary biology, 18(1), 1-19.

305. García-Machado, E., Hernández, D., García-Debrás, A., Chevalier-Monteagudo, P., Metcalfe, C., Bernatchez, L., & Casane, D. (2011). Molecular phylogeny and phylogeography of the Cuban cave-fishes of the genus Lucifuga: evidence for cryptic allopatric diversity. Molecular Phylogenetics and Evolution, 61(2), 470–483.

306. Gatesy, J. & Springer, M.S. (2014) Phylogenetic analysis at deep timescales: unreliable gene trees, bypassed hidden support, and the coalescence/concatalescence conundrum. Molecular Phylogenetics and Evolution, 80, 231–266.

307. Gómez, R. A., Reddell, J., Will, K., & Moore, W. (2016) Up high and down low: Molecular systematics and insight into the diversification of the ground beetle genus Rhadine LeConte. Molecular Phylogenetics and Evolution, 98, 161–175.

308. Grant, E. H. C., Mulder, K. P., Brand, A. B., Chambers, D. B., Wynn, A. H., Capshaw, G., Niemiller, M.L., Phillips, J.G., Jacobs, J.F., Kuchta, S.R. and Bell, R.C. (2022). Speciation with gene flow in a narrow endemic West Virginia cave salamander (Gyrinophilus subterraneus). Conservation Genetics, 1-18.

309. Gray, H. H. (1991). Origin and history of the Teays drainage system: the view from midstream. Geology and hydrogeology of the Teays-Mahomet Bedrock Valley System. Geological Society of America Special Paper, 258, 43–49.

310. Gustafson, G. T., Baca, S. M., Alexander, A. M., & Short, A. E. (2020). Phylogenomic analysis of the beetle suborder Adephaga with comparison of tailored and generalized ultraconserved element probe performance. Systematic Entomology, 45(3), 552–570.

311. Hack, J. T. (1969) The area, its geology: Cenozoic development of the southern Appalachians. The distributional history of the biota of the southern Appalachians, Part I: Invertebrates: Blacksburg, Virginia Polytechnic Institute, Research Division Monograph, 1, 1–17.

312. Harden, C. W., Caterino, M. S., & Malabad, T. E. (2022). First Records of the Palearctic Species Trechus obtusus Erichson from the Appalachian Region of the Southeastern United States (Coleoptera: Carabidae: Trechinae: Trechini). The Coleopterists Bulletin, 76(1), 61-69.

313. Harris, R.S. (2007). Improved Pairwise Alignment of Genomic DNA. Pennsylvania State University, State College, Pennsylvania.

314. Hart, P. B., Niemiller, M. L., Burress, E. D., Armbruster, J. W., Ludt, W. B., & Chakrabarty, P. (2020). Cave-adapted evolution in the North American amblyopsid

fishes inferred using phylogenomics and geometric morphometrics. Evolution, 74(5), 936-949.

315. Harvey, M. S. (2002). Short-range endemism amongst the Australian fauna: some examples from non-marine environments. Invertebrate Systematics, 16(4), 555-570.

316. Hedin, M. (1997). Molecular phylogenetics at the population/species interface in cave spiders of the Southern Appalachians (Araneae: Nesticidae: Nesticus). Molecular Biology and Evoution,14, 309–324.

317. Hedin, M. (2015). High-stakes species delimitation in eyeless cave spiders (Cicurina, Dictynidae, Araneae) from central Texas. Molecular ecology, 24(2), 346-361.

318. Hedin, M. C. (1997). Speciational history in a diverse clade of habitat-specialized spiders (Araneae: Nesticidae: Nesticus): inferences from geographic-based sampling. Evolution, 51(6), 1929-1945.

319. Hedin, M., & Thomas, S. M. (2010). Molecular systematics of eastern North American Phalangodidae (Arachnida: Opiliones: Laniatores), demonstrating convergent morphological evolution in caves. Molecular Phylogenetics and Evolution, 54(1), 107-121.

320. Hedin, M., Derkarabetian, S., Blair, J., & Paquin, P. (2018). Sequence capture phylogenomics of eyeless Cicurina spiders from Texas caves, with emphasis on US federally-endangered species from Bexar County (Araneae, Hahniidae). ZooKeys, (769), 49.

321. Hillis, D. M., Chamberlain, D. A., Wilcox, T. P., & Chippindale, P. T. (2001). A new species of subterranean blind salamander (Plethodontidae: Hemidactyliini: Eurycea:

Typhlomolge) from Austin, Texas, and a systematic revision of central Texas paedomorphic salamanders. Herpetologica, 266-280.

322. Hobbs, H. H. (2012) Diversity patterns in the United States. Pp. 251–264 in Encyclopedia of Caves, 2nd edition. White WB, Culver DC, eds. Academic Press, Amsterdam.

323. Holsinger, J. R. (1988) Troglobites: the evolution of cave-dwelling organisms. American Scientist, 76(2), 146–153.

324. Holsinger, J. R. (2000) Ecological derivation, colonization, and speciation. In: Wilkens H, Culver D, Humphreys W (Eds.), Subterranean ecosystems. Elsevier, Amsterdam, pp. 399–415.

325. Hosner, P. A., Faircloth, B. C., Glenn, T. C., Braun, E. L. & Kimball, R. T. (2016) Avoiding missing data biases in phylogenomic inference: an empirical study in the land fowl (Aves: Galliformes). Molecular Biology and Evolution, 33, 1110–1125.

326. Jeannel, R. (1924). Biospeologica. Monographie des Bathysciinae. Archives de zoologie experimentale et generale.

327. Jeannel, R. (1926). Monographie des Trechinae. Morphologie comparée et distribution géographique d'un group de Coléoptères. Première Livraison. L' Abeille 32:221–550.

328. Jeannel, R. (1927). Monographie des Trechinae. Morphologie comparée et distribution géographique d'un group de Coléoptères. Deuxième Livraison. L' Abeille 33:1–592.

329. Jeannel, R. (1928). Monographie des Trechinae. Morphologie comparée et distribution géographique d'un group de Coléoptères. Troisième Livraison. L' Abeille 34:1–808.

330. Jeannel, R. (1930). Monographie des Trechinae. Morphologie comparée et distribution géographique d'un group de Coléoptères. Quatrième Livraison.Supplément. L' Abeille 34:59–122.

331. Jeannel, R. (1943). Les fossiles vivants des cavernes (No. 1). Gallimard.

332. Jeannel, R. (1949). Les Coléoptères cavernicoles de la région des Appalaches. Étude systématique. Notes Biospéologiques 4:37–104.

333. Jeffery, W. R. (2009) Regressive evolution in Astyanax cavefish. Annual Review of Genetics, 43: 25–47.

334. Juan, C., & Emerson, B.C. (2010). Evolution underground: shedding light on the diversification of subterranean insects. Journal of Biology, 9(3), 1–5.

335. Juan, C., Guzik, M. T., Jaume, D., Cooper, S. J. B. (2010). Evolution in caves: Darwin's 'wrecks of ancient life' in the molecular era. Molecular Ecology, 19, 3865–3880.

336. Katoh, K. & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution, 30, 772–780.

337. Katz, A. D., Taylor, S. J., & Davis, M. A. (2018). At the confluence of vicariance and dispersal: Phylogeography of cavernicolous springtails (Collembola: Arrhopalitidae, Tomoceridae) codistributed across a geologically complex karst landscape in Illinois and Missouri. Ecology and Evolution, 8(20), 10306-10325.

338. Krekeler, C. H. (1959). Dispersal of cavernicolous beetles. Systematic Zoology, 8, 119-130.

339. Krijgsman, W., Garcés, M., Agustı, J., Raffi, I., Taberner, C., & Zachariasse, W. J. (2000). The 'Tortonian salinity crisis' of the eastern Betics (Spain). Earth and Planetary Science Letters, 181(4), 497-511.

340. Lanfear, R., Calcott, B., Ho, S. Y., & Guindon, S. (2012). PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Molecular biology and evolution, 29(6), 1695–1701.

341. Lanfear, R., Frandsen, P. B., Wright, A. M., Senfeld, T., & Calcott, B. (2017). PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. Molecular Biology and Evolution, 34(3), 772–773.

342. Leray, V. L., Caravas, J., Friedrich, M., & Zigler, K. S. (2019). Mitochondrial sequence data indicate "Vicariance by Erosion" as a mechanism of species diversification in North American Ptomaphagus (Coleoptera, Leiodidae, Cholevinae) cave beetles. Subterranean Biology, 29, 35-57.

343. Leys, R., Cooper, S. J., Strecker, U., & Wilkens, H. (2005). Regressive evolution of an eye pigment gene in independently evolved eyeless subterranean diving beetles. Biology Letters, 1(4), 496–499.

344. Liu, W., Golovatch, S., Wesener, T., & Tian, M. (2017). Convergent evolution of unique morphological adaptations to a subterranean environment in cave millipedes (Diplopoda). PloS one, 12(2), e0170717.

345. Loria, S., Zigler, K., & Lewis, J. (2011). Molecular phylogeography of the troglobiotic millipede Tetracion Hoffman, 1956 (Diplopoda, Callipodida, Abacionidae). International Journal of Myriapodology, 5, 35-48.

346. Maddison, D. R., And, M. D. B., & Ober, K. A. (1999). Phylogeny of carabid beetles as inferred from 18S ribosomal DNA (Coleoptera: Carabidae). Systematic Entomology, 24(2), 103–138.

347. Maddison, D. R., Kanda, K., Boyd, O. F., Faille, A., Porch, N., Erwin, T. L., & Roig-Juñent, S. (2019). Phylogeny of the beetle supertribe Trechitae (Coleoptera: Carabidae): Unexpected clades, isolated lineages, and morphological convergence. Molecular Phylogenetics and Evolution, 132: 151-176.

348. Maddison, W. P. (1997) Gene trees in species trees. Systematic Biology, 46, 525–536.

349. Malabad, T. E., Kosič, F., Katarina, & William D. O. (2021). Review of rare cave beetles of the genus Pseudanophthalmus in Virginia. With a revised status assessment of 17 candidate species (January 14, 2015 through September 8, 2021): Natural Heritage Technical Report 2021-18. Virginia Department of Conservation and Recreation, Division of Natural Heritage, Richmond, Virginia. 221 pp including 5 appendices.

350. Mammola, S. (2019). Finding answers in the dark: caves as models in ecology fifty years after Poulson and White. Ecography, 42(7), 1331-1351.

351. Mammola, S., P. Cardoso, D.C. Culver, *et al*. 2019. Scientists' Warning on the Conservation of Subterranean Ecosystems. BioScience 69: 641-650.

352. Matzke, N. J. (2013) BioGeoBEARS: Biogeography with Bayesian (and Likelihood) Evolutionary Analysis in R Scripts (R package, version 0.2.1).

353. Matzke, N. J. (2014) Model selection in historical biogeography reveals that founder-event speciation is a crucial process in island clades. Systematic Biology, 63, 951–970.

354. McKenney, R., & Jacobson, R. B. (1996). Erosion and deposition at the riffle-pool scale in gravel-bed streams, Ozark Plateaus, Missouri and Arkansas, 1990-95. US Department of the Interior, US Geological Survey.

355. Meiklejohn, K. A., Faircloth, B.C., Glenn, T.C., Kimball, R.T. & Braun, E.L. (2016) Analysis of a rapid evolutionary radiation using ultraconserved elements: evidence for a bias in some multispecies coalescent methods. Systematic Biology, 65, 612–627.

356. Micheels, A., Eronen, J., & Mosbrugger, V. (2009). The Late Miocene climate response to a modern Sahara desert. Global and Planetary Change, 67(3-4), 193-204.

357. Mirarab, S., & Warnow, T. (2015). ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. Bioinformatics, 31(12), i44-i52.

358. Morvan, C., Malard, F., Paradis, E., Lefébure, T., Konecny-Dupré, L., & Douady, C. J. (2013). Timetree of Aselloidea reveals species diversification dynamics in groundwater. Systematic Biology, 62(4), 512-522.

359. Moyle, R. G., Oliveros, C. H., Andersen, M. J., Hosner, P. A., Benz, B. W., Manthey, J. D., Travers, S.L., Brown, R.M. and Faircloth, B.C. (2016). Tectonic collision and uplift of Wallacea triggered the global songbird radiation. Nature Communications, 7(1), 1-7.

360. Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular biology and evolution, 32(1), 268-274.

361. Niemiller, M. L. and Zigler, K. S. (2013). Patterns of cave biodiversity and endemism in the Appalachians and Interior Plateau of Tennessee, USA. PLoS One 8:e64177.

362. Niemiller, M. L., Fitzpatrick, B. M., Shah, P., Schmitz, L., & Near, T. J. (2013). Evidence for repeated loss of selective constraint in rhodopsin of amblyopsid cavefishes (Teleostei: Amblyopsidae). Evolution, 67(3), 732–748.

363. Niemiller, M. L., Helf, K., & Toomey, R. S. (2021). Mammoth cave: a hotspot of subterranean biodiversity in the United States. Diversity, 13(8), 373.

364. Niemiller, M. L., Higgs, D. M., & Soares, D. (2013). Evidence for hearing loss in amblyopsid cavefishes. Biology letters, 9(3), 20130104.

365. Niemiller, M. L., Miller, B. T., Fitzpatrick, B. M., & White, W. B. (2009). Systematics and evolutionary history of subterranean Gyrinophilus salamanders. In Proceedings of the 15th International Congress of Speleology (Vol. 1, pp. 242-248).

366. Niemiller, M. L., Near, T. J., Fitzpatrick, B. M. (2012). Delimiting species using multilocus data: diagnosing cryptic diversity in the southern cavefish Typhlichthys subterraneus (Teleostei: Amblyopsidae). Evolution, 66, 846–866.

367. Niemiller, M. L., Taylor, S. J., Slay, M. E., & Hobbs III, H. H. (2019). Biodiversity in the United States and Canada. In Encyclopedia of Caves (pp. 163-176). Academic Press.

368. Niemiller, M. L., Zigler, K. S., Ober, K. A., Carter, E. T., Engel, A. S., Moni, G., Moni, G., Philips, T.K. and Stephen, C. D. (2017). Rediscovery and conservation

status of six short-range endemic Pseudanophthalmus cave beetles (Carabidae: Trechini). Insect Conservation and Diversity, 10(6), 495-501.

369. Niemiller, M., Fitzpatrick, B., Miller, B. (2008) Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: Gyrinophilus) inferred from gene genealogies. Molecular Ecology, 17, 2258–2275.

370. Ober, K. A., Niemiller, M. L., & Philips, T. K. (2022) Cave trechine (Coleoptera: Carabidae) diversity and biogeography in North America. In: Cave Life – Drivers of Diversity and Diversification (Wynne JJ, ed.), John Hopkins Press.

371. Paquin, P., & Hedin, M. C. (2004) The power and perils of 'molecular taxonomy' a case study of eyeless and endangered Cicurina (Araneae: Dictynidae) from Texas caves. Molecular Ecology, 13, 3239–3255.

372. Peck, S. B. (1998). A summary of diversity and distribution of the obligate cave inhabiting faunas of the United States and Canada. Journal of Cave and Karst Studies, 60:18–26.

373. Phillips, J. G., Fenolio, D. B., Emel, S. L., & Bonett, R. M. (2017). Hydrologic and geologic history of the Ozark Plateau drive phylogenomic patterns in a cave-obligate salamander. Journal of Biogeography, 44(11), 2463-2474.

374. Porter, M. (2007). Subterranean biogeography: what have we learned from molecular techniques? Journal of Cave and Karst Studies, 69: 179–186.

375. Poulson, T. L., & White, W. B. (1969). The Cave Environment: Limestone caves provide unique natural laboratories for studying biological and geological processes. Science, 165(3897), 971-981.

376. R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

377. Racovitza, E. G. (1907). Essai sur les problemes biospeologiques. Archive de zoologie experimentale et generale , v. 6, p. 371–488.

378. Rambaut, A. (2010). FigTree v1.3.1. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh. http://tree.bio.ed.ac.uk/software/figtree/

379. Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Systematic biology, 67(5), 901-904.S

380. Reaz, R., Bayzid, M. S., & Rahman, M. S. (2014). Accurate phylogenetic tree reconstruction from quartets: A heuristic approach. PloS One, 9(8), e104008.

381. Ree, R. H., & Smith, S. A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. Systematic Biology, 57(1), 4–14.

382. Ree, R. H., Moore, B. R., Webb, C. O., & Donoghue, M. J. (2005). A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution, 59(11), 2299–2311.

383. Ribera, I., Fresneda, J., Bucur, R., Izquierdo, A., Vogler, A.P., Salgado, J.M., Cieslak, A., (2010). Ancient origin of a Western Mediterranean radiation of subterranean beetles. BMC Evolutionary Biology, 10, 29.

384. Rizzo, V., Comas, J., Fadrique, F., Fresneda, J., Ribera, I., (2013). Early Pliocene range expansion of a clade of subterranean Pyrenean beetles. Journal of Biogeography, 40, 1861–1873.

385. Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. and Huelsenbeck, J.P., & Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Systematic Biology, 61(3), 539–542.

386. Salgado, J. M., Blas, M. & Fresneda, J. (2008). Coleoptera, Cholevidae. Fauna Iberica, Vol. 31. MNCN, CSIC, Madrid, Spain.

387. Sasowsky, I. D., White, W. B., & Schmidt, V. A. (1995). Determination of stream-incision rate in the Appalachian plateaus by using cave-sediment magnetostratigraphy. Geology, 23(5), 415–418.

388. Shofner, G. A., Mills, H. H., & Duke, J. E. (2001). A simple map index of karstification and its relationship to sinkhole and cave distribution in Tennessee. Journal of Cave and Karst Studies, 63(2), 67-75.

389. Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J., & Birol, I. (2009). ABySS: a parallel assembler for short read sequence data. Genome research, 19(6), 1117–1123.

390. Snowman, C. V., Zigler, K. S., & Hedin, M. (2010). Caves as islands: mitochondrial phylogeography of the cave-obligate spider species Nesticus barri (Araneae: Nesticidae). The Journal of Arachnology, 38(1), 49-56.

391. Soares, D., & Niemiller, M. L. (2020). Extreme adaptation in caves. The Anatomical Record, 303(1), 15-23.

392. Sota, T., Takami, Y., Ikeda, H., Liang, H., Karagyan, G., Scholtz, C., & Hori, M. (2022). Global dispersal and diversification in ground beetles of the subfamily Carabinae. Molecular Phylogenetics and Evolution, 167, 107355.

393. Springer, M. S. & Gatesy, J. (2014). Land plant origins and coalescence confusion. Trends in Plant Science, 19, 267–269.

394. Springer, M. S. & Gatesy, J. (2016). The gene tree delusion. Molecular Phylogenetics and Evolution, 94, 1–33.

395. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, 30(9), 1312–1313.

396. Stern, D. B., Breinholt, J., Pedraza-Lara, C., López-Mejía, M., Owen, C. L., Bracken-Grissom, H., Fetzner Jr, J.W. & Crandall, K. A. (2017). Phylogenetic evidence from freshwater crayfishes that cave adaptation is not an evolutionary dead-end. Evolution, 71(10), 2522-2532.

397. Strecker, U., Hausdorf, B., & Wilkens, H. (2012). Parallel speciation in Astyanax cave fish (Teleostei) in Northern Mexico. Molecular phylogenetics and evolution, 62(1), 62-70.

398. Streicher, J. W., Shulte, J. A. & Wiens, J. J. (2016). How should genes and taxa be sampled for phylogenetic analyses with missing data? An empirical study in iguanian lizards. Systematic Biology, 65, 128–145.

399. Swofford, D. L. (1998) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

400. Talavera, G. & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Systematic Biology, 56, 564–577.

401. Teller, J. T., & Goldthwait, R. P. (1991). The old Kentucky River; a major tributary to the Teays River. Geological Society of America Special Papers, 258, 29–42.

402. Thomson, R. C., Shedlock, A. M., Edwards, S. V. & Shaffer, H. B. (2008). Developing markers for multilocus phylogenetics in non-model organisms: a test case with turtles. Molecular Phylogenetics and Evolution, 49, 514–525.

403. Tierney, S. M., Cooper, S. J., Saint, K. M., Bertozzi, T., Hyde, J., Humphreys, W. F., & Austin, A. D. (2015). Opsin transcripts of predatory diving beetles: a comparison of surface and subterranean photic niches. Royal Society open science, 2(1), 140386.

404. Torres-Dowdall, J., Karagic, N., Plath, M., & Riesch, R. (2018). Evolution in caves: selection from darkness causes spinal deformities in teleost fishes. Biology letters, 14(6), 20180197.

405. Towns, J., Cockerill, T., Dahan, M., Foster, I., Gaither, K., Grimshaw, A., Hazlewood, V., Lathrop, S., Lifka, D., Peterson, G.D. and Roskies, R. (2014). XSEDE: accelerating scientific discovery. Computing in science & engineering, 16(5), 62-74.

406. Valentine, J. M. (1952). New genera of anophthalmid beetles from Cumberland caves (Carabidae, Trechinae). Geological Survey of Alabama: Museum Paper, 34, 1–41.

407. Van Dam, M. H., Lam, A. W., Sagata, K., Gewa, B., Laufa, R., Balke, M., Faircloth, B.C., & Riedel, A. (2017). Ultraconserved elements (UCEs) resolve the phylogeny of Australasian smurf-weevils. PloS one, 12(11), e0188044.

408. Vandel, A., (1964). Biospeologie. La Biologie des Animaux Cavernicoles. Gauthier-Villars, Paris. 619 pp.

409. Ward, B. (2006). Geologic History of South Central Texas. Boerne Chapter, NPSOT.

410. Weary, D. J., & Doctor, D. H. (2014). Karst in the United States: A digital map compilation and database.

411. White, K., Davidson, G. R., & Paquin, P. (2009). Hydrologic evolution of the Edwards Aquifer recharge zone (Balcones fault zone) as recorded in the DNA of eyeless Cicurina cave spiders, south-central Texas. Geology, 37(4), 339–342.

412. White, W. B. (2009). The evolution of Appalachian fluviokarst: competition between stream erosion, cave development, surface denudation, and tectonic uplift. Journal of Cave and Karst Studies, 71(3), 159-167.

413. Wiens, J. J., Chippindale, P. T., Hillis, D. M. (2003). When are phylogenetic analyses misled by convergence? A case study in Texas cave salamanders. Systematic Biology, 52, 501–514.

414. Xi, Z., Liu, L. & Davis, C.C. (2015). Genes with minimal phylogenetic information are problematic for coalescent analyses when gene tree estimation is biased. Molecular Phylogenetics and Evolution, 92, 63–71.

415. Zhang, D., Gao, F., Jakovlić, I., Zou, H., Zhang, J., Li, W. X., & Wang, G. T. (2020). PhyloSuite: an integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Molecular ecology resources, 20(1), 348-355.

Appendix A. Chapter 1 Figures and Tables

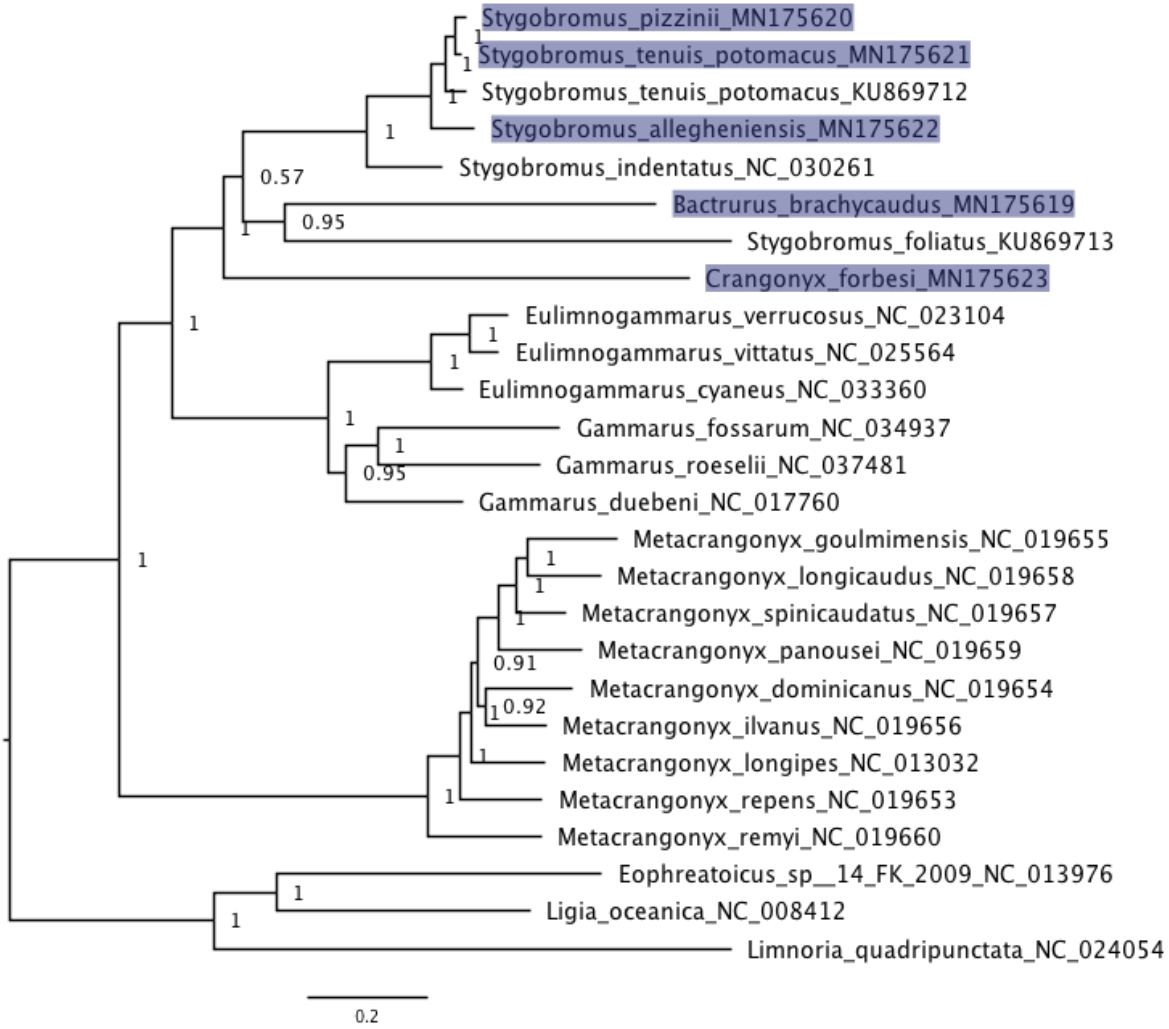
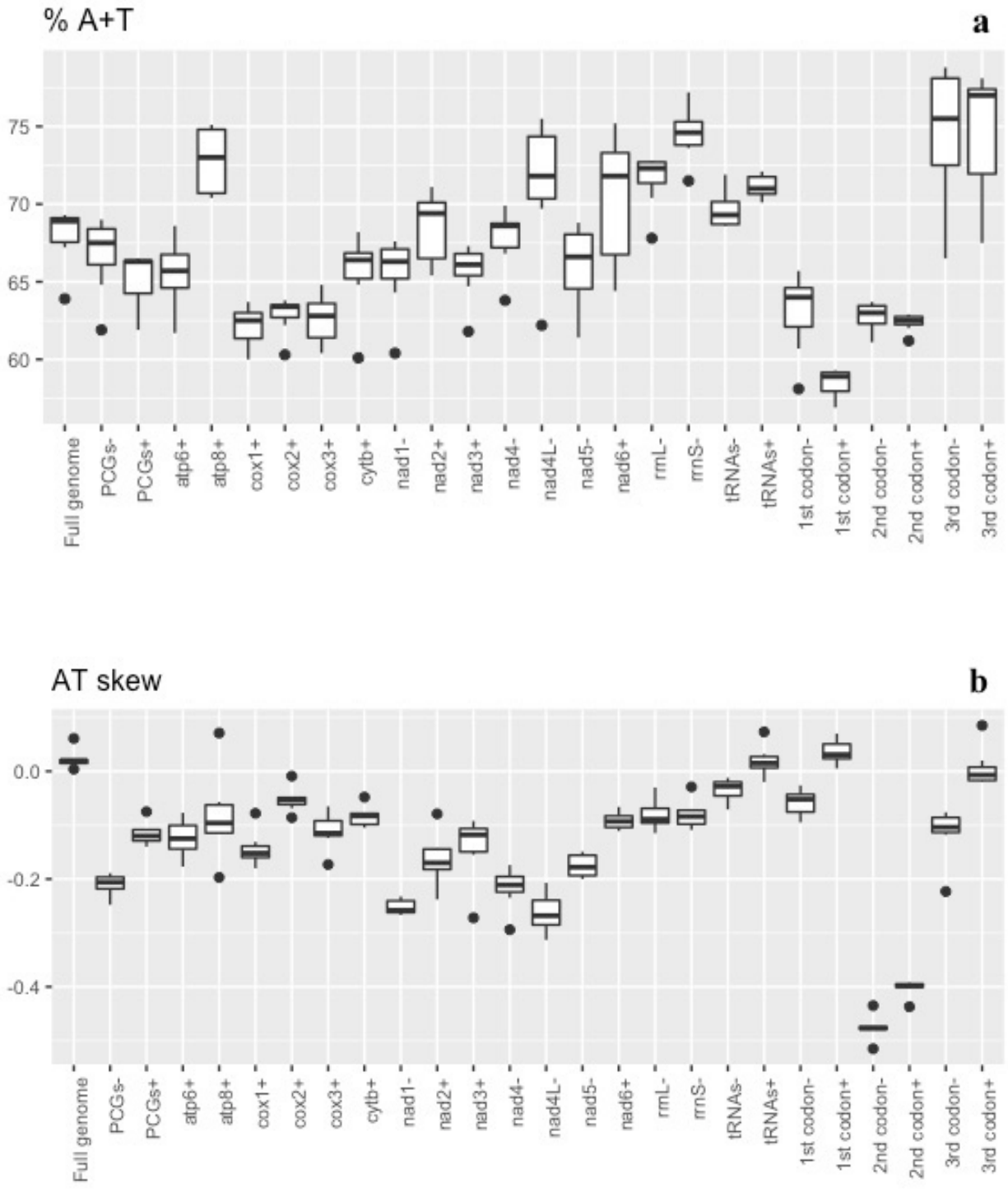


Figure A.1 Bayesian phylogeny of aligned protein-coding loci (3169 aa) for five new amphipod mitogenomes (*Stygobromus allegheniensis*, *S. pizzinii*, *S. tenuis potomacus*, *Bactrurus brachycaudus*, and *Crangonyx forbesi*) in addition to 18 additional amphipod mitogenomes available on Genbank. The three isopods *Ligia oceanica*, *Limnoria quadripunctata*, and *Eophreatoicus sp.14 FK-2009* were included as an outgroup to root the phylogeny. New mitogenomes generated in this study are highlighted. GenBank accession numbers were included as suffix next to the species names. Values at nodes represent posterior probabilities.

Table A.1 List of amphipod mitogenomes, including GenBank accession numbers, taxonomy, and length in bp used for comparative analyses, including new mitogenomes generated in this study (in bold). Partial mitogenome is indicated with an asterisk. Isopod mitogenomes are indicated with a plus sign.

GenBank Accession	Species	Family	Full length (bp)
KU869712	*Stygobromus tenuis potomacus*	Crangonyctidae	14915
KU869713	*Stygobromus foliatus*	Crangonyctidae	15563*
MN175619	***Bactrurus brachycaudus***	**Crangonyctidae**	**14661**
MN175620	***Stygobromus pizzinii***	**Crangonyctidae**	**15176**
MN175621	***Stygobromus tenuis potomacus***	**Crangonyctidae**	**14712**
MN175622	***Stygobromus allegheniensis***	**Crangonyctidae**	**15164**
MN175623	***Crangonyx forbesi***	**Crangonyctidae**	**15469**
NC_008412	*Ligia oceanica*+	Ligiidae	15289
NC_013032	*Metacrangonyx longipes*	Metacrangonyctidae	14113
NC_013976	*Eophreatoicus sp. 14 FK-2009*+	Phreatoicidae	14994
NC_017760	*Gammarus duebeni*	Gammaridae	15651
NC_019653	*Metacrangonyx repens*	Metacrangonyctidae	14355
NC_019654	*Metacrangonyx dominicanus*	Metacrangonyctidae	14543
NC_019655	*Metacrangonyx goulmimensis*	Metacrangonyctidae	14507
NC_019656	*Metacrangonyx ilvanus*	Metacrangonyctidae	14770
NC_019657	*Metacrangonyx spinicaudatus*	Metacrangonyctidae	15037
NC_019658	*Metacrangonyx longicaudus*	Metacrangonyctidae	14711
NC_019659	*Metacrangonyx panousei*	Metacrangonyctidae	14478
NC_019660	*Metacrangonyx remyi*	Metacrangonyctidae	14787
NC_023104	*Eulimnogammarus verrucosus*	Eulimnogammaridae	15315
NC_024054	*Limnoria quadripunctata*+	Limnoriidae	16515
NC_025564	*Eulimnogammarus vittatus*	Eulimnogammaridae	15534
NC_030261	*Stygobromus indentatus*	Crangonyctidae	14638
NC_033360	*Eulimnogammarus cyaneus*	Eulimnogammaridae	14370
NC_034937	*Gammarus fossarum*	Gammaridae	15989
NC_037481	*Gammarus roeselii*	Gammaridae	16073

Appendix B. Chapter 2 Figures and Tables



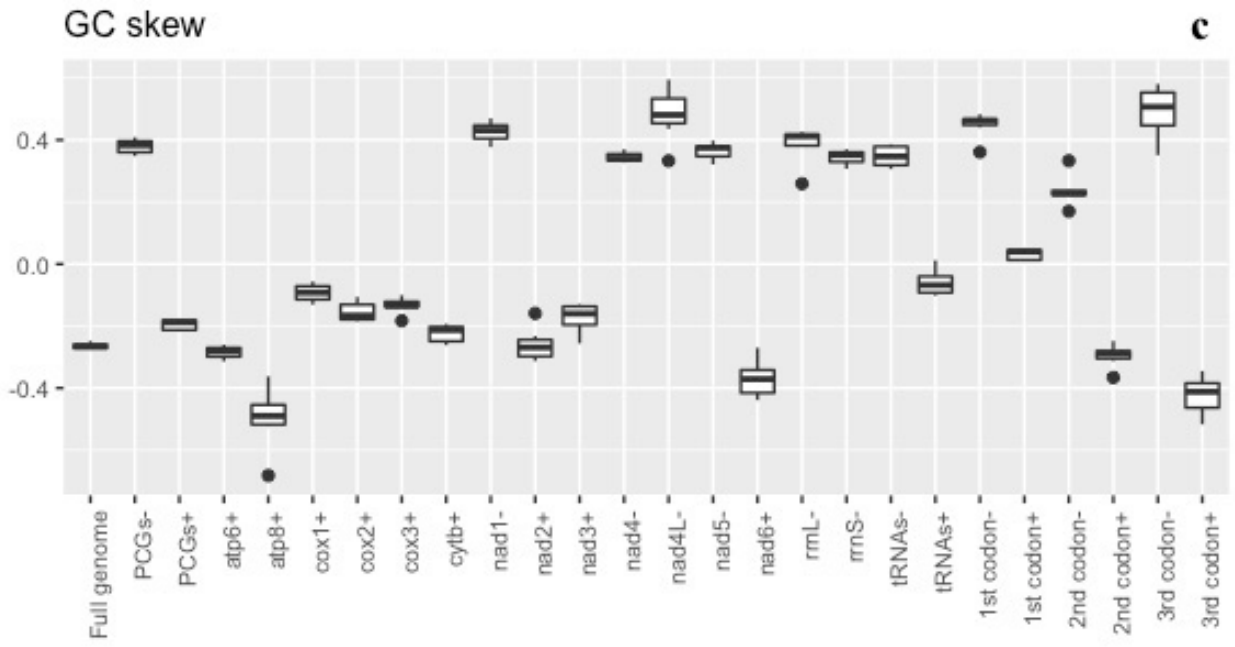


Figure B.1 Crangonyctidae mitochondrial nucleotide composition. Box plots showing values of nucleotide composition (A + T percentage) (a), AT-skew (b), and GC-skew (c) across mitogenomes, protein coding genes (PCG), and ribosomal (rRNA) and transfer ribosomal (tRNA) RNA. The same features are shown for each protein-coding gene and pooled by codon position and coding strand. Genes coded on the (-) strand are represented by a "-" sign and genes coded on the (+) strand are represented by "+" sign at the end of the gene label.

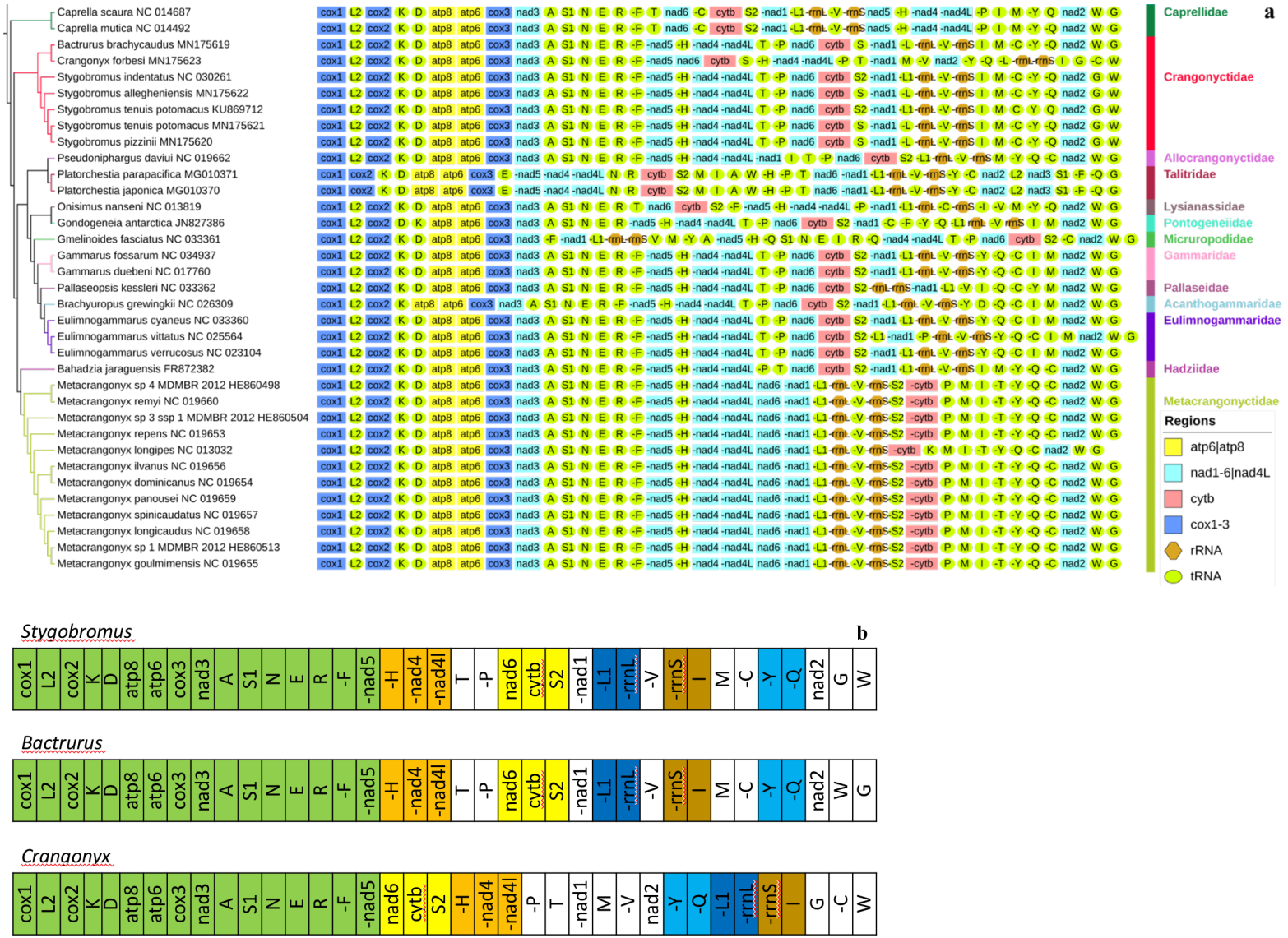


Figure B.2 Mitochondrial phylogenomics and gene orders: (a) Bayesian phylogram inferred using amino acid sequences of all mitochondrial PCGs (left) and gene orders (right). Three isopod outgroups are not shown. GenBank accession numbers are included as suffix next to the species names; (b) gene orders of mitochondrial genomes in three genera of crangonyctid amphipods, including *Stygobromus*, *Bactrurus*, and *Crangonyx*. Conserved gene clusters are indicated by different colors.

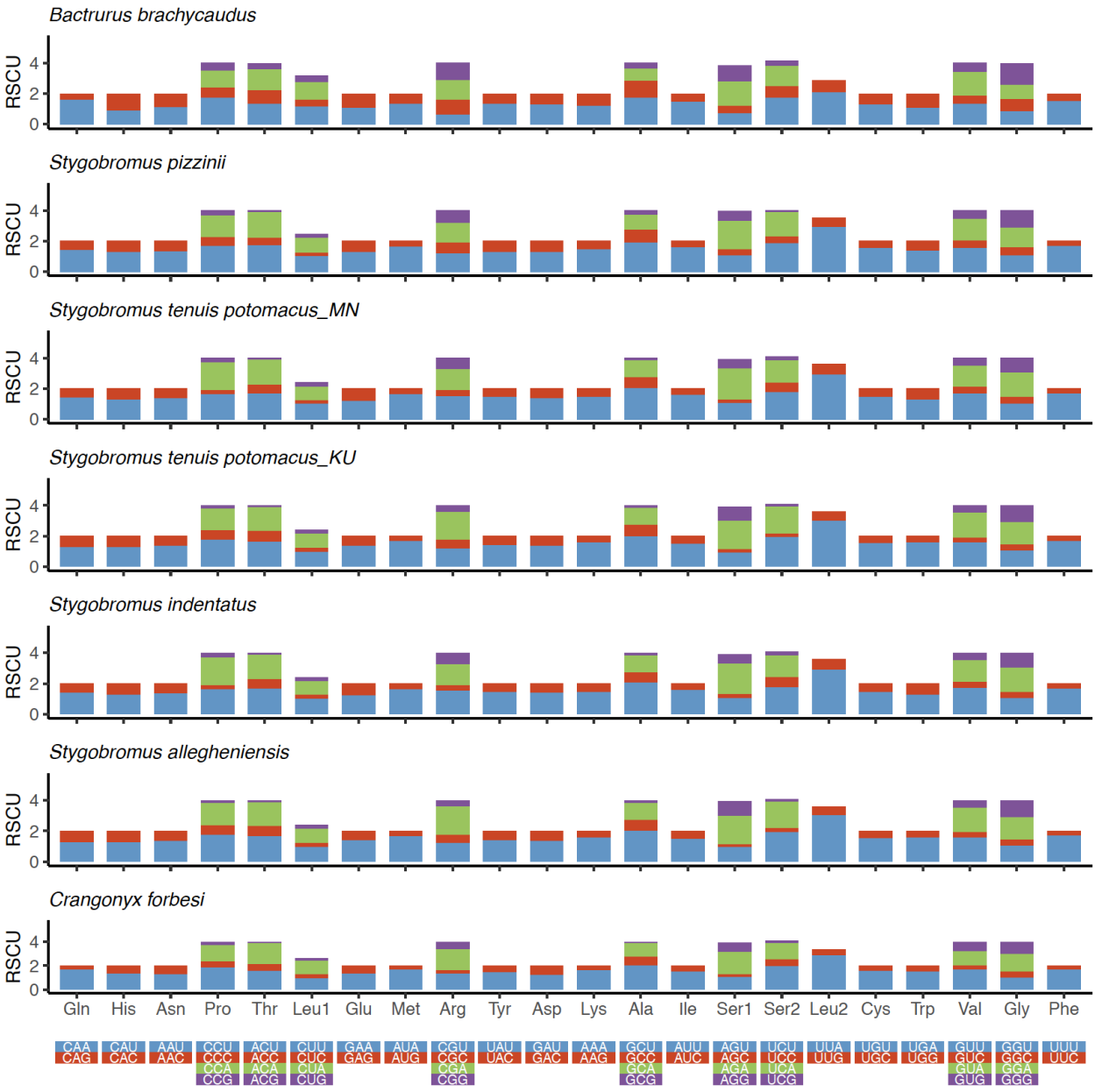


Figure B.3 The relative synonymous codon usage (RSCU) of the mitogenome of all crangonyctid amphipods. The RSCU value are provided on the Y-axis and the codon families are provided on the X-axis.

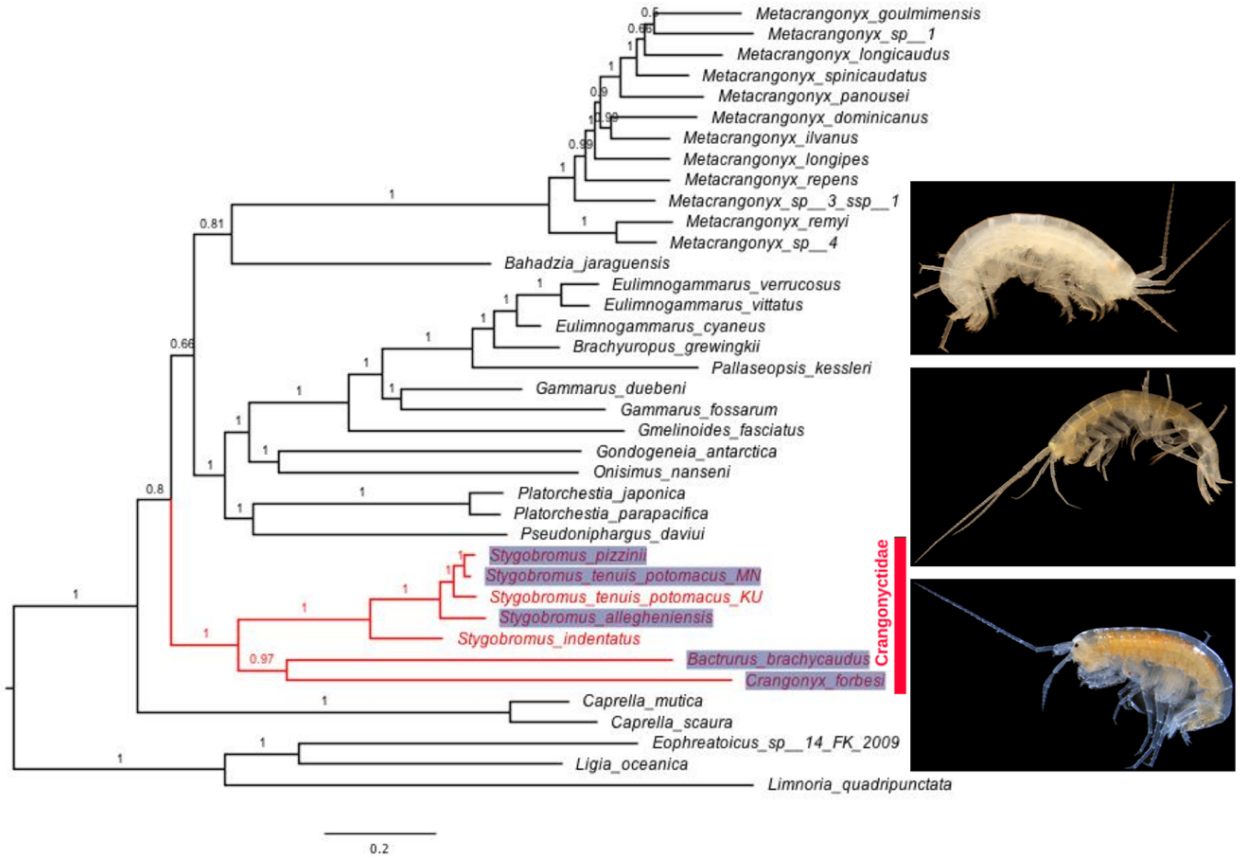


Figure B.4 Bayesian phylogeny of aligned protein-coding loci (3,607 amino acids) for five new amphipod mitogenomes (*Stygobromus allegheniensis, S. pizzinii, S. tenuis potomacus, Bactrurus brachycaudus,* and *Crangonyx forbesi*) in addition to 30 additional amphipod mitogenomes available on Genbank. The three isopods *Ligia oceanica, Limnoria quadripunctata,* and *Eophreatoicus sp.14 FK-2009* are included as an outgroup to root the phylogeny. New mitogenomes generated in this study are highlighted. GenBank accession numbers are included as suffix next to the species names. Values at nodes represent posterior probabilities.

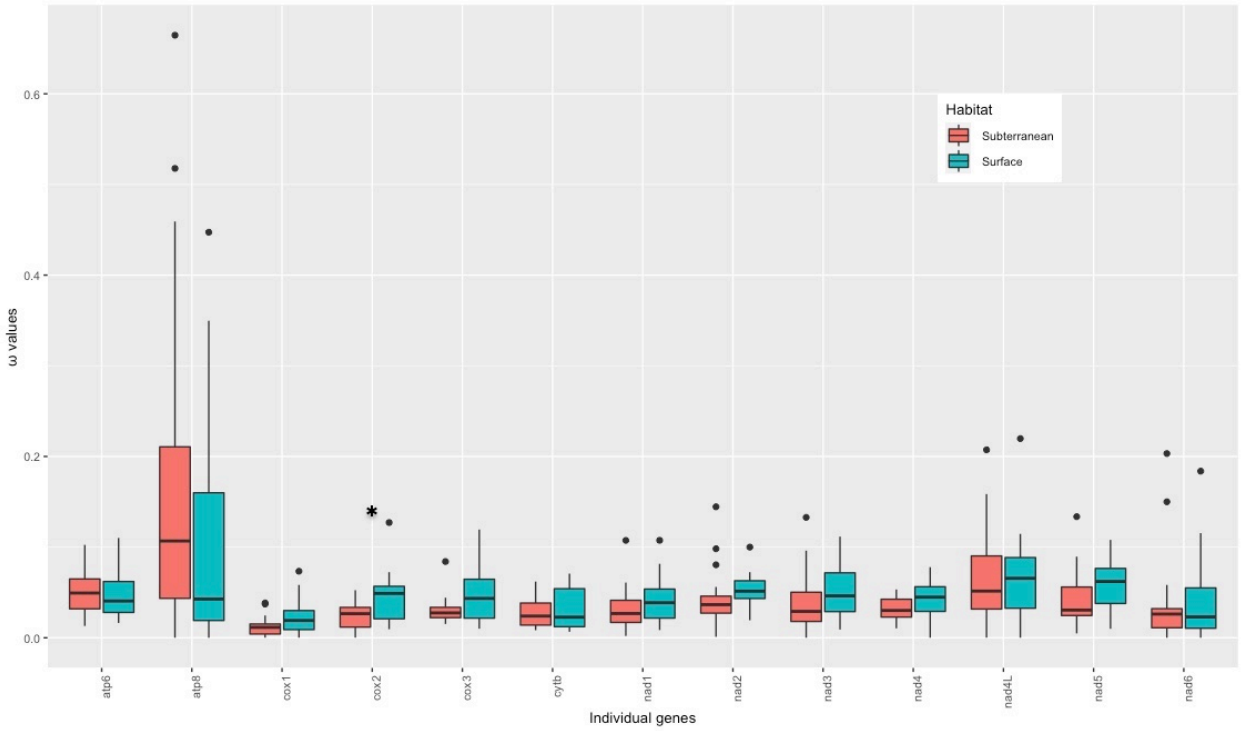


Figure B.5 Ratio of non-synonymous to synonymous substitutions (ω) in the 13 PCGs of subterranean (coral color) and surface (cyan color) amphipods based on the free-ratio model. Boxes include 50% of values; ω is not significantly different between subterranean and surface amphipods for any gene except *cox2** (P value = 0.02).

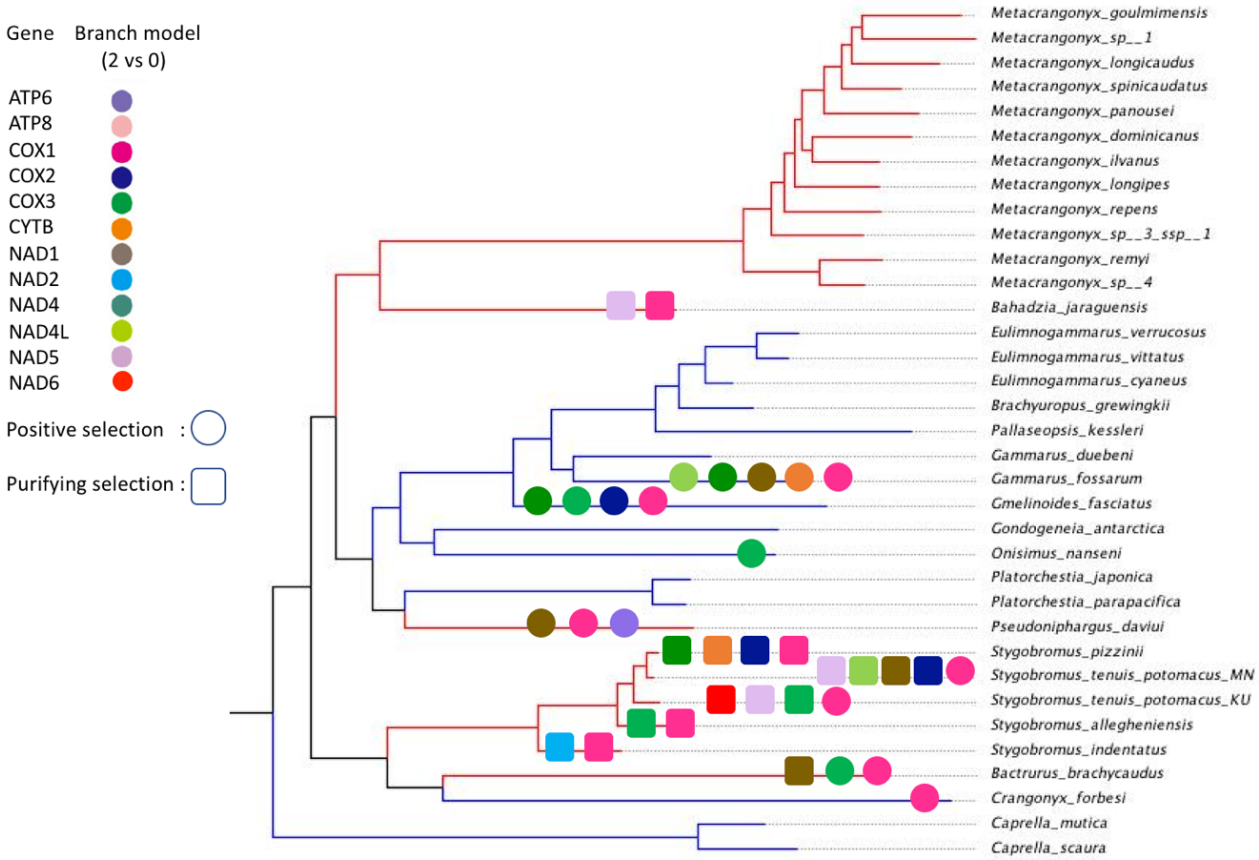


Figure B.6 Results of selective pressure analysis of mitochondrial PCGs with LRT P-value < 0.05 in subterranean and surface-dwelling lineages of amphipods based on branch 2 vs. 0 model. Different colored shapes represent different mitochondrial genes. Squares represent purifying selection and circles represent positive selection. Surface amphipod branches are colored blue and subterranean amphipod branches are colored red.

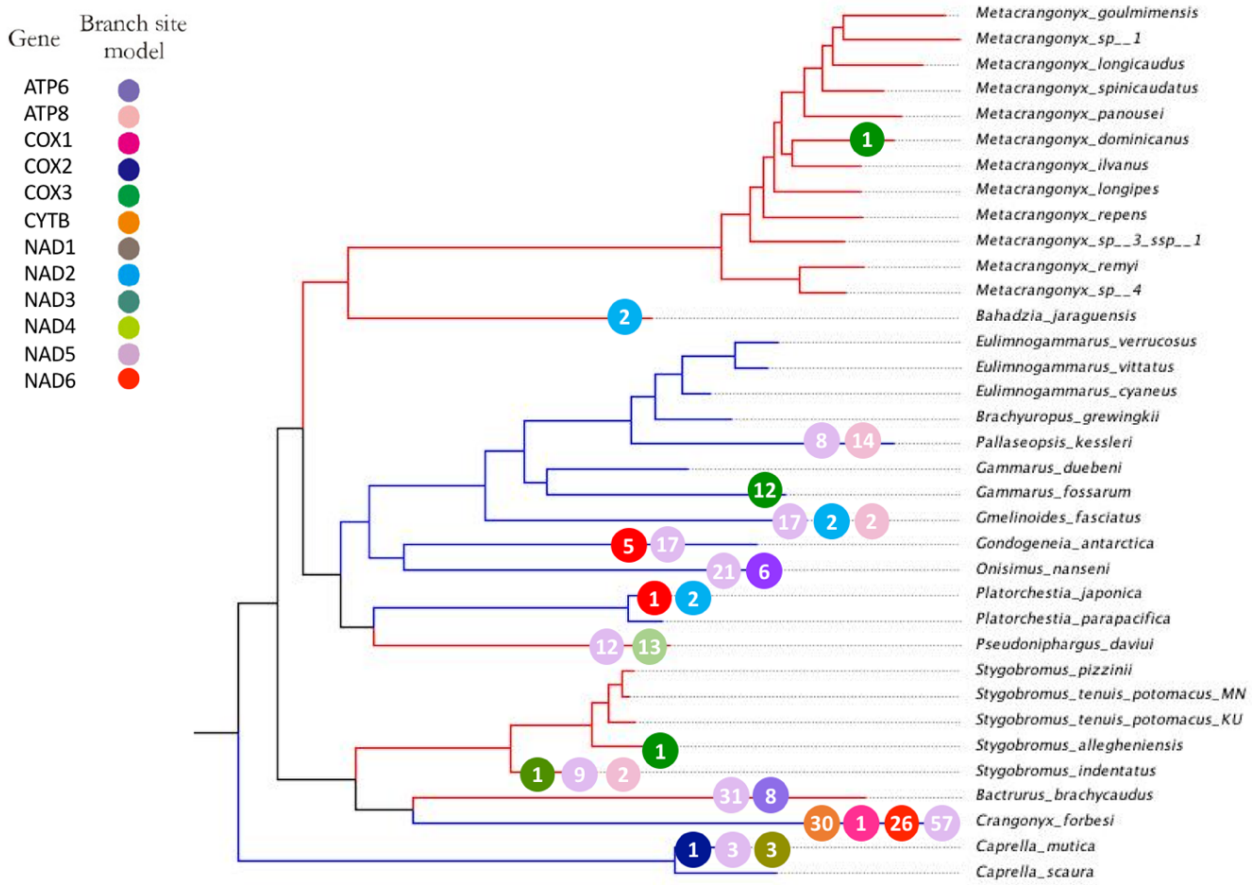


Figure B.7 Evidence of positive selection on the mitochondrial PCGs with LRT P-value < 0.05 and positively selected site (BEB: P≥95%) in subterranean and surface-dwelling lineages of amphipods based on branch-site models. Different colored circles represent different mitochondrial genes. The number within each circle represents the number of positive selection sites detected for the gene. Surface amphipod branches are colored blue and subterranean amphipod branches are colored red.

Table B.1 Summary of mitogenomic characteristics, location, and habitat of subterranean and surface amphipods included for comparative mitogenome analyses.

Accession number	Organism	Family	Full length (bp)	A+T (%)	AT skew	GC skew	Habitat/Locality	Surface vs. Subterranean	References
NC_026309	*Brachyuropus grewingkii*	Acanthogammaridae	17118	62.2	0.003	-0.307	Lake Baikal, deep-water	Surface	Romanova *et al*. 2016
NC_019662	*Pseudoniphargus daviui*	Allocrangonyctidae	15157	68.7	-0.002	-0.314	Spain, well	Subterranean	Bauzà-Ribot *et al*. 2012
NC_014492	*Caprella mutica*	Caprellidae	15427	68.0	-0.023	-0.171	Shallow protected bodies of water in the Sea of Japan	Surface	Kilpert and Podsiadlowski 2010
NC_014687	*Caprella scaura*	Caprellidae	15079	66.4	-0.015	-0.134	Western Indian Ocean	Surface	Ito *et al*. 2010
MN175619	*Bactrurus brachycaudus*	Crangonyctidae	14661	63.9	0.004	-0.258	Fogelpole Cave, Monroe County, Illinois,	Subterranean	Benito *et al*. 2021
MN175623	*Crangonyx forbesi*	Crangonyctidae	15469	67.9	0.061	-0.266	Unidentified spring, Monroe County, Illinois	Surface	Benito *et al*. 2021
MN175622	*Stygobromus allegheniensis*	Crangonyctidae	15164	67.2	0.020	-0.261	Caskey Spring, Berkeley County, West Virginia	Subterranean	Benito *et al*. 2021
NC_030261	*Stygobromus indentatus*	Crangonyctidae	14638	69.3	0.016	-0.270	Fort A.P. Hill, Caroline County, VA, seepage springs	Subterranean	Aunins *et al*. 2016
MN175620	*Stygobromus pizzinii*	Crangonyctidae	15176	68.9	0.014	-0.248	Pimmit Run Seepage Spring, Arlington County, Virginia	Subterranean	Benito *et al*. 2021
KU869712	*Stygobromus tenuis potomacus*	Crangonyctidae	14915	69.1	0.020	-0.275	Fort A.P. Hill, Caroline County, VA, seepage springs	Subterranean	Aunins *et al*. 2016
MN175621	*Stygobromus tenuis potomacus*	Crangonyctidae	14712	69.1	0.022	-0.272	Pimmit Run Seepage Spring, Arlington County, Virginia	Subterranean	Benito *et al*. 2021
NC_033360	*Eulimnogammarus cyaneus*	Eulimnogammaridae	14370	67.6	-0.019	-0.251	Lake Baikal, 0–3.5 m	Surface	Romanova *et al*. 2016
NC_023104	*Eulimnogammarus verrucosus*	Eulimnogammaridae	15315	69.0	-0.008	-0.238	Lake Baikal, 0–12 m	Surface	Rivarola-Duarte *et al*. 2014
NC_025564	*Eulimnogammarus vittatus*	Eulimnogammaridae	15534	67.4	-0.015	-0.222	Lake Baikal, 0–30 m	Surface	Romanova *et al*. 2016
NC_017760	*Gammarus duebeni*	Gammaridae	15651	64.0	-0.016	-0.223	Intertidal zone of the North Atlantic region	Surface	Krebes and Bastrop 2012
NC_034937	*Gammarus fossarum*	Gammaridae	15989	65.2	0.018	-0.261	Europe, freshwater	Surface	Macher *et al*. 2017
FR872382	*Bahadzia jaraguensis*	Hadziidae	14657	69.7	0.037	-0.431	Dominican Rep, cave	Subterranean	Bauzà-Ribot *et al*. 2012
NC_013819	*Onisimus nanseni*	Lysianassidae	14734	70.3	-0.004	-0.198	Below arctic pack ice near the Svalbard archipelago	Surface	Ki *et al*. 2010
NC_019654	*Metacrangonyx dominicanus*	Metacrangonyctidae	14543	73.6	-0.016	-0.026	Dominican Rep, well	Subterranean	Bauzà-Ribot *et al*. 2012
NC_019655	*Metacrangonyx goulmimensis*	Metacrangonyctidae	14507	69.7	-0.016	-0.028	Morocco, well	Subterranean	Bauzà-Ribot *et al*. 2012
NC_019656	*Metacrangonyx ilvanus*	Metacrangonyctidae	14770	74.5	-0.014	-0.012	Italy, well	Subterranean	Bauzà-Ribot *et al*. 2012
NC_019658	*Metacrangonyx longicaudus*	Metacrangonyctidae	14711	75.8	-0.014	-0.051	Morocco, well	Subterranean	Bauzà-Ribot *et al*. 2012
NC_013032	*Metacrangonyx longipes*	Metacrangonyctidae	14113	76.1	-0.017	-0.035	Spain, Cala Figuera cave	Subterranean	Bauzà-Ribot *et al*. 2012
NC_019659	*Metacrangonyx panousei*	Metacrangonyctidae	14478	76.1	-0.012	-0.051	Morocco, well	Subterranean	Bauzà-Ribot *et al*. 2012
NC_019660	*Metacrangonyx remyi*	Metacrangonyctidae	14787	70.8	-0.014	0.017	Morocco, spring at maison forestière	Subterranean	Bauzà-Ribot *et al*. 2012
NC_019653	*Metacrangonyx repens*	Metacrangonyctidae	14355	76.9	-0.025	-0.014	Spain, well	Subterranean	Bauzà-Ribot *et al*. 2012

HE860513	*Metacrangonyx sp. 1 MDMBR-2012*	Metacrangonyctidae	14277	74.4	-0.019	-0.043	Not available	Subterranean	Bauzà-Ribot *et al*. 2012
HE860504	*Metacrangonyx sp. 3 ssp. 1 MDMBR-2012*	Metacrangonyctidae	14644	75.1	-0.062	0.120	Not available	Subterranean	Bauzà-Ribot *et al*. 2012
HE860498	*Metacrangonyx sp. 4 MDMBR-2012*	Metacrangonyctidae	15012	72.6	-0.009	0.005	Not available	Subterranean	Bauzà-Ribot *et al*. 2012
NC_019657	*Metacrangonyx spinicaudatus*	Metacrangonyctidae	15037	74.8	0.010	-0.139	Morocco, well	Subterranean	Bauzà-Ribot *et al*. 2012
NC_033361	*Gmelinoides fasciatus*	Micruropodidae	18114	65.9	-0.001	-0.303	Lake Baikal, 0–192 m	Surface	Romanova *et al*. 2016
NC_033362	*Pallaseopsis kessleri*	Pallaseidae	15759	63.1	0.011	-0.182	Lake Baikal, 0–61 m	Surface	Romanova *et al*. 2016
JN827386	*Gondogeneia antarctica*	Pontogeneiidae	18424	70.1	-0.006	-0.290	Coast of Antarctica, seawater	Surface	Shin *et al*. 2012
MG010370	*Platorchestia japonica*	Talitridae	14780	72.5	0.015	-0.237	Pacific region esp. northeast Asia, terrestrial and supra-littoral habitats	Surface	Yang *et al*. 2017
MG010371	*Platorchestia parapacifica*	Talitridae	14787	74.8	0.011	-0.253	Pacific region esp. northeast Asia, terrestrial and supra-littoral habitats	Surface	Yang *et al*. 2017

Table B.2 Comparison of mitogenomic characteristics of 35 amphipods discussed in this study.

Species	Acession number	PCGs				rRNAs				tRNAs			
		Length (bp)	A+T (%)	AT skew	GC skew	Length (bp)	A+T (%)	AT skew	GC skew	Length (bp)	A+T (%)	AT skew	GC skew
Bactrurus brachycaudus	MN175619	11028	61.9	-0.177	0.090	1705	69.3	-0.030	0.374	1275	69.8	-0.028	0.192
Bahadzia jaraguensis	FR872382	11073	68.7	-0.145	0.109	1788	72.4	-0.076	0.477	1332	71.7	0.005	0.174
Brachyuropus grewingkii	NC_026309	11046	60.2	-0.156	0.067	1608	66.3	-0.074	0.383	1304	65.4	0.012	0.137
Caprella mutica	NC_014492	10989	66.2	-0.195	0.019	1742	72.2	-0.050	0.176	1338	71.8	0.011	0.116
Caprella scaura	NC_014687	10986	64.6	-0.190	0.048	1739	71.7	-0.024	0.149	1318	70.4	-0.015	0.149
Crangonyx forbesi	MN175623	11304	65.9	-0.162	0.065	1785	73.1	-0.072	0.297	1317	71.5	0.001	0.177
Eulimnogammarus cyaneus	NC_033360	11043	67.0	-0.139	0.092	1607	71.8	-0.095	0.377	1300	66.8	0.026	0.132
Eulimnogammarus verrucosus	NC_023104	11019	68.0	-0.141	0.097	1602	69.6	-0.072	0.348	1335	67.1	-0.002	0.123
Eulimnogammarus vittatus	NC_025564	11046	65.7	-0.144	0.072	1606	71.3	-0.072	0.341	1373	66.9	0.020	0.131
Gammarus duebeni	NC_017760	11019	61.6	-0.165	0.074	1623	65.0	-0.037	0.345	1319	63.6	0.029	0.124
Gammarus fossarum	NC_034937	11031	62.7	-0.163	0.073	2614	72.4	-0.022	0.269	1302	66.4	0.011	0.136
Gmelinoides fasciatus	NC_033361	11091	63.5	-0.141	0.081	1594	69.0	-0.031	0.332	1348	66.2	0.025	0.147

Gondogeneia antarctica	JN827386	10794	67.3	-0.158	0.088	800	70.3	-0.007	-0.261	1364	70.0	0.019	0.107
Metacrangonyx dominicanus	NC_019654	11064	72.1	-0.189	0.075	1691	77.6	-0.036	0.232	1301	77.7	0.039	0.177
Metacrangonyx goulmimensis	NC_019655	11067	67.7	-0.175	0.034	1755	75.2	-0.015	0.292	1298	74.8	0.055	0.145
Metacrangonyx ilvanus	NC_019656	11064	72.9	-0.181	0.057	1750	78.1	-0.026	0.260	1306	77.2	0.029	0.211
Metacrangonyx longicaudus	NC_019658	11055	74.8	-0.170	0.070	1757	78.5	-0.002	0.270	1309	77.4	0.017	0.209
Metacrangonyx longipes	NC_013032	11070	75.4	-0.170	0.082	1832	78.8	-0.026	0.263	1287	78.2	0.047	0.204
Metacrangonyx panousei	NC_019659	11025	75.2	-0.174	0.086	1751	78.4	-0.031	0.314	1349	76.5	0.073	0.188
Metacrangonyx remyi	NC_019660	11055	68.6	-0.184	0.044	1728	73.6	0.002	0.189	1296	74.7	0.039	0.198
Metacrangonyx repens	NC_019653	11064	76.0	-0.172	0.101	1752	79.3	-0.022	0.280	1299	79.1	0.057	0.196
Metacrangonyx sp. 1 MDMBR-2012	HE860513	11085	73.4	-0.172	0.077	1758	77.4	-0.011	0.315	1305	76.7	0.040	0.191
Metacrangonyx sp. 3 ssp. 1 MDMBR-2012	HE860504	11076	73.9	-0.167	0.068	1752	78.2	-0.015	0.139	1303	76.8	0.045	0.208
Metacrangonyx sp. 4 MDMBR-2012	HE860498	11076	70.3	-0.190	0.054	1731	75.2	-0.024	0.252	1298	75.4	0.021	0.207
Metacrangonyx spinicaudatus	NC_019657	11067	73.3	-0.155	0.045	1749	77.4	-0.013	0.352	1310	78.0	0.032	0.161
Onisimus nanseni	NC_013819	11046	69.0	-0.170	0.102	1840	76.3	-0.009	0.286	1396	73.4	0.024	0.137
Pallaseopsis kessleri	NC_033362	11028	61.2	-0.147	0.027	1597	64.9	-0.071	0.241	1302	67.8	0.009	0.114
Platorchestia japonica	MG010370	11043	70.9	-0.201	0.109	1610	75.4	-0.069	0.338	1342	76.9	0.027	0.183
Platorchestia parapacifica	MG010371	11043	73.6	-0.185	0.115	1609	77.1	-0.066	0.330	1330	76.9	0.022	0.195
Pseudoniphargus daviui	NC_019662	10998	66.6	-0.176	0.097	1710	73.8	-0.076	0.433	1307	70.2	0.031	0.139
Stygobromus allegheniensis	MN175622	10980	64.6	-0.178	0.106	1722	71.9	-0.112	0.390	1303	69.4	0.004	0.119
Stygobromus indentatus	NC_030261	11100	67.7	-0.149	0.085	1704	74.5	-0.081	0.356	1304	71.5	-0.016	0.170
Stygobromus pizzinii	MN175620	11088	66.9	-0.159	0.086	1715	73.3	-0.086	0.386	1309	70.3	-0.003	0.112
Stygobromus tenuis potomacus	KU869712	11112	67.4	-0.154	0.099	1717	73.2	-0.095	0.383	1300	70.2	-0.005	0.109
Stygobromus tenuis potomacus	MN175621	11091	67.4	-0.163	0.111	1715	74.1	-0.088	0.405	1306	70.2	-0.007	0.129

Table B.3 Likelihood ratio tests of selective pressures (ω ratio) on mitochondrial PCGs between subterranean and surface amphipods. The terminal branches of surface amphipods are assigned as foreground branches and the terminal branches of subterranean amphipods are assigned as foreground branches. PCGs with significant LRT P-value < 0.05 are highlighted in red color.

Model	np	LnL	Estimates of parameters			Model compared	LRT P-value	Omega for Foreground Branch	Gene
Two ratio Model 2	70	-15282.77	ω:	ω0=0.05659	ω1=0.04105	Model 0 vs. Two ratio Model 2	0.0184	ω1=0.04105	*atp6*
Model 0	69	-15285.55	ω=	0.05174					
Two ratio Model 2	70	-3926.02	ω:	ω0=0.12582	ω1=0.14171	Model 0 vs. Two ratio Model 2	0.7207	ω1=0.14171	*atp8*
Model 0	69	-3926.08	ω=	0.13019					
Two ratio Model 2	70	-24614.85	ω:	ω0=0.01988	ω1=0.02323	Model 0 vs. Two ratio Model 2	0.1009	ω1=0.02323	*cox1*
Model 0	69	-24616.19	ω=	0.02099					
Two ratio Model 2	70	-12959.59	ω:	ω0=0.02831	ω1=0.03729	Model 0 vs. Two ratio Model 2	0.0720	ω1=0.03729	*cox2*
Model 0	69	-12961.21	ω=	0.03033					
Two ratio Model 2	70	-16206.66	ω:	ω0=0.04062	ω1=0.04744	Model 0 vs. Two ratio Model 2	0.1739	ω1=0.04744	*cox3*
Model 0	69	-16207.58	ω=	0.04274					
Two ratio Model 2	70	-23399.90	ω:	ω0=0.03742	ω1=0.02615	Model 0 vs. Two ratio Model 2	0.0024	ω1=0.02615	*cytb*
Model 0	69	-23404.50	ω=	0.03337					
Two ratio Model 2	70	-19418.37	ω:	ω0=0.03385	ω1=0.02885	Model 0 vs. Two ratio Model 2	0.2450	ω1=0.02885	*nad1*
Model 0	69	-19419.05	ω=	0.03239					
Two ratio Model 2	70	-19836.87	ω:	ω0=0.04288	ω1=0.04833	Model 0 vs. Two ratio Model 2	0.3826	ω1=0.04833	*nad2*
Model 0	69	-19837.25	ω=	0.04445					
Two ratio Model 2	70	-7403.62	ω:	ω0=0.04379	ω1=0.05134	Model 0 vs. Two ratio Model 2	0.4299	ω1=0.05134	*nad3*
Model 0	69	-7403.93	ω=	0.04560					
Two ratio Model 2	70	-29490.78	ω:	ω0=0.04233	ω1=0.04466	Model 0 vs. Two ratio Model 2	0.6063	ω1=0.04466	*nad4*
Model 0	69	-29490.91	ω=	0.04296					

Two ratio Model 2	70	-6866.17	ω:	ω0=0.03397	ω1=0.04991	Model 0 vs. Two ratio Model 2	0.1902	ω1=0.04991	*nad4l*
Model 0	69	-6867.03	ω=	0.03788					
Two ratio Model 2	70	-42647.82	ω:	ω0=0.06189	ω1=0.05843	Model 0 vs. Two ratio Model 2	0.5037	ω1=0.05843	*nad5*
Model 0	69	-42648.04	ω=	0.06077					
Two ratio Model 2	70	-11163.67	ω:	ω0=0.04176	ω1=0.04888	Model 0 vs. Two ratio Model 2	0.4948	ω1=0.04888	*nad6*
Model 0	69	-11163.91	ω=	0.04373					

Table B.4 Evidence of positive selection on the mitochondrial PCGs of subterranean and surface-dwelling amphipods based on site models.

Model	**np**	**Ln L**	**Estimates of parameters**				**Model compared**	**LRT P-value**	**Positive sites**	**Gene**
M2a	72	-4021.125381	p:	0.32097	0.56570	0.11334	M1a vs. M2a	0.019437929	31 S 0.977*, 37 N 0.997**	*atp8*
			ω:	0.09324	1.00000	2.44333				
M1a	70	-4025.065910	p:	0.37651	0.62349					
			ω:	0.09961	1.00000					
M8	72	-3950.635672	p0=0.85323	p=0.54776	q=3.55619		M8a vs.M8	0.000013424	37 N 0.875	
			(p1= 0.14677)	ω= 1.00000						
M8a	71	-3941.161040	p0=0.97880	p=0.73512	q=3.74996					
			(p1= 0.02120)	ω= 1.00000						
M2a	72	-42238.654620	p:	0.85660	0.08524	0.05815	M1a vs. M2a	1.000000000	482 Q 0.524	*nad5*
			ω:	0.07211	1.00000	1.00000				
M1a	70	-42238.654620	p:	0.85660	0.14340					
			ω:	0.07211	1.00000					
M8	72	-40545.492521	p0=0.98460	p=0.53211	q=7.22545		M8a vs.M8	0.000000000	482 Q 0.856	
			(p1= 0.01540)	ω= 1.00000						
M8a	71	-40454.428911	p0=0.99711	p=0.50047	q=6.79227					
			(p1= 0.00289)	ω= 1.00000						

Table B.5 Selection signals in the mitogenomes of amphipods inferred using aBSREL, BUSTED, and RELAX algorithms. The dataset comprising all 13 concatenated protein-coding genes with 3,607 amino acid sites in the alignment. K column: a statistically significant K > 1 indicates that selection strength has been intensified, and K < 1 indicates that selection strength has been relaxed. LR is likelihood ratio and D indicates the direction of selection pressure change: intensified (I) or relaxed (R), where * highlights a statistically significant ($p < 0.05$) result. Mitogenomes with significant LRT P-value < 0.05 are highlighted in red color.

		aBSREL	BUSTED	RELAX			
Species	**Genes-Sites**	**p-value**	**p-value**	**K**	**p-value**	**LR**	**D**
Crangonyx forbesi	13PCG-3607	0.000	0.011	0.00	0.000	146.46	R*
Gmelinoides fasciatus	13PCG-3607	0.030	0.033	0.24	0.000	174.12	R*
Onisimus nanseni	13PCG-3607	0.001	0.500	0.25	0.000	80.59	R*
Gammarus fossarum	13PCG-3607	0.045	0.500	0.43	0.000	94.94	R*
Caprella mutica	13PCG-3607	0.500	0.500	1.12	0.015	5.97	I*
Eulimnogammarus cyaneus	13PCG-3607	0.500	0.448	2.20	0.000	77.75	I*
Platorchestia japonica	13PCG-3607	0.500	0.500	1.90	0.000	57.20	I*
Gondogeneia antarctica	13PCG-3607	0.500	0.009	0.28	0.000	49.83	R*
Pallaseopsis kessleri	13PCG-3607	0.500	0.500	1.00	0.000	15.68	R*
Brachyuropus grewingkii	13PCG-3607	0.500	0.500	1.03	0.624	0.24	I
Pseudoniphargus daviui	13PCG-3607	0.000	0.020	0.63	0.000	2721.06	R*
Metacrangonyx dominicanus	13PCG-3607	0.500	0.293	0.88	0.015	5.89	R*
Bahadzia jaraguensis	13PCG-3607	0.002	0.500	0.90	0.116	2.47	R
Bactrurus brachycaudus	13PCG-3607	0.000	0.500	0.90	0.119	2.43	R
Stygobromus tenuis potomacus_MN	13PCG-3607	0.500	0.500	48.44	0.000	66.82	I*
Stygobromus tenuis potomacus_KU	13PCG-3607	0.500	0.500	1.17	0.739	0.11	I
Stygobromus allegheniensis	13PCG-3607	0.500	0.500	49.37	0.025	5.02	I*
Stygobromus pizzinii	13PCG-3607	0.500	0.500	1.27	0.000	1852.16	I*
Stygobromus indentatus	13PCG-3607	0.500	0.500	1.28	0.000	16.42	I*

Table B.6 Selection signals in the mitochondrial PCGs of crangonyctid amphipods sequenced in this study inferred using aBSREL, BUSTED, and RELAX algorithms. K column: a statistically significant K > 1 indicates that selection strength has been intensified, and K < 1 indicates that selection strength has been relaxed. LR is likelihood ratio and D indicates the direction of selection pressure change: intensified (I) or relaxed (R), where * highlights a statistically significant ($p < 0.05$) result. PCGs with significant LRT P-value < 0.05 are highlighted in red color.

		Crangonyx forbesi				*Bactrurus brachycaudus*				*Stygobromus tenuis potomacus*				*Stygobromus allegheniensis*				*Stygobromus pizzinii*			
		aBSREL	**BUSTED**	**RELAX**		**aBSREL**	**BUSTED**	**RELAX**		**aBSREL**	**BUSTED**	**RELAX**		**aBSREL**	**BUSTED**	**RELAX**		**aBSREL**	**BUSTED**	**RELAX**	
Gene	**Sites**	**p-value**	**p-value**	**p-value**	**D**	**p-value**	**p-value**	**p-value**	**D**	**p-value**	**p-value**	**p-value**	**D**	**p-value**	**p-value**	**p-value**	**D**	**p-value**	**p-value**	**p-value**	**D**
atp6	218	0.500	0.272	0.001	I*	0.500	0.357	0.648	R	1.000	0.500	1.000	R	0.069	0.500	0.000	I*	1.000	0.500	0.127	I
atp8	50	0.026	0.229	0.625	I	1.000	0.500	0.214	I	0.500	0.500	0.510	R	0.435	0.478	0.142	R	0.444	0.500	0.336	R
cox1	511	0.056	0.146	0.000	R*	0.105	0.500	1.000	R	1.000	0.500	0.000	I*	1.000	0.500	0.150	I	1.000	0.500	0.044	I*
cox2	222	0.500	0.476	0.634	I	0.188	0.280	0.262	R	1.000	0.500	0.017	I*	0.500	0.500	0.359	I	1.000	0.500	0.096	I
cox3	262	1.000	0.500	0.032	R*	0.199	0.029	0.023	R*	1.000	0.500	1.000	R	1.000	0.500	0.000	I*	0.367	0.500	0.018	I*
cytb	377	0.143	0.500	0.000	R*	0.500	0.500	0.112	I	1.000	0.500	0.080	I	0.230	0.500	0.672	I	1.000	0.500	0.536	I
nad1	305	0.093	0.500	0.000	R*	1.000	0.500	0.067	I	1.000	0.500	0.001	I*	0.151	0.500	0.034	I*	1.000	0.500	0.125	I
nad2	321	1.000	0.500	0.386	R	1.000	0.500	0.363	R	1.000	0.500	0.220	I	1.000	0.500	0.002	I*	1.000	0.500	0.141	I
nad3	116	0.414	0.308	0.000	R*	1.000	0.500	0.752	R	1.000	0.500	0.000	I*	0.500	0.500	0.118	I	0.041	0.500	0.149	R
nad4	435	1.000	0.146	0.068	R	1.000	0.500	0.482	I	1.000	0.500	0.076	I	1.000	0.500	0.105	I	1.000	0.500	0.008	I*
nad4l	94	1.000	0.500	0.978	I	1.000	0.500	0.854	I	1.000	0.500	0.018	I*	0.500	0.500	0.980	I	1.000	0.500	1.000	I
nad5	552	1.000	0.500	0.774	I	0.088	0.479	0.024	R*	1.000	0.500	0.005	I*	0.447	0.500	0.012	I*	1.000	0.500	0.003	I*
nad6	144	0.500	0.500	0.250	R	1.000	0.500	0.036	R*	1.000	0.500	0.643	R	1.000	0.500	0.517	I	0.500	0.500	0.680	R

Appendix C. Chapter 2 Supplemental Figures and Tables

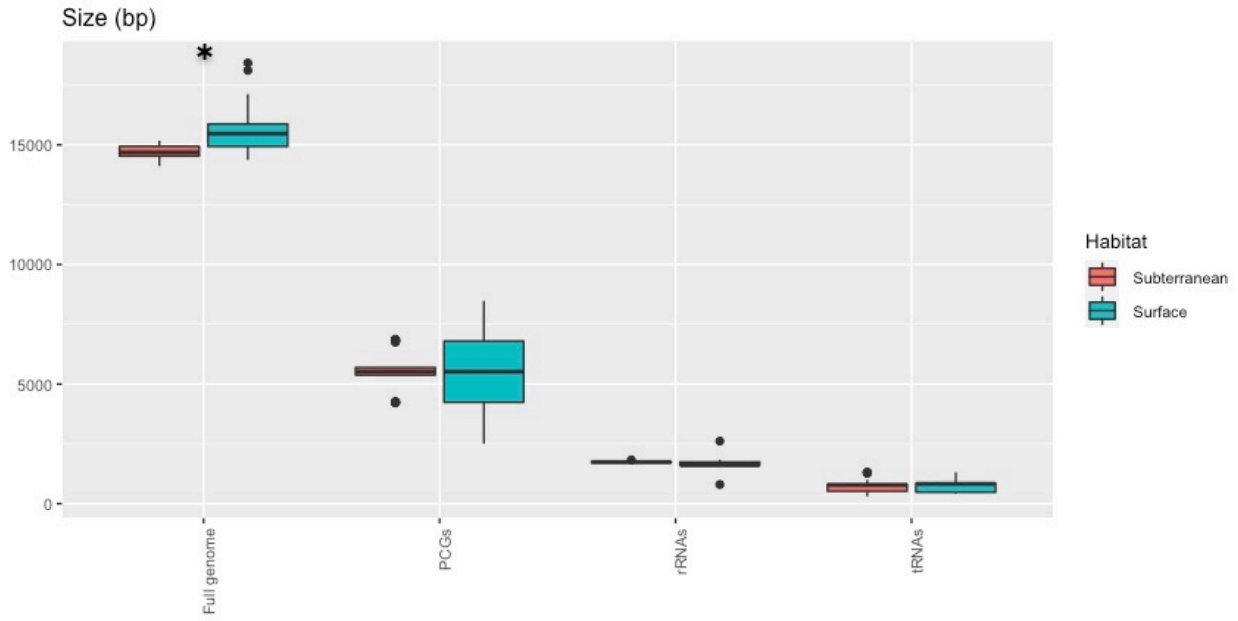


Figure C.1 Box plot showing size (*i.e.*, length) in bp of mitogenomes, protein coding genes (PCG), ribosomal (rRNAs) loci, and transfer ribosomal (tRNAs) loci between subterranean and surface amphipods. Significant P-value < 0.05 is indicated using *.

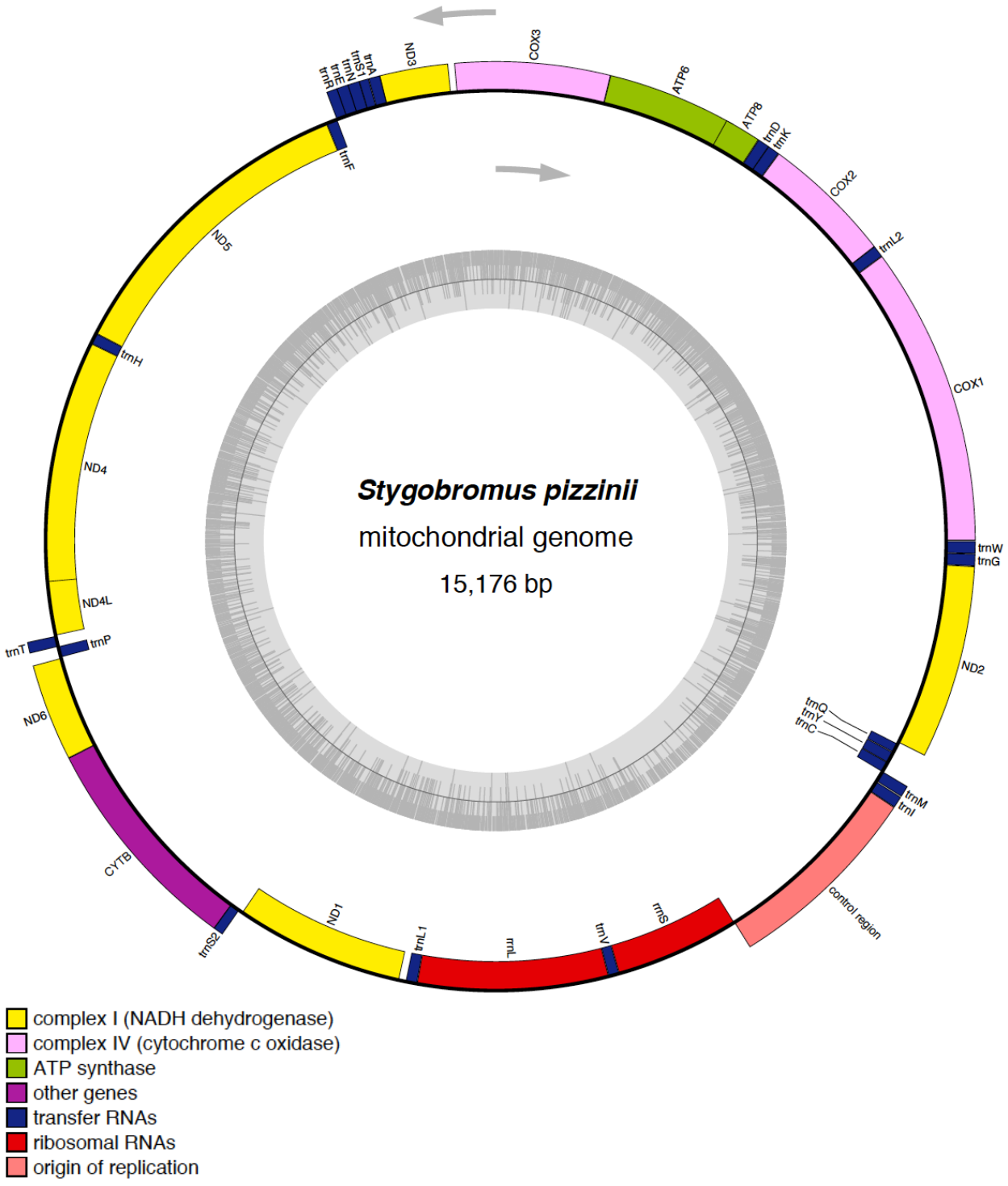


Figure C.2 Map of the mitochondrial genome of *Stygobromus pizzinii*.

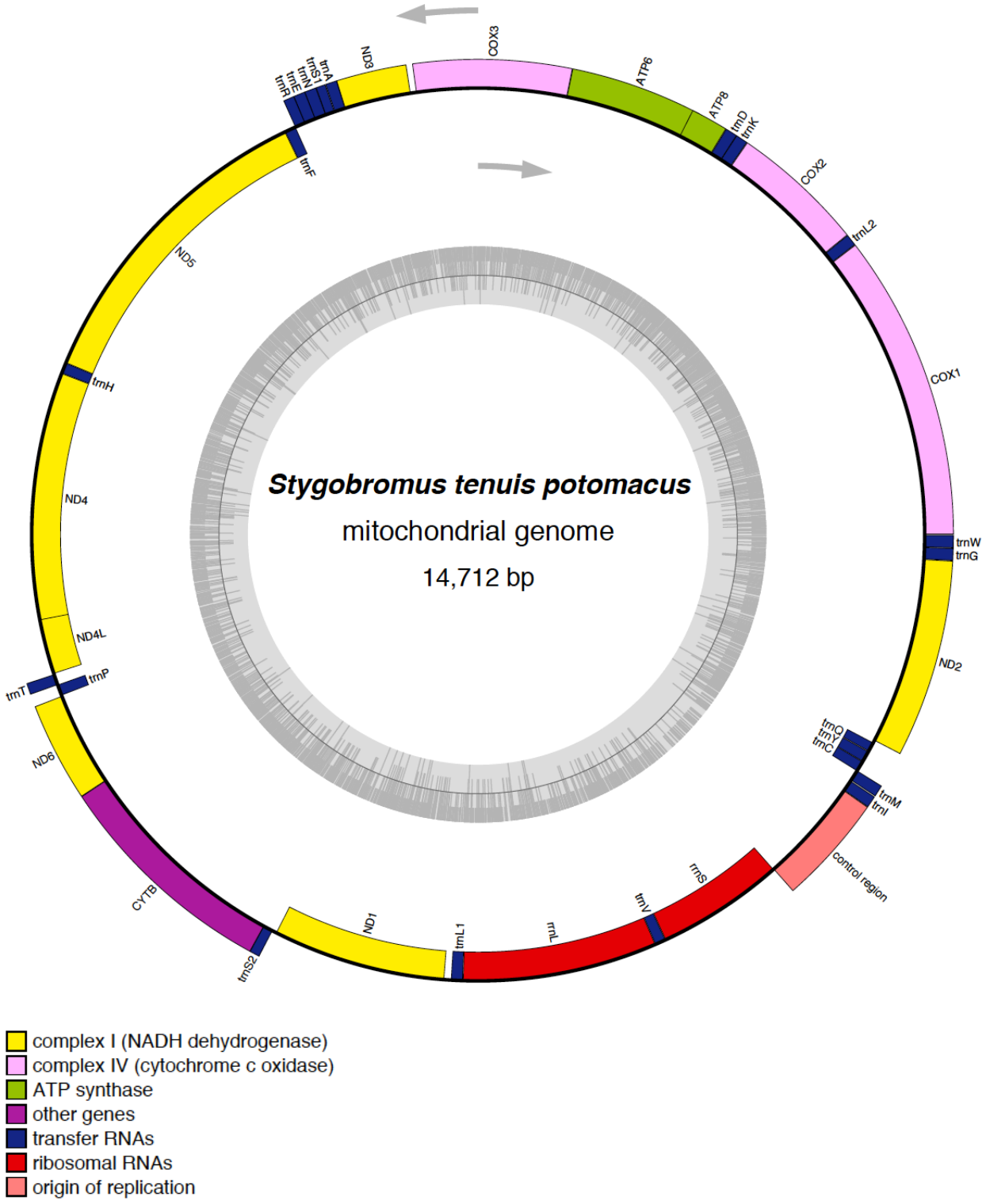


Figure C.3 Map of the mitochondrial genome of *Stygobromus tenuis potomacus*.

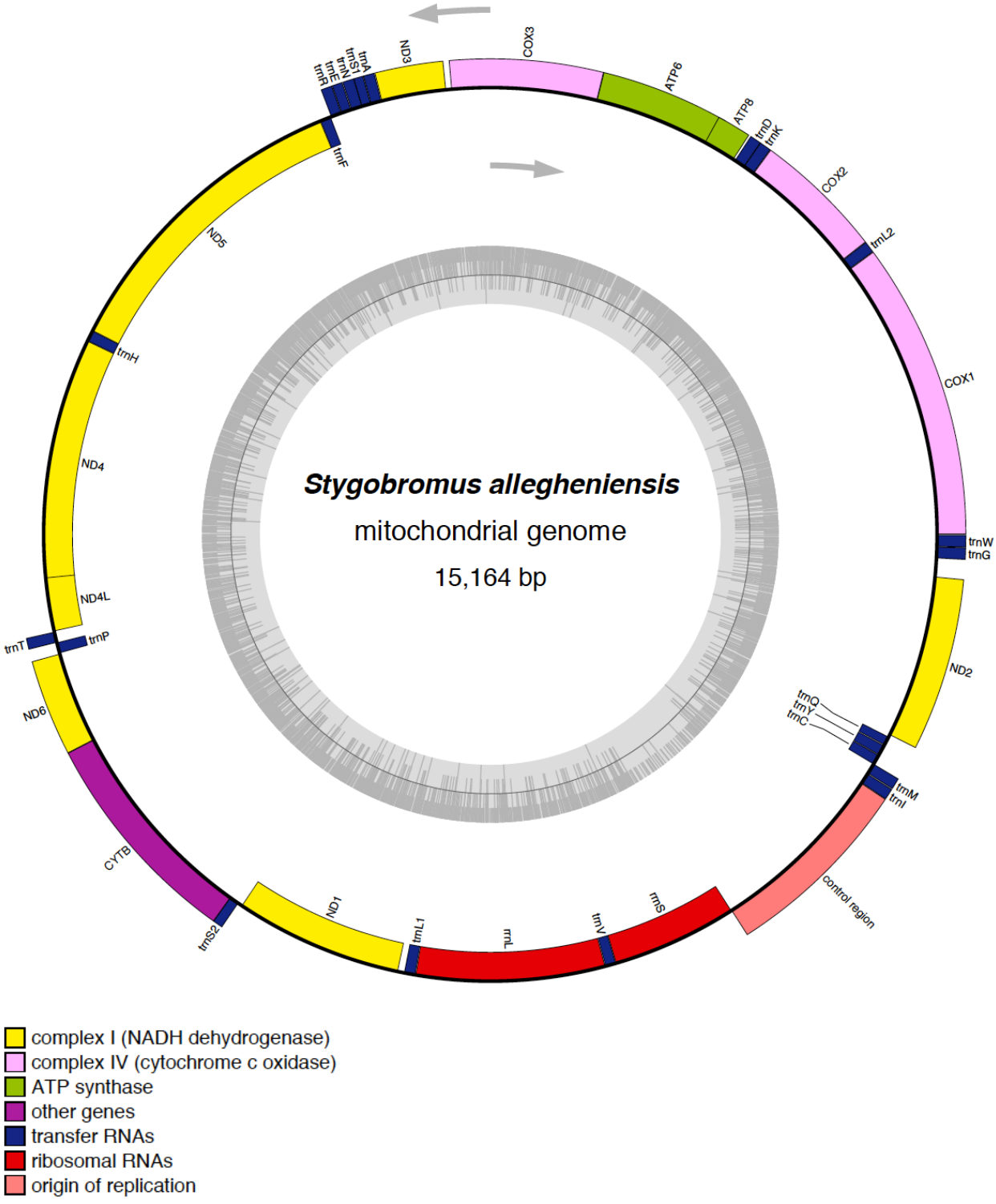


Figure C.4 Map of the mitochondrial genome of *Stygobromus allegheniensis*.

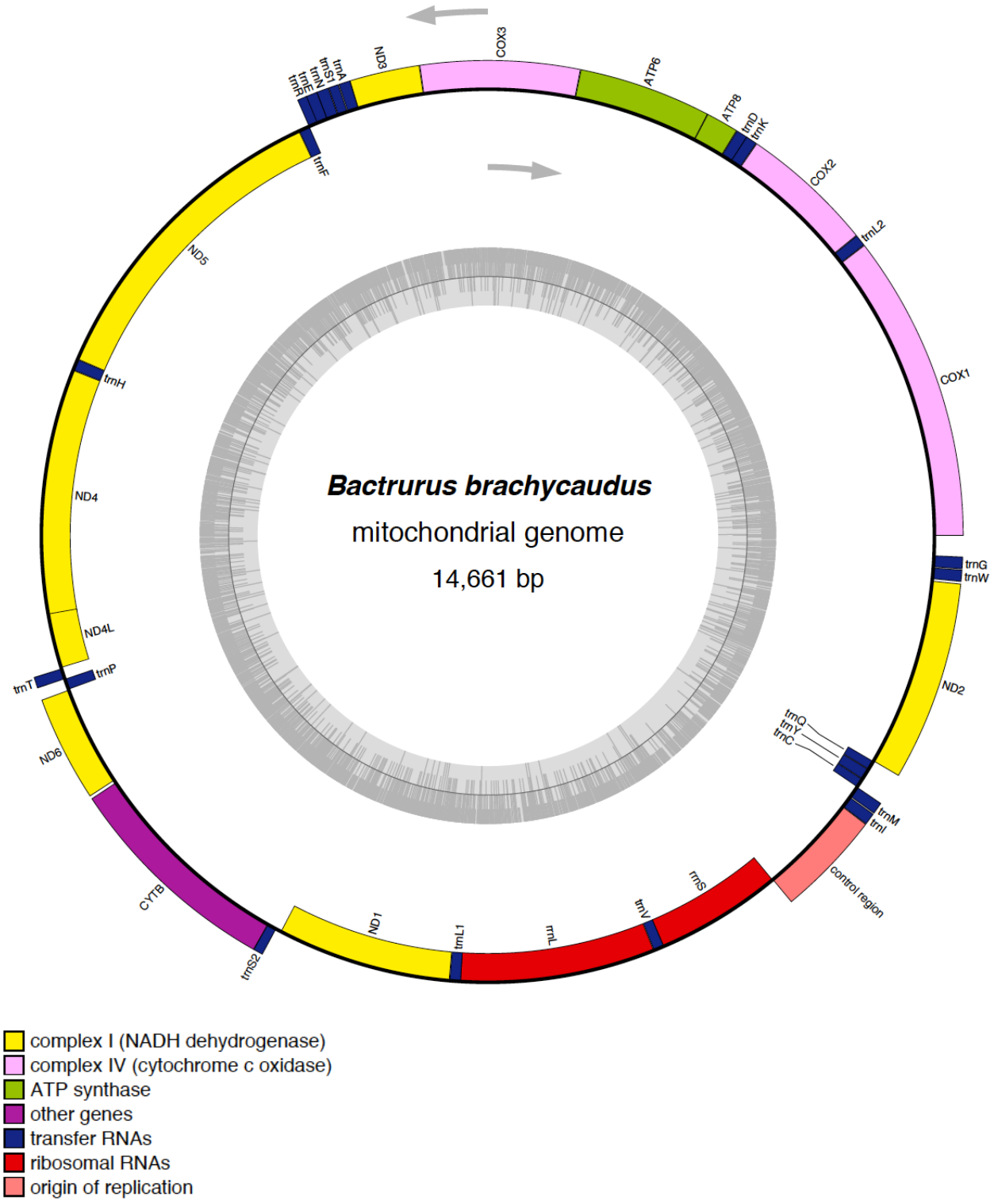


Figure C.5 Map of the mitochondrial genome of *Bactrurus brachycaudus*.

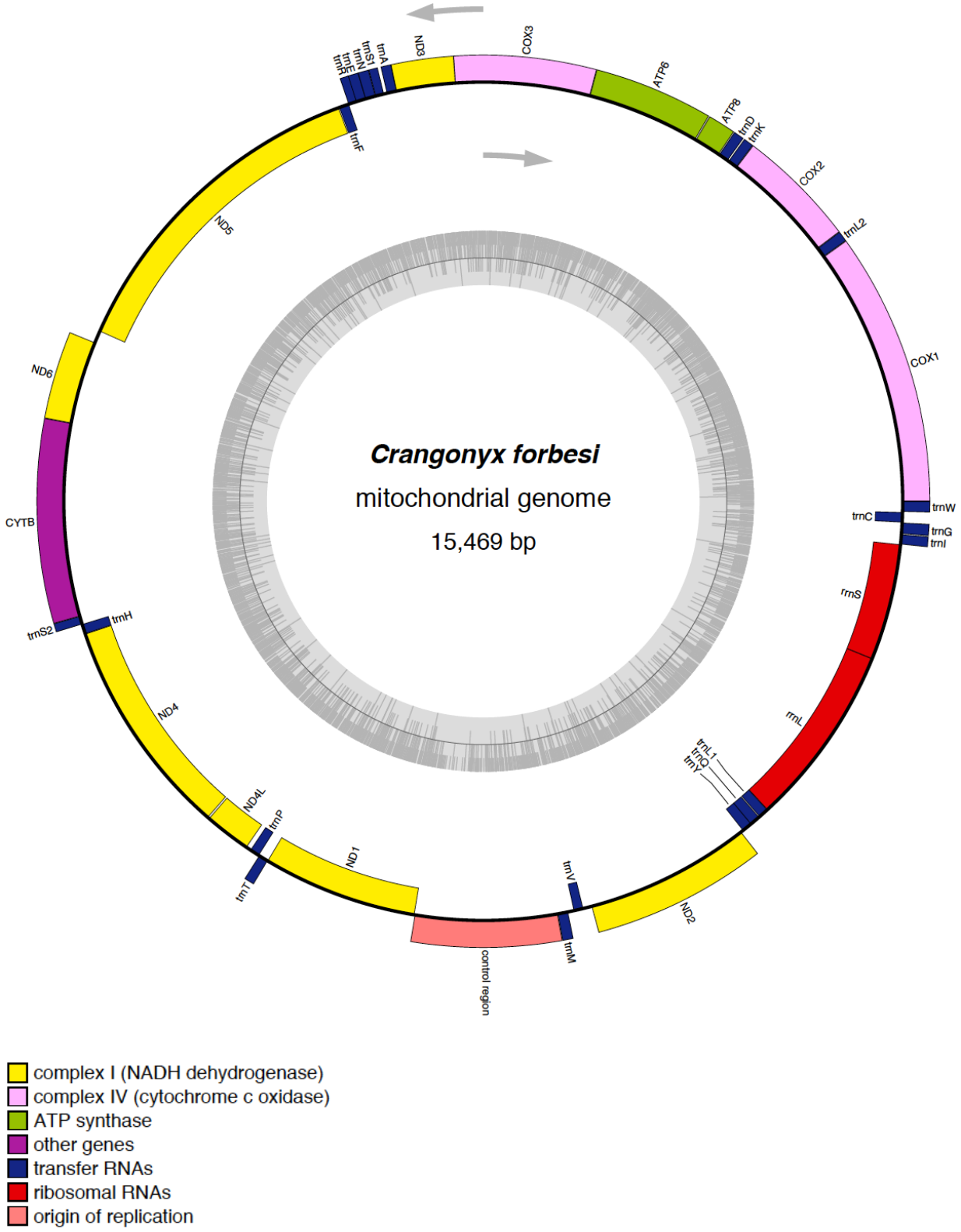


Figure C.6 Map of the mitochondrial genome of *Crangonyx forbesi*.

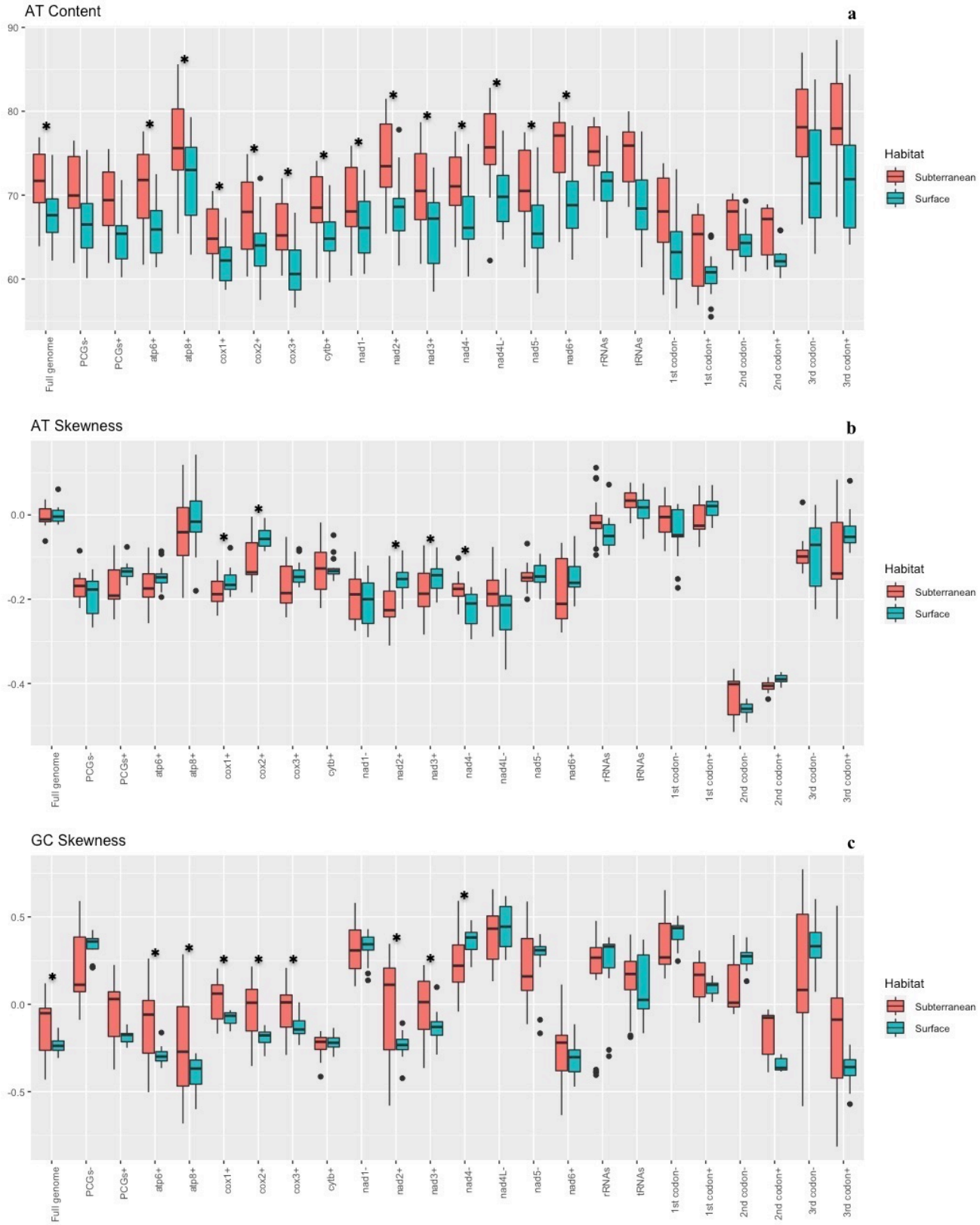


Figure C.7 Box plot showing AT percentage (a), AT-skew (b), and GC-skew (c) between subterranean and surface amphipods across mitogenomes, protein coding genes (PCG), ribosomal (rRNA) loci, and transfer ribosomal (tRNA) loci. The same features are shown for each protein-coding gene and pooled by codon position and coding strand. Genes coded on the (-) strand are represented by a "-" sign and genes coded on the (+) strand are represented by "+" sign at the end of the gene label. Mitogenome and PCG with significant P-value < 0.05 are indicated using *.

Figure C.8 CREx analysis showing the possible scenarios for the evolution of gene rearrangements in the crangonyctid amphipod genus *Stygobromus* from the ancestral pan-crustacean pattern.

Figure C.9 CREx analysis showing the possible scenarios for the evolution of gene rearrangements in the crangonyctid amphipod genus *Bactrurus* from the ancestral pan-crustacean pattern.

Figure C.10 CREx analysis showing the possible scenarios for the evolution of gene rearrangements in the crangonyctid amphipod genus *Crangonyx* from the ancestral pan-crustacean pattern.

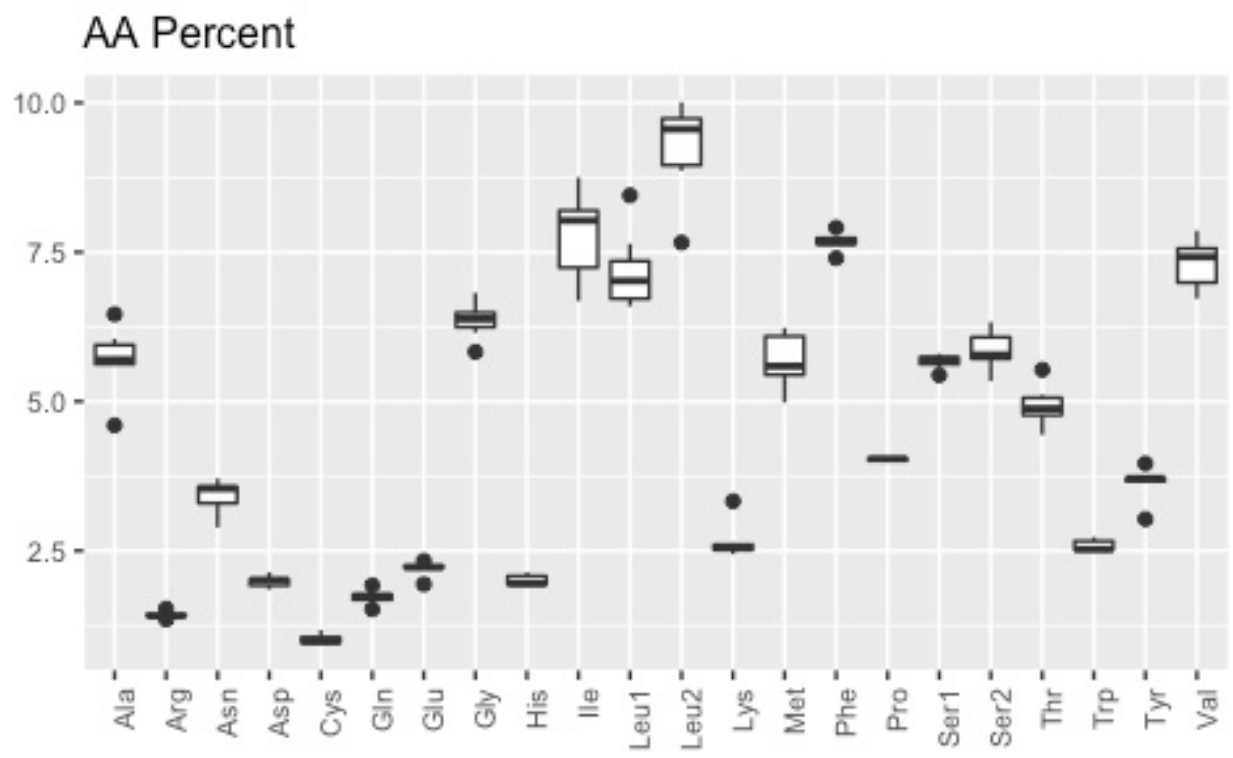
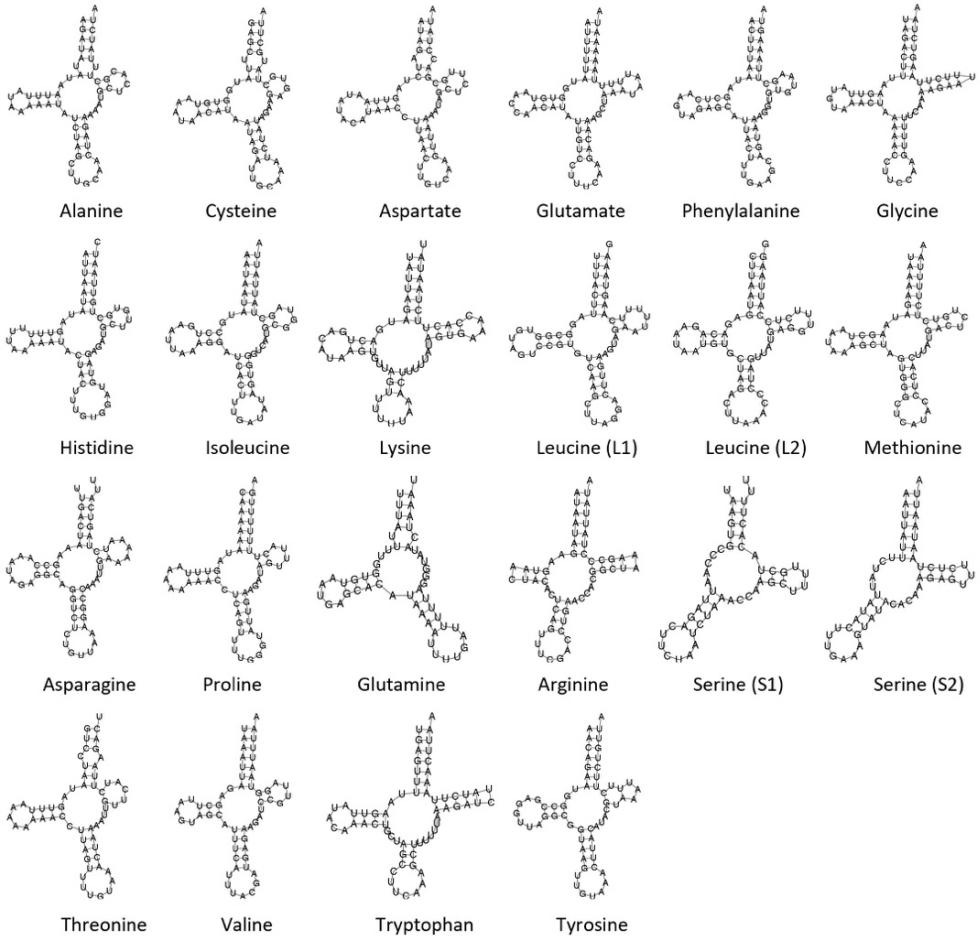
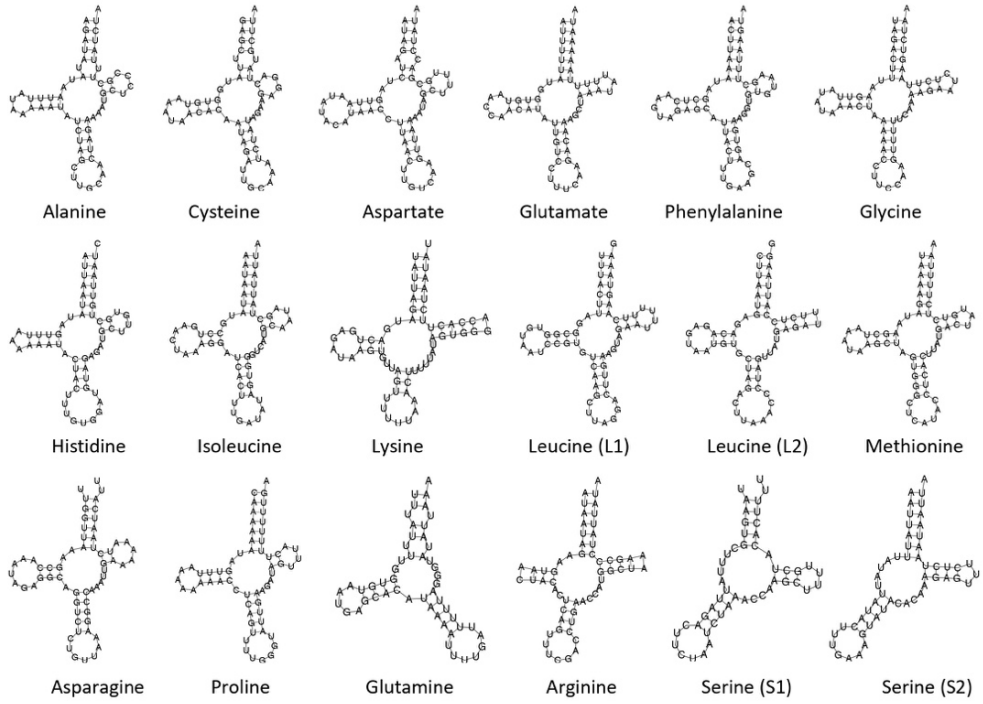


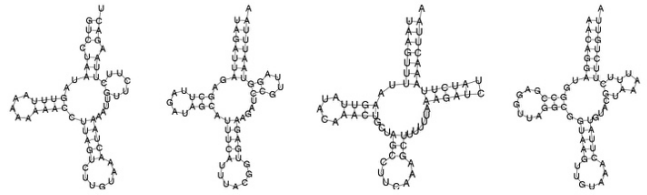
Figure C.11 Box plot showing amino acid composition for mitochondrial protein-coding gene across crangonyctid mitogenomes.

Stygobromus pizzinii

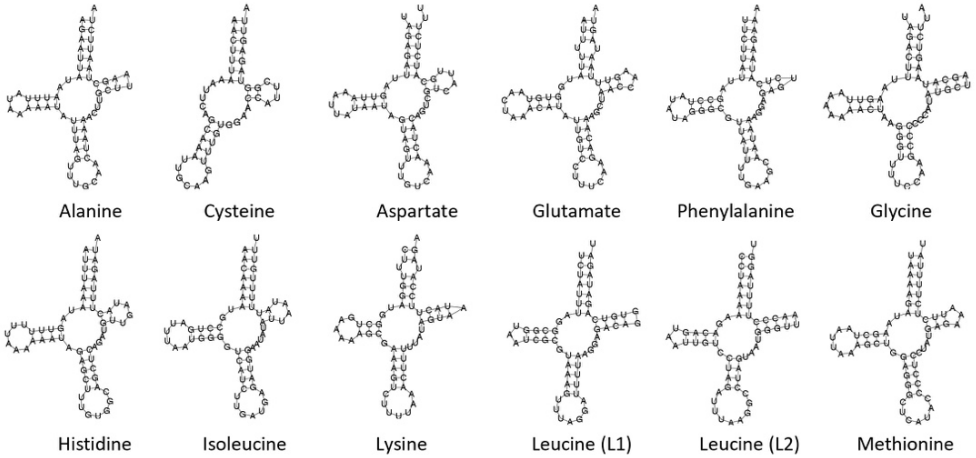


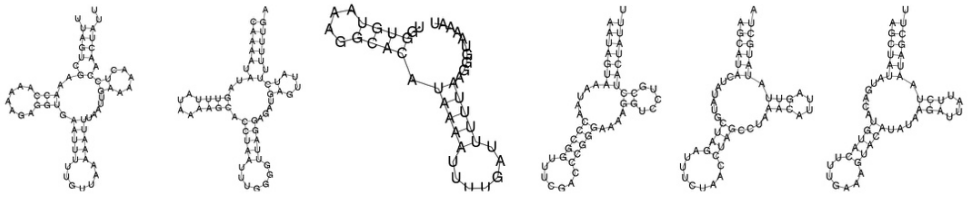
Stygobromus tenuis potomacus

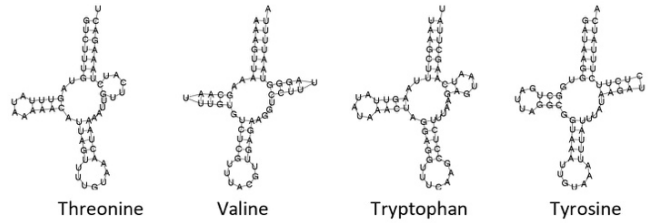




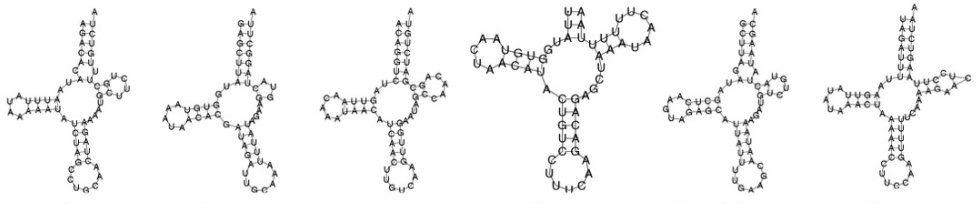
Bactrurus brachycaudus

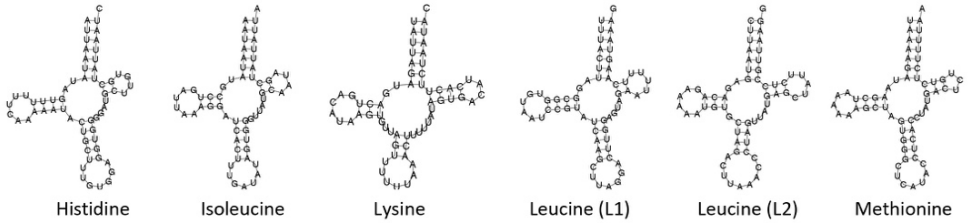






Stygobromus allegheniensis





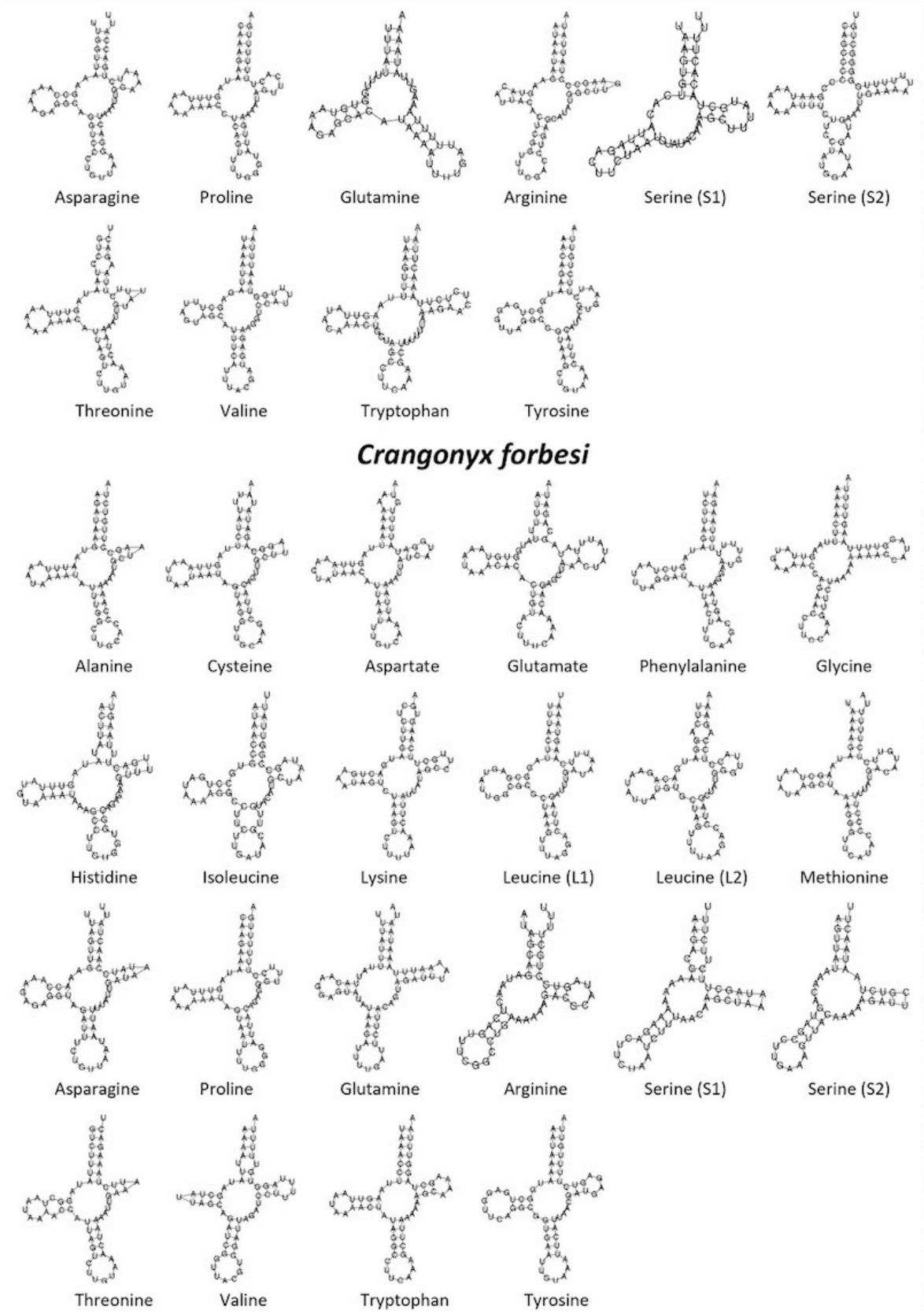


Figure C.12 The predicted mitochondrial tRNAs secondary structures of crangonyctid amphipods under study.

Table C.1 Organization of the mitochondrial genomes of crangonyctid amphipods under study.

Bactrurus brachycaudus

Gene	Position		Size	Intergenic nucleotides	Codon		
	From	To			Start	Stop	Strand
cox1	1	1534	1534		ATC	T	H
trnL2	1535	1594	60				H
cox2	1595	2264	670		ATG	T	H
trnK	2265	2323	59				H
trnD	2320	2379	60	-4			H
atp8	2380	2544	165		ATC	TAA	H
atp6	2544	3206	663	-1	ATG	TAA	H
cox3	3207	4002	796		ATG	T	H
nad3	4003	4353	351		ATC	TAG	H
trnA	4352	4410	59	-2			H
trnS1	4412	4463	52	1			H
trnN	4464	4525	62				H
trnE	4523	4584	62	-3			H
trnR	4578	4627	50	-7			H
trnF	4625	4683	59	-3			L
nad5	4684	6382	1699		ATT	T	L
trnH	6383	6443	61				L
nad4	6444	7755	1312		ATC	T	L
nad4L	7740	8030	291	-16	ATG	TAG	L
trnT	8035	8092	58	4			H
trnP	8092	8150	59	-1			L
nad6	8153	8683	531	2	ATG	TAG	H
cytb	8697	9794	1098	13	ATC	TAG	H
S_copy2	9793	9843	51	-2			H
nad1	9871	10794	924	27	TTG	TAA	L
L_copy2	10795	10854	60				L
rrnL	10851	11884	1034	-4			L
trnV	11885	11941	57				L
rrnS	11939	12609	671	-3			L
control region	12610	13140	531				H
trnI	13141	13202	62				H
trnM	13206	13267	62	3			H
trnC	13267	13316	50	-1			L
trnY	13317	13376	60				L

trnQ	13377	13426	50				L
nad2	13427	14425	999		ATT	TAA	H
trnW	14435	14494	60	9			H
trnG	14497	14558	62	2			H
Overlap:	12	gap:	8				

Stygobromus pizzinii

Gene	**Position**		**Size**	**Intergenic nucleotides**	**Codon**		
	From	**To**			**Start**	**Stop**	**Strand**
cox1	1	1534	1534		ATA	T	H
trnL2	1535	1596	62				H
cox2	1596	2270	675	-1	GTG	TAA	H
trnK	2271	2333	63				H
trnD	2332	2393	62	-2			H
atp8	2394	2582	189		ATC	TAG	H
atp6	2542	3210	669	-41	ATG	TAA	H
cox3	3210	4004	795	-1	ATG	TAA	H
nad3	4036	4386	351	31	ATT	TAG	H
trnA	4385	4443	59	-2			H
trnS1	4444	4493	50				H
trnN	4492	4553	62	-2			H
trnE	4551	4612	62	-3			H
trnR	4610	4666	57	-3			H
trnF	4665	4724	60	-2			L
nad5	4725	6432	1708		ATT	T	L
trnH	6433	6492	60				L
nad4	6493	7813	1321		ATG	T	L
nad4L	7807	8100	294	-7	ATG	TAA	L
trnT	8104	8162	59	3			H
trnP	8162	8221	60	-1			L
nad6	8226	8726	501	4	ATG	TAA	H
cytb	8726	9865	1140	-1	ATG	TAA	H
S_copy2	9864	9916	53	-2			H
nad1	9937	10857	921	20	TTG	TAA	L
L_copy2	10895	10956	62	37			L
rrnL	10957	11993	1037				L
trnV	11994	12051	58				L
rrnS	12052	12729	678				L
control region	12730	13750	1021				H

trnI	13751	13811	61				H
trnM	13815	13875	61	3			H
trnC	13873	13930	58	-3			L
trnY	13931	13991	61				L
trnQ	13993	14044	52	1			L
nad2	14053	15046	994	8	ATA	T	H
trnG	15047	15109	63				H
trnW	15110	15173	64				H
Overlap:	14	gap:	8				

Stygobromus tenuis potomacus

Gene	Position		Size	Intergenic nucleotides	Codon		
	From	To			Start	Stop	Strand
cox1	1	1534	1534		ATG	T	H
trnL2	1535	1595	61				H
cox2	1596	2270	675		GTG	TAA	H
trnK	2271	2333	63				H
trnD	2332	2393	62	-2			H
atp8	2394	2582	189		ATT	TAG	H
atp6	2542	3210	669	-41	ATG	TAA	H
cox3	**3210**	4004	795	-1	ATG	TAA	H
nad3	4034	4387	354	29	ATT	TAG	H
trnA	4386	4444	59	-2			H
trnS1	4445	4494	50				H
trnN	**4493**	4554	62	-2			H
trnE	4552	4612	61	-3			H
trnR	**4610**	4666	57	-3			H
trnF	4665	4724	60	-2			L
nad5	4725	6432	1708		ATG	T	L
trnH	6433	6490	58				L
nad4	6491	7811	1321		ATG	T	L
nad4L	7805	8098	294	-7	ATG	TAA	L
trnT	8102	8160	59	3			H
trnP	8160	8219	60	-1			L
nad6	8223	8723	501	3	ATG	TAA	H
cytb	8723	9862	1140	-1	ATG	TAA	H
S_copy2	9861	9913	53	-2			H
nad1	9935	10855	921	21	TTG	TAG	L
L_copy2	10891	10953	63	35			L

rrnL	**10954**	11989	1036				L
trnV	11990	12047	58				L
rrnS	12048	12726	679				L
control region	12727	13282	556				H
trnI	**13283**	13343	61				H
trnM	13347	13407	61	3			H
trnC	13405	13462	58	-3			L
trnY	13463	13523	61				L
trnQ	13525	13576	52	1			L
nad2	13589	14582	994	12	ATA	T	H
trnG	14583	14645	63				H
trnW	14646	14709	64				H
Overlap:	13	gap:	8				

Stygobromus allegheniensis

Gene	**Position**		**Size**	**Intergenic nucleotides**	**Codon**		
	From	**To**			**Start**	**Stop**	**Strand**
cox1	1	1534	1534		ATC	T	H
trnL2	1535	1596	62				H
cox2	1597	2271	675		ATG	TAG	H
trnK	2272	2334	63				H
trnD	2333	2393	61	-2			H
atp8	2403	2582	180	9	ATA	TAA	H
atp6	2542	3210	669	-41	ATG	TAA	H
cox3	3210	4002	793	-1	ATG	T	H
nad3	4031	4384	354	28	ATT	TAA	H
trnA	4386	4444	59	1			H
trnS1	4445	4495	51				H
trnN	4494	4554	61	-2			H
trnE	4556	4610	55	1			H
trnR	4611	4670	60				H
trnF	4673	4732	60	2			L
nad5	4733	6440	1708		GTG	T	L
trnH	6441	6501	61				L
nad4	6502	7822	1321		ATG	T	L
nad4L	7816	8109	294	-7	ATG	TAG	L
trnT	8113	8172	60	3			H
trnP	8172	8231	60	-1			L
nad6	8235	8735	501	3	ATG	TAA	H

cytb	8735	9874	1140	-1	ATG	TAA	H
S_copy2	9873	9926	54	-2			H
nad1	9949	10869	921	22	TTG	TAG	L
L_copy2	10904	10965	62	34			L
rrnL	10967	12000	1034	1			L
trnV	12002	12060	59	1			L
rrnS	12061	12748	688				L
control region	12749	13739	991				H
trnI	13740	13799	60				H
trnM	13802	13862	61	2			H
trnC	13860	13916	57	-3			L
trnY	13918	13976	59	1			L
trnQ	13979	14029	51	2			L
nad2	14041	14935	895	11	ATA	T	H
trnG	15035	15097	63	99			H
trnW	15098	15161	64				H
Overlap:	9	gap:	16				

Crangonyx forbesi

Gene	**Position**		**Size**	**Intergenic nucleotides**	**Codon**		
	From	**To**			**Start**	**Stop**	**Strand**
cox1	1	1534	1534		ATT	T	H
trnL2	1535	1595	61				H
cox2	1596	2271	676		ATT	T	H
trnK	2272	2330	59				H
trnD	2337	2397	61	6			H
atp8	2401	2559	159	3	ATG	TAA	H
atp6	2564	3235	672	4	ATG	TAA	H
cox3	3235	4033	799	-1	ATG	T	H
nad3	4031	4381	351	-3	ATT	TAG	H
trnA	4380	4438	59	-2			H
trnS1	4456	4507	52	17			H
trnN	4507	4568	62	-1			H
trnE	4566	4631	66	-3			H
trnR	4622	4672	51	-10			H
trnF	4671	4730	60	-2			L
nad5	4734	6707	1974	3	ATT	TAA	L
nad6	6779	7273	495	71	ATG	TAA	H
cytb	7273	8418	1146	-1	ATG	TAA	H

S_copy2	8418	8467	50	-1			H
trnH	8467	8526	60	-1			L
nad4	8527	9838	1312		ATG	T	L
nad4L	9847	10125	279	8	ATA	TAA	L
trnP	10153	10211	59	27			L
trnT	10214	10272	59	2			H
nad1	10272	11195	924	-1	GTG	TAA	L
control region	11198	12040	843	2			H
trnM	12041	12102	62				H
trnV	12137	12194	58	34			L
nad2	12245	13231	987	50	ATA	TAA	H
trnY	13245	13307	63	13			L
trnQ	13304	13368	65	-4			L
L_copy2	13372	13433	62	3			L
rrnL	13433	14522	1090	-1			L
rrnS	14521	15215	695	-2			L
trnI	15216	15274	59				H
trnG	15278	15343	66	3			H
trnC	15345	15405	61	1			L
trnW	15408	15469	62	2			H
Overlap:	14	gap:	17				

Table C.2 Summary of putative start codons in PCG of mitochondrial genomes of all amphipods.

		Putative start codon												
Species	**Accession**	***atp6***	***atp8***	***cox1***	***cox2***	***cox3***	***cytb***	***nad1***	***nad2***	***nad3***	***nad4***	***nad4L***	***nad5***	***nad6***
Stygobromus pizzinii	MN175620	ATG	ATC	ATA	GTG	ATG	ATG	TTG	ATA	ATT	ATG	ATG	ATT	ATG
Stygobromus tenuis potomacus	MN175621	ATG	ATT	ATG	GTG	ATG	ATG	TTG	ATA	ATT	ATG	ATG	ATG	ATG
Bactrurus brachycaudus	MN175619	ATG	ATC	ATC	ATG	ATG	ATC	TTG	ATT	ATC	ATC	ATG	ATT	ATG
Stygobromus allegheniensis	MN175622	ATG	GTA	ATC	ATG	ATG	ATG	TTG	ATA	ATT	ATG	ATG	GTG	ATG
Crangonyx forbesi	MN175623	ATA	ATG	ATT	ATT	ATG	ATG	GTG	ATT	ATT	ATG	ATA	ATT	ATG
Gondogeneia antarctica	JN827386.1	ATG	ATT	ATG	ATA	ATG	ATG	ATT	ATT	ATA	ATG	ATG	ATT	ATG
Gmelinoides fasciatus	NC_033361.1	ATG	ATC	ATA	TTG	ATG	ATG	ATA	TTG	ATG	GTG	ATG	ATG	ATG
Brachyuropus grewingkii	NC_026309.1	ATG	GTG	ATT	TTG	ATG	ATG	TTG	TTG	ATG	ATG	ATG	ATT	GTG
Gammarus fossarum	NC_034937.1	ATA	ATC	ATA	TTG	ATG	ATA	TTG	TTG	ATG	ATA	ATG	TTG	ATG
Pallaseopsis kessleri	NC_033362.1	ATG	ATT	ATT	GTG	ATG	ATG	TTG	TTG	ATG	TTG	ATG	TTG	ATT
Gammarus duebeni	NC_017760.1	ATG	ATA	ATA	TTG	ATG	ATG	ATA	TTG	ATG	ATA	ATG	GTG	ATG
Eulimnogammarus vittatus	NC_025564.1	ATG	GTG	ATT	TTG	ATG	ATG	TTG	TTG	ATG	ATG	ATG	TTG	ATG
Caprella mutica	NC_014492.1	ATG	ATA	ATT	ATA	ATA	ATG	ATA	ATT	ATG	ATG	TTG	ATA	ATT
Eulimnogammarus verrucosus	NC_023104.1	ATG	GTG	ATT	TTG	ATG	ATG	ATT	TTG	ATG	ATA	ATA	ATA	ATG
Pseudoniphargus daviui	NC_019662.2	ATG	ATA	ATT	TTG	ATG	ATG	ATA	ATC	TTG	ATA	ATG	TTG	AAT
Caprella scaura	NC_014687.1	ATG	ATA	ATT	ATA	ATG	ATG	CTG	ATG	ATG	ATG	CTG	ATA	ATA
Metacrangonyx spinicaudatus	NC_019657.1	ATG	ATA	ATT	ATC	ATG	ATG	ATA	ATT	ATT	ATG	ATG	TTG	ATT
Stygobromus tenuis potomacus	KU869712.1	ATG	ATT	ATT	ATC	ATG	ATG	TTG	ATA	ATA	ATG	ATG	ATA	ATG
Metacrangonyx remyi	NC_019660.1	ATG	GTG	ATC	ATT	ATG	ATG	ATA	ATA	ATT	ATG	ATG	CTG	ATT
Platorchestia parapacifica	MG010371.1	ATA	ATG	ATT	ATT	ATG	ATA	ATT	ATC	ATG	ATG	ATT	ATA	ATC

Platorchestia japonica	MG010370.1	ATA	ATG	ATT	ATA	ATG	ATA	ATT	ATA	ATG	ATG	ATA	ATA	ATT
Metacrangonyx ilvanus	NC_019656.1	ATG	ATC	ATT	ATT	ATG	ATG	ATA	ATT	ATT	ATG	ATG	TTG	ATT
Onisimus nanseni	NC_013819.1	ATG	ATT	ATG	TTG	ATG	ATG	ATA	ATG	ATG	TTG	ATG	GTG	ATT
Metacrangonyx longicaudus	NC_019658.1	ATG	ATT	ATT	ATC	ATG	ATG	ATA	ATT	ATT	ATG	ATG	TTG	ATT
Bahadzia jaraguensis	FR872382.1	ATG	ATC	ATT	ATT	ATG	ATG	ATG	ATT	ATT	ATG	ATG	ATA	ATT
Stygobromus indentatus	NC_030261.1	ATA	ATT	ATT	ATA	ATG	ATG	GTG	ATT	ATA	ATT	ATG	GTG	ATG
Metacrangonyx dominicanus	NC_019654.1	ATG	ATC	ATT	ATT	ATG	ATG	ATA	ATT	ATC	ATG	ATG	TTG	ATT
Metacrangonyx goulmimensis	NC_019655.1	ATG	ATC	ATT	ATC	ATG	ATG	ATA	ATT	ATC	ATG	ATG	TTG	ATT
Metacrangonyx panousei	NC_019659.1	ATG	ATC	ATT	ATT	ATG	ATG	ATA	ATT	ATT	ATG	ATG	TTG	ATT
Eulimnogammarus cyaneus	NC_033360.1	ATG	GTG	ATC	TTG	ATG	ATG	TTG	TTG	ATG	ATG	ATG	GTG	ATG
Metacrangonyx repens	NC_019653.1	ATG	GTG	ATT	ATT	ATG	ATG	ATA	ATT	ATT	ATG	ATG	TTG	ATT
Metacrangonyx longipes	NC_013032.1	ATG	ATC	ATT	ATT	ATG	ATG	ATA	ATT	ATT	ATA	ATA	ATT	ATT
Metacrangonyx sp. 3 ssp. 1 MDMBR-2012	HE860504.1	ATG	ATC	ATT	ATT	ATG	ATG	ATA	ATT	ATT	ATG	ATG	TTG	ATT
Metacrangonyx sp. 4 MDMBR-2012	HE860498.1	ATG	ATC	ATC	ATC	ATG	ATG	ATA	ATA	ATC	GTG	ATG	TTG	ATT
Metacrangonyx sp. 1 MDMBR-2012	HE860513.1	ATG	ATC	ATT	ATC	ATG	ATG	ATA	TTT	ATT	ATG	ATG	TTG	ATT

Table C.3 Summary of codon usage in PCG of mitochondrial genomes of all crangonyctid amphipods. Sign "#" indicates total number of certain codon in protein-coding sequences of every species; "%" indicates percent of certain codon in total coding sequence in every species; RSCU indicates the calculated RSCU value of certain codon in total coding sequence in every species.

AA	Codon	Stygobromus indentatus			Crangonyx forbesi			Stygobromus allegheniensis			Stygobromus tenuis potomacus			Stygobromus pizzinii			Bactrurus brachycaudus			Stygobromus tenuis potomacus		
		RSCU	#	%	RSCU	#	%	RSCU	#	%	RSCU	#	%	RSCU	#	%	RSCU	#	%	RSCU	#	%
Phe	UUU	1.74	238	6.45	1.76	255	6.78	1.64	237	6.49	1.70	240	6.51	1.69	241	6.54	1.55	218	5.94	1.72	243	6.57
	UUC	0.26	35	0.95	0.24	35	0.93	0.36	52	1.42	0.30	42	1.14	0.31	45	1.22	0.45	63	1.72	0.28	39	1.05
Leu2	UUA	2.89	286	7.75	2.58	267	7.10	2.56	259	7.09	2.95	298	8.08	2.96	297	8.06	2.10	207	5.64	3.04	314	8.49
	UUG	0.49	48	1.30	0.64	66	1.76	0.89	90	2.46	0.64	65	1.76	0.58	58	1.57	0.75	74	2.02	0.54	56	1.51

Leu1	CUU	1.00	99	2.68	1.05	109	2.90	0.97	98	2.68	1.02	103	2.79	1.03	103	2.79	1.18	116	3.16	0.99	102	2.76
	CUC	0.33	33	0.89	0.32	33	0.88	0.27	27	0.74	0.26	26	0.70	0.24	24	0.65	0.46	45	1.23	0.27	28	0.76
	CUA	1.10	109	2.95	1.07	111	2.95	0.99	100	2.74	0.88	89	2.41	0.97	97	2.63	1.12	110	3.00	0.91	94	2.54
	CUG	0.18	18	0.49	0.33	34	0.90	0.33	33	0.90	0.25	25	0.68	0.23	23	0.62	0.40	39	1.06	0.24	25	0.68
Ile	AUU	1.57	253	6.85	1.47	185	4.92	1.45	188	5.15	1.62	240	6.51	1.61	245	6.64	1.48	201	5.48	1.53	230	6.22
	AUC	0.43	70	1.90	0.53	66	1.76	0.55	71	1.94	0.38	56	1.52	0.39	59	1.60	0.52	70	1.91	0.47	71	1.92
Met	AUA	1.71	197	5.34	1.61	183	4.87	1.66	151	4.13	1.64	169	4.58	1.66	169	4.58	1.37	153	4.17	1.70	168	4.54
	AUG	0.29	33	0.89	0.39	45	1.20	0.34	31	0.85	0.36	37	1.00	0.34	35	0.95	0.63	71	1.94	0.30	30	0.81
Val	GUU	1.71	106	2.87	1.61	112	2.98	1.55	111	3.04	1.72	118	3.20	1.57	101	2.74	1.36	96	2.62	1.60	103	2.79
	GUC	0.32	20	0.54	0.46	32	0.85	0.50	36	0.99	0.42	29	0.79	0.50	32	0.87	0.54	38	1.04	0.34	22	0.60
	GUA	1.24	77	2.09	1.38	96	2.55	1.17	84	2.30	1.39	95	2.58	1.40	90	2.44	1.52	107	2.92	1.61	104	2.81
	GUG	0.73	45	1.22	0.56	39	1.04	0.78	56	1.53	0.47	32	0.87	0.54	35	0.95	0.58	41	1.12	0.45	29	0.78
Ser2	UCU	2.01	108	2.93	1.93	110	2.93	1.81	92	2.52	1.80	94	2.55	1.89	100	2.71	1.74	95	2.59	1.95	100	2.70
	UCC	0.56	30	0.81	0.47	27	0.72	0.47	24	0.66	0.63	33	0.89	0.43	23	0.62	0.75	41	1.12	0.25	13	0.35
	UCA	1.34	72	1.95	1.39	79	2.10	1.38	70	1.92	1.45	76	2.06	1.61	85	2.31	1.37	75	2.04	1.74	89	2.41
	UCG	0.17	9	0.24	0.39	22	0.59	0.18	9	0.25	0.19	10	0.27	0.09	5	0.14	0.31	17	0.46	0.14	7	0.19
Pro	CCU	1.88	69	1.87	1.87	71	1.89	1.95	71	1.94	1.66	62	1.68	1.72	64	1.74	1.76	65	1.77	1.79	67	1.81
	CCC	0.54	20	0.54	0.84	32	0.85	0.44	16	0.44	0.27	10	0.27	0.56	21	0.57	0.65	24	0.65	0.61	23	0.62
	CCA	1.31	48	1.30	0.84	32	0.85	1.37	50	1.37	1.80	67	1.82	1.40	52	1.41	1.11	41	1.12	1.44	54	1.46
	CCG	0.27	10	0.27	0.45	17	0.45	0.25	9	0.25	0.27	10	0.27	0.32	12	0.33	0.49	18	0.49	0.16	6	0.16
Thr	ACU	1.59	70	1.90	1.42	74	1.97	1.81	83	2.27	1.70	75	2.03	1.76	79	2.14	1.37	56	1.53	1.67	79	2.14
	ACC	0.59	26	0.70	0.73	38	1.01	0.52	24	0.66	0.59	26	0.70	0.47	21	0.57	0.88	36	0.98	0.68	32	0.87
	ACA	1.73	76	2.06	1.52	79	2.10	1.51	69	1.89	1.61	71	1.93	1.67	75	2.03	1.37	56	1.53	1.54	73	1.97
	ACG	0.09	4	0.11	0.33	17	0.45	0.15	7	0.19	0.09	4	0.11	0.11	5	0.14	0.37	15	0.41	0.11	5	0.14
Ala	GCU	2.05	106	2.87	1.66	72	1.92	1.86	103	2.82	2.08	108	2.93	1.92	101	2.74	1.77	105	2.86	2.02	109	2.95
	GCC	0.77	40	1.08	0.90	39	1.04	0.89	49	1.34	0.67	35	0.95	0.86	45	1.22	1.08	64	1.74	0.72	39	1.05
	GCA	1.10	57	1.54	1.13	49	1.30	0.98	54	1.48	1.13	59	1.60	0.95	50	1.36	0.81	48	1.31	1.11	60	1.62
	GCG	0.08	4	0.11	0.30	13	0.35	0.27	15	0.41	0.12	6	0.16	0.27	14	0.38	0.34	20	0.55	0.15	8	0.22
Tyr	UAU	1.46	100	2.71	1.11	83	2.21	1.31	88	2.41	1.49	102	2.77	1.32	89	2.41	1.35	75	2.04	1.42	97	2.62

	UAC	0.54	37	1.00	0.89	66	1.76	0.69	46	1.26	0.51	35	0.95	0.68	46	1.25	0.65	36	0.98	0.58	40	1.08
His	CAU	1.37	54	1.46	0.99	36	0.96	1.21	47	1.29	1.30	46	1.25	1.32	49	1.33	0.94	33	0.90	1.28	46	1.24
	CAC	0.63	25	0.68	1.01	37	0.98	0.79	31	0.85	0.70	25	0.68	0.68	25	0.68	1.06	37	1.01	0.72	26	0.70
Gln	CAA	1.70	51	1.38	1.19	34	0.90	1.20	42	1.15	1.42	46	1.25	1.44	46	1.25	1.63	53	1.44	1.31	42	1.14
	CAG	0.30	9	0.24	0.81	23	0.61	0.80	28	0.77	0.58	19	0.52	0.56	18	0.49	0.37	12	0.33	0.69	22	0.60
Asn	AAU	1.31	81	2.19	1.19	80	2.13	1.25	74	2.03	1.40	96	2.60	1.37	90	2.44	1.15	61	1.66	1.39	91	2.46
	AAC	0.69	43	1.16	0.81	55	1.46	0.75	44	1.20	0.60	41	1.11	0.63	41	1.11	0.85	45	1.23	0.61	40	1.08
Lys	AAA	1.63	79	2.14	1.60	100	2.66	1.60	71	1.94	1.47	69	1.87	1.47	69	1.87	1.25	57	1.55	1.60	76	2.06
	AAG	0.37	18	0.49	0.40	25	0.67	0.40	18	0.49	0.53	25	0.68	0.53	25	0.68	0.75	34	0.93	0.40	19	0.51
Asp	GAU	1.24	44	1.19	1.17	42	1.12	1.13	44	1.20	1.41	52	1.41	1.30	48	1.30	1.32	45	1.23	1.39	52	1.41
	GAC	0.76	27	0.73	0.83	30	0.80	0.87	34	0.93	0.59	22	0.60	0.70	26	0.71	0.68	23	0.63	0.61	23	0.62
Glu	GAA	1.35	58	1.57	1.01	37	0.98	1.31	53	1.45	1.24	51	1.38	1.33	55	1.49	1.09	44	1.20	1.40	58	1.57
	GAG	0.65	28	0.76	0.99	36	0.96	0.69	28	0.77	0.76	31	0.84	0.67	28	0.76	0.91	37	1.01	0.60	25	0.68
Cys	UGU	1.62	30	0.81	1.59	35	0.93	1.25	25	0.68	1.49	26	0.70	1.60	28	0.76	1.33	24	0.65	1.56	28	0.76
	UGC	0.38	7	0.19	0.41	9	0.24	0.75	15	0.41	0.51	9	0.24	0.40	7	0.19	0.67	12	0.33	0.44	8	0.22
Trp	UGA	1.53	74	2.00	1.29	65	1.73	1.28	59	1.62	1.30	60	1.63	1.40	65	1.76	1.10	55	1.50	1.61	75	2.03
	UGG	0.47	23	0.62	0.71	36	0.96	0.72	33	0.90	0.70	32	0.87	0.60	28	0.76	0.90	45	1.23	0.39	18	0.49
Arg	CGU	1.36	17	0.46	1.11	15	0.40	1.46	19	0.52	1.54	20	0.54	1.23	16	0.43	0.64	9	0.25	1.23	16	0.43
	CGC	0.32	4	0.11	0.44	6	0.16	0.54	7	0.19	0.38	5	0.14	0.69	9	0.24	1.00	14	0.38	0.54	7	0.19
	CGA	1.76	22	0.60	1.19	16	0.43	1.08	14	0.38	1.38	18	0.49	1.31	17	0.46	1.29	18	0.49	1.85	24	0.65
	CGG	0.56	7	0.19	1.26	17	0.45	0.92	12	0.33	0.69	9	0.24	0.77	10	0.27	1.07	15	0.41	0.38	5	0.14
Ser1	AGU	1.10	59	1.60	1.18	67	1.78	1.08	55	1.51	1.09	57	1.55	1.10	58	1.57	0.75	41	1.12	0.96	49	1.33
	AGC	0.21	11	0.30	0.39	22	0.59	0.41	21	0.58	0.23	12	0.33	0.38	20	0.54	0.46	25	0.68	0.21	11	0.30
	AGA	1.86	100	2.71	1.47	84	2.23	1.97	100	2.74	2.03	106	2.87	1.87	99	2.69	1.59	87	2.37	1.85	95	2.57
	AGG	0.75	40	1.08	0.79	45	1.20	0.69	35	0.96	0.57	30	0.81	0.62	33	0.90	1.03	56	1.53	0.90	46	1.24
Gly	GGU	1.02	58	1.57	1.11	61	1.62	1.12	67	1.83	1.06	62	1.68	1.08	64	1.74	0.86	54	1.47	1.07	63	1.70
	GGC	0.53	30	0.81	0.82	45	1.20	0.43	26	0.71	0.43	25	0.68	0.54	32	0.87	0.83	52	1.42	0.39	23	0.62
	GGA	1.44	82	2.22	1.11	61	1.62	1.17	70	1.92	1.57	92	2.49	1.29	76	2.06	0.91	57	1.55	1.47	87	2.35
	GGG	1.00	57	1.54	0.95	52	1.38	1.28	77	2.11	0.94	55	1.49	1.08	64	1.74	1.39	87	2.37	1.07	63	1.70

Table C.4 Summary of amino acid composition in PCGs of mitochondrial genomes of all crangonyctid amphipods. Sign "#" indicates total number of certain amino acid in protein-coding sequences of every species; "%" indicates percent of certain amino acid in total coding sequence in every species.

AA	Stygobromus tenuis potomacus		Stygobromus indentatus		Stygobromus allegheniensis		Bactrurus brachycaudus		Crangonyx forbesi		Stygobromus pizzinii		Stygobromus tenuis potomacus	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%
Phe	282	7.63	273	7.40	289	7.91	281	7.66	290	7.71	286	7.76	282	7.65
Leu2	370	10.01	334	9.05	349	9.56	281	7.66	333	8.86	355	9.63	363	9.84
Leu1	249	6.74	259	7.02	258	7.06	310	8.45	287	7.64	247	6.70	243	6.59
Ile	301	8.14	323	8.75	259	7.09	271	7.39	251	6.68	304	8.25	296	8.03
Met	198	5.36	230	6.23	182	4.98	224	6.11	228	6.07	204	5.53	206	5.59
Val	258	6.98	248	6.72	287	7.86	282	7.69	279	7.42	258	7.00	274	7.43
Ser2	209	5.65	219	5.93	195	5.34	228	6.22	238	6.33	213	5.78	213	5.78
Pro	150	4.06	147	3.98	146	4.00	148	4.03	152	4.04	149	4.04	149	4.04
Thr	189	5.11	176	4.77	183	5.01	163	4.44	208	5.53	180	4.88	176	4.77
Ala	216	5.84	207	5.61	221	6.05	237	6.46	173	4.60	210	5.70	208	5.64
Tyr	137	3.71	137	3.71	134	3.67	111	3.03	149	3.96	135	3.66	137	3.71
His	72	1.95	79	2.14	78	2.14	70	1.91	73	1.94	74	2.01	71	1.93
Gln	64	1.73	60	1.63	70	1.92	65	1.77	57	1.52	64	1.74	65	1.76
Asn	131	3.54	124	3.36	118	3.23	106	2.89	135	3.59	131	3.55	137	3.71
Lys	95	2.57	97	2.63	89	2.44	91	2.48	125	3.33	94	2.55	94	2.55
Asp	75	2.03	71	1.92	78	2.14	68	1.85	72	1.92	74	2.01	74	2.01
Glu	83	2.25	86	2.33	81	2.22	81	2.21	73	1.94	83	2.25	82	2.22
Cys	36	0.97	37	1.00	40	1.10	36	0.98	44	1.17	35	0.95	35	0.95
Trp	93	2.52	97	2.63	92	2.52	100	2.73	101	2.69	93	2.52	92	2.49
Arg	52	1.41	50	1.35	52	1.42	56	1.53	54	1.44	52	1.41	52	1.41
Ser1	201	5.44	210	5.69	211	5.78	209	5.70	218	5.80	210	5.70	205	5.56
Gly	236	6.38	227	6.15	240	6.57	250	6.82	219	5.83	236	6.40	234	6.34

Table C.5 Nucleotide composition statistics for ribosomal RNA genes in mitochondrial genomes of all crangonyctid amphipods.

Species	Gene	Size (bp)	AT %	AT-skew	GC-skew
	rrnL-	1034	67.8	-0.030	0.381
Bactrurus brachycaudus	*rrnS*-	671	71.5	-0.029	0.361
	rrnL-	1090	72.7	-0.064	0.259
Crangonyx forbesi	*rrnS*-	695	73.6	-0.084	0.359
	rrnL-	1034	70.4	-0.115	0.412
Stygobromus allegheniensis	*rrnS*-	688	74.0	-0.108	0.352
	rrnL-	1035	72.7	-0.089	0.383
Stygobromus indentatus	*rrnS*-	669	77.2	-0.070	0.307
	rrnL-	1037	72.3	-0.093	0.415
Stygobromus pizzinii	*rrnS*-	678	74.6	-0.075	0.337
	rrnL-	1036	72.8	-0.074	0.426
Stygobromus tenuis potomacus	*rrnS*-	679	76.0	-0.109	0.370
	rrnL-	1034	72.3	-0.099	0.420
Stygobromus tenuis potomacus	*rrnS*-	683	74.6	-0.088	0.322

Table C.6 Results of selective pressure (ω ratio) analyses of mitochondrial PCGs with LRT P-value < 0.05 in subterranean and surface lineages of amphipods based on 2 vs. 1 ratio model.

Model	np	Ln L	Estimates of parameters			Model compared	LRT P-value	Omega for Foreground Branch	Gene	Species
Two ratio Model 2	70	-24608.883248	ω:	ω0=0.02037	ω1=0.05379	Model 0 vs. Two ratio Model 2	0.000131407	ω1=0.05379	*cox1*	*C. forbesi*
Model 0	69	-24616.193996	ω=	0.02099						
Two ratio Model 2	70	-15283.383385	ω:	ω0=0.05091	ω1=0.18347	Model 0 vs. Two ratio Model 2	0.037178941	ω1=0.18347	*atp6*	*P. daviui*
Model 0	69	-15285.554483	ω=	0.05174						
Two ratio Model 2	70	-24611.896377	ω:	ω0=0.02052	ω1=0.04222	Model 0 vs. Two ratio Model 2	0.003370432	ω1=0.04222	*cox1*	
Model 0	69	-24616.193996	ω=	0.02099						
Two ratio Model 2	70	-19414.614776	ω:	ω0=0.03225	ω1=0.19046	Model 0 vs. Two ratio Model 2	0.002904709	ω1=0.19046	*nad1*	
Model 0	69	-19419.047965	ω=	0.03239						

Two ratio Model 2	70	-24599.610909	ω:	ω0=0.02012	ω1=0.07488	Model 0 vs. Two ratio Model 2	0.000000008	ω1=0.07488	*cox1*	*G. fasciatus*
Model 0	69	-24616.193996	ω=	0.02099						
Two ratio Model 2	70	-12953.332620	ω:	ω0=0.03009	ω1=0.21013	Model 0 vs. Two ratio Model 2	0.000072299	ω1=0.21013	*cox2*	
Model 0	69	-12961.207473	ω=	0.03033						
Two ratio Model 2	70	-16204.441186	ω:	ω0=0.04169	ω1=0.09747	Model 0 vs. Two ratio Model 2	0.012180256	ω1=0.09747	*cox3*	
Model 0	69	-16207.583406	ω=	0.04274						
Two ratio Model 2	70	-29486.407012	ω:	ω0=0.04262	ω1=0.12215	Model 0 vs. Two ratio Model 2	0.002680018	ω1=0.12215	*nad4*	
Model 0	69	-29490.913731	ω=	0.04296						
Two ratio Model 2	70	-16205.223658	ω:	ω0=0.04182	ω1=0.10051	Model 0 vs. Two ratio Model 2	0.029822483	ω1=0.10051	*cox3*	*O. nanseni*
Model 0	69	-16207.583406	ω=	0.04274						
Two ratio Model 2	70	-24619.041265	ω:	ω0=0.01766	ω1=0.04082	Model 0 vs. Two ratio Model 2	0.017017792	ω1=0.04082	*cox1*	*G. fossarum*
Model 0	69	-24616.193996	ω=	0.02099						
Two ratio Model 2	70	-23400.542887	ω:	ω0=0.03254	ω1=0.08246	Model 0 vs. Two ratio Model 2	0.004931290	ω1=0.08246	*cytb*	
Model 0	69	-23404.495120	ω=	0.03337						
Two ratio Model 2	70	-19414.218221	ω:	ω0=0.03259	ω1=0.13042	Model 0 vs. Two ratio Model 2	0.001883761	ω1=0.13042	*nad1*	
Model 0	69	-19419.047965	ω=	0.03239						
Two ratio Model 2	70	-29487.580618	ω:	ω0=0.04235	ω1=0.09448	Model 0 vs. Two ratio Model 2	0.009825704	ω1=0.09448	*nad4*	
Model 0	69	-29490.913731	ω=	0.04296						
Two ratio Model 2	70	-6864.685694	ω:	ω0=0.03733	ω1=0.19416	Model 0 vs. Two ratio Model 2	0.030276845	ω1=0.19416	*nad4l*	
Model 0	69	-6867.032446	ω=	0.03788						
Two ratio Model 2	70	-24613.676186	ω:	ω0=0.02143	ω1=0.01086	Model 0 vs. Two ratio Model 2	0.024831197	ω1=0.01086	*cox1*	*B. jaraguensis*
Model 0	69	-24616.193996	ω=	0.02099						
Two ratio Model 2	70	-42645.830655	ω:	ω0=0.06114	ω1=0.02138	Model 0 vs. Two ratio Model 2	0.035443809	ω1=0.02138	*nad5*	
Model 0	69	-42648.042488	ω=	0.06077						
Two ratio Model 2	70	-24610.442091	ω:	ω0=0.02042	ω1=0.04738	Model 0 vs. Two ratio Model 2	0.000694537	ω1=0.04738	*cox1*	*B. brachycaudus*
Model 0	69	-24616.193996	ω=	0.02099						
Two ratio Model 2	70	-16205.427054	ω:	ω0=0.04211	ω1=0.09064	Model 0 vs. Two ratio Model 2	0.037828788	ω1=0.09064	*cox3*	
Model 0	69	-16207.583406	ω=	0.04274						
Two ratio Model 2	70	-19416.309168	ω:	ω0=0.03286	ω1=0.00207	Model 0 vs. Two ratio Model 2	0.019261754	ω1=0.00207	*nad1*	

Model 0	69	-19419.047965	ω=	0.03239						
Two ratio Model 2	70	-24607.788487	ω:	ω0=0.02132	ω1=0.00010	Model 0 vs. Two ratio Model 2	0.000041293	ω1=0.00010	*cox1*	*S. tenuis_MN*
Model 0	69	-24616.193996	ω=	0.02099						
Two ratio Model 2	70	-12967.565600	ω:	ω0=0.02987	ω1=2.22100	Model 0 vs. Two ratio Model 2	0.000362491	ω1=2.22100	*cox2*	
Model 0	69	-12961.207473	ω=	0.03033						
Two ratio Model 2	70	-19417.045819	ω:	ω0=0.03302	ω1=0.00531	Model 0 vs. Two ratio Model 2	0.045384555	ω1=0.00531	*nad1*	
Model 0	69	-19419.047965	ω=	0.03239						
Two ratio Model 2	70	-6864.477274	ω:	ω0=0.03988	ω1=0.00010	Model 0 vs. Two ratio Model 2	0.023783604	ω1=0.00010	*nad4l*	
Model 0	69	-6867.032446	ω=	0.03788						
Two ratio Model 2	70	-42645.097762	ω:	ω0=0.06190	ω1=0.02375	Model 0 vs. Two ratio Model 2	0.015231840	ω1=0.02375	*nad5*	
Model 0	69	-42648.042488	ω=	0.06077						
Two ratio Model 2	70	-24643.013443	ω:	ω0=0.02059	ω1=37.72058	Model 0 vs. Two ratio Model 2	0.000000000	ω1=37.72058	*cox1*	*S. tenuis_KU*
Model 0	69	-24616.193996	ω=	0.02099						
Two ratio Model 2	70	-16204.089657	ω:	ω0=0.04368	ω1=0.01396	Model 0 vs. Two ratio Model 2	0.008208101	ω1=0.01396	*cox3*	
Model 0	69	-16207.583406	ω=	0.04274						
Two ratio Model 2	70	-42643.975206	ω:	ω0=0.06250	ω1=0.02993	Model 0 vs. Two ratio Model 2	0.004342929	ω1=0.02993	*nad5*	
Model 0	69	-42648.042488	ω=	0.06077						
Two ratio Model 2	70	-11161.910962	ω:	ω0=0.04608	ω1=0.01015	Model 0 vs. Two ratio Model 2	0.045732562	ω1=0.01015	*nad6*	
Model 0	69	-11163.906671	ω=	0.04373						
Two ratio Model 2	70	-24609.090352	ω:	ω0=0.02152	ω1=0.00488	Model 0 vs. Two ratio Model 2	0.000163735	ω1=0.00488	*cox1*	*S. allegheniensis*
Model 0	69	-24616.193996	ω=	0.02099						
Two ratio Model 2	70	-16205.109405	ω:	ω0=0.04350	ω1=0.01645	Model 0 vs. Two ratio Model 2	0.026120839	ω1=0.01645	*cox3*	
Model 0	69	-16207.583406	ω=	0.04274						
Two ratio Model 2	70	-24613.689587	ω:	ω0=0.02117	ω1=0.00359	Model 0 vs. Two ratio Model 2	0.025218520	ω1=0.00359	*cox1*	*S. pizzinii*
Model 0	69	-24616.193996	ω=	0.02099						
Two ratio Model 2	70	-12958.490043	ω:	ω0=0.03112	ω1=0.00477	Model 0 vs. Two ratio Model 2	0.019738673	ω1=0.00477	*cox2*	
Model 0	69	-12961.207473	ω=	0.03033						
Two ratio Model 2	70	-23400.497996	ω:	ω0=0.03415	ω1=0.00760	Model 0 vs. Two ratio Model 2	0.004692619	ω1=0.00760	*cytb*	
Model 0	69	-23404.495120	ω=	0.03337						

Two ratio Model 2	70	-29487.040979	ω:	ω0=0.04390	ω1=0.01007	Model 0 vs. Two ratio Model 2	0.005384643	ω1=0.01007	*nad4*	
Model 0	69	-29490.913731	ω=	0.04296						
Two ratio Model 2	70	-24606.402409	ω:	ω0=0.02171	ω1=0.00333	Model 0 vs. Two ratio Model 2	0.000009631	ω1=0.00333	*cox1*	*S. indentatus*
Model 0	69	-24616.193996	ω=	0.02099						
Two ratio Model 2	70	-19831.073084	ω:	ω0=0.04595	ω1=0.00102	Model 0 vs. Two ratio Model 2	0.000440467	ω1=0.00102	*nad2*	
Model 0	69	-19837.249185	ω=	0.04445						

Table C.7 Evidence of positive selection on the mitochondrial PCGs with LRT P-value < 0.05 and positively selected site (BEB: P≥95%) in subterranean and surface dwelling lineages of amphipods based on branch-site models.

Species	**Gene**	**Model**	**np**	**Ln L**	**Estimates of parameters**					**Model compared**	**LRT P-value**	**Positive sites (BEB: P≥95%)**
P. kessleri	*atp8*	Model A	72	-3843.85	Site class	0	1	2a	2b	Model A vs.Model A null	0.0007	4 S 1.000**,5 N 0.978*,6 W 1.000**,8 F 1.000**,14 L 0.969*,16 I 0.981*,20 M 0.976*,21 N 0.952*,24 L 0.999**,28 S 0.988*,33 L 0.951*,34 N 0.958*,35 N 0.993**,37 A 0.986*
					f	0.23149	0.32042	0.18795	0.26014			
					ω0	0.12034	1.00000	0.12034	1.00000			
					ω1	0.12034	1.00000	999.00000	999.00000			
		Model A null	71	-3849.58	1							Not Allowed
	nad5	Model A	72	-42215.62	Site class	0	1	2a	2b	Model A vs.Model A null	0.0468	12 G 0.992**,34 W 0.993**,60 T 0.964*,91 M 0.966*,358 L 0.971*,457 K 0.964*,477 S 0.979*,511 S 0.960*
					f	0.76401	0.11614	0.10404	0.01581			
					ω0	0.06917	1.00000	0.06917	1.00000			
					ω1	0.06917	1.00000	2.17571	2.17571			
		Model A null	71	-42217.60	1							Not Allowed

G. antarctica	*nad5*	Model A	72	-42207.04	Site class	0	1	2a	2b	Model A vs.Model A null	0.0045	35 G 0.996**,40 N 0.971*,46 F 0.994**,47 S 0.957*,59 S 0.985*,131 S 0.986*,171 S 0.953*,192 S 0.989*,239 I 0.965*,283 E 0.993**,381 V 0.980*,389 M 0.993**,429 W 0.953*,472 G 0.987*,473 L 0.991**,477 S 0.976*,488 K 0.962*
					f	0.69413	0.11323	0.16562	0.02702			
					ω0	0.07063	1.00000	0.07063	1.00000			
					ω1	0.07063	1.00000	42.86420	42.86420			
		Model A null	71	-42211.07	1							Not Allowed
	nad6	Model A	72	-11115.29	Site class	0	1	2a	2b	Model A vs.Model A null	0.0019	51 G 0.989*,86 N 0.982*,100 S 0.964*,108 L 0.954*,128 G 0.974*
					f	0.60982	0.15741	0.18502	0.04776			
					ω0	0.07302	1.00000	0.07302	1.00000			
					ω1	0.07302	1.00000	1.50615	1.50615			
		Model A null	71	-11110.47	1							Not Allowed
P. japonica	*nad2*	Model A	72	-19775.06	Site class	0	1	2a	2b	Model A vs.Model A null	0.0172	116 D 0.979*, 232 S 0.970*
					f	0.92964	0.04229	0.02685	0.00122			
					ω0	0.04908	1.00000	0.04908	1.00000			
					ω1	0.04908	1.00000	855.94976	855.94976			
		Model A null	71	-19777.90	1							Not Allowed
	nad6	Model A	72	-11114.94	Site class	0	1	2a	2b	Model A vs.Model A null	0.0420	86 N 0.982*
					f	0.80724	0.14323	0.04207	0.00746			
					ω0	0.06706	1.00000	0.06706	1.00000			
					ω1	0.06706	1.00000	998.99536	998.99536			
		Model A null	71	-11117.00	1							Not Allowed

C. mutica	*nad1*	Model A	72	-19367.63	Site class	0	1	2a	2b	Model A vs.Model A null	0.0007	163 I 0.997**,243 S 0.984*,249 V 0.989*
					f	0.94079	0.02466	0.03367	0.00088			
					ω0	0.03489	1.00000	0.03489	1.00000			
					ω1	0.03489	1.00000	14.28437	14.28437			
		Model A null	71	-19373.41	1							Not Allowed
	nad5	Model A	72	-42225.68	Site class	0	1	2a	2b	Model A vs.Model A null	0.0030	280 G 0.975*,326 G 0.975*,400 S 0.989*
					f	0.79382	0.13139	0.06417	0.01062			
					ω0	0.07070	1.00000	0.07070	1.00000			
					ω1	0.07070	1.00000	9.66169	9.66169			
		Model A null	71	-42230.09	1							Not Allowed
	cox2	Model A	72	-12930.83	Site class	0	1	2a	2b	Model A vs.Model A null	0.0115	142 K 0.986*
					f	0.97663	0.00901	0.01423	0.00013			
					ω0	0.03012	1.00000	0.03012	1.00000			
					ω1	0.03012	1.00000	268.52413	268.52413			
		Model A null	71	-12934.02	1							Not Allowed
G. fossarum	*nad3*	Model A	72	-7315.91	Site class	0	1	2a	2b	Model A vs.Model A null	0.0439	6 F 0.991**,17 L 0.999**,22 H 0.979*,23 S 0.999**,25 P 0.969*,26 S 1.000**,76 P 0.978*,82 T 0.972*,84 L 0.998**,91 V 0.974*,93 L 0.966*,94 I 0.999**
					f	0.59041	0.08730	0.28077	0.04152			
					ω0	0.03895	1.00000	0.03895	1.00000			
					ω1	0.03895	1.00000	11.02113	11.02113			
		Model A null	71	-7317.94	1							Not Allowed

O. nanseni	*atp6*	Model A	72	-15167.07	Site class	0	1	2a	2b	Model A vs.Model A null	0.0478	69 M 0.986*, 146 N 0.981*, 172 S 0.974*, 186 A 0.969*, 188 G 0.981*, 189 L 0.996**
					f	0.76666	0.07611	0.14304	0.01420			
					ω0	0.05002	1.00000	0.05002	1.00000			
					ω1	0.05002	1.00000	18.13180	18.13180			
		Model A null	71	-15169.03	1							Not Allowed
	nad5	Model A	72	-42207.40	Site class	0	1	2a	2b	Model A vs.Model A null	0.0004	3 S 0.991**, 6 A 0.988*, 73 S 0.985*, 98 F 0.972*, 130 S 0.968*, 162 W 0.961*, 179 I 0.960*, 242 M 0.955*, 325 W 0.976* , 341 C 0.995**, 367 L 0.990**, 373 V 0.953*, 377 G 0.974*, 427 M 0.987*, 465 N 0.989*, 479 W 0.977*, 514 L 0.969*, 515 Q 0.989*, 518 Q 0.999**, 519 S 0.985*, 523 S 0.959*
					f	0.72687	0.12217	0.12924	0.02172			
					ω0	0.07034	1.00000	0.07034	1.00000			
					ω1	0.07034	1.00000	62.60443	62.60443			
		Model A null	71	-42213.74	1							Not Allowed
G. fasciatus	*atp8*	Model A	72	-4021.06	Site class	0	1	2a	2b	Model A vs.Model A null	0.0000	7 S 0.999**, 12 F 0.976*
					f	0.25262	0.42481	0.12029	0.20228			
					ω0	0.09609	1.00000	0.09609	1.00000			
					ω1	0.09609	1.00000	327.33424	327.33424			
		Model A null	71	-3851.18	1							Not Allowed
	nad2	Model A	72	-19783.73	Site class	0	1	2a	2b	Model A vs.Model A null	0.0000	117 M 0.960*, 191 K 0.967*
					f	0.80648	0.03554	0.15131	0.00667			
					ω0	0.04922	1.00000	0.04922	1.00000			
					ω1	0.04922	1.00000	1.00000	1.00000			
		Model A null	71	-19775.46	1							Not Allowed
	nad5	Model A	72	-42209.09	Site class	0	1	2a	2b	Model A vs.Model A null	0.0001	17 V 0.993**, 27 N 0.977*, 72 S 0.978*, 159 S 0.972*, 161 G 0.994**, 183 V 0.961*, 242 M 0.976*, 243 M 0.978*, 313 S 0.966*, 323 G 0.985*, 368 C 0.960*, 376 S 0.962*, 388 S 0.996**, 389 M 0.852, 391 V 0.984*, 474 A 0.997**, 514 L 0.976*
					f	0.75033	0.12088	0.11092	0.01787			
					ω0	0.06986	1.00000	0.06986	1.00000			
					ω1	0.06986	1.00000	94.94243	94.94243			
		Model A null	71	-42216.31	1							Not Allowed

C. forbesi	*nad5*	Model A	72	-42180.92	Site class	0	1	2a	2b	Model A vs.Model A null	0.0000	2 F 0.982*, 20 S 0.986*, 23 V 0.994**, 40 N 0.978*, 48 L 0.979*, 56 V 0.970*, 78 G 0.991**, 82 I 0.993**, 88 I 0.958*, 121 A 0.997**, 131 S 0.992**, 132 Q 0.999**, 152 S 0.981*, 157 S 0.979*, 159 S 0.989*, 165 V 0.965*, 178 I 0.962*, 181 A 0.990*, 192 S 0.998**, 200 A 0.997**, 250 L 0.989*, 253 N 0.980*, 254 Y 0.991**, 271 G 0.995**, 278 S 0.989*, 283 E 0.989*, 323 G 0.978*, 333 F 0.966*, 336 C 0.991**, 337 N 0.999**, 344 P 0.957*, 347 S 0.996**, 354 L 0.983*, 355 V 0.956*, 359 M 0.725, 360 L 0.956*, 361 S 0.996**, 373 V 0.989*, 375 S 0.988*, 384 V 0.989*, 411 G 0.999**, 412 G 0.956*, 417 F 0.984*, 444 S 0.981*, 448 L 0.984*, 465 N 0.986*, 472 G 0.996**, 473 L 0.994**, 474 A 0.998**, 495 L 0.997**, 502 G 0.970*, 503 G 0.992**, 507 F 0.991**, 511 S 0.994**, 519 S 0.982*, 523 S 0.996**, 538 V 0.977*
					f	0.58855	0.10272	0.26286	0.04588			
					ω0	0.07187	1.00000	0.07187	1.00000			
					ω1	0.07187	1.00000	63.07884	63.07884			
		Model A null	71	-42191.63	1							Not Allowed
	nad6	Model A	72	-11102.90	Site class	0	1	2a	2b	Model A vs.Model A null	0.0077	3 I 1.000**, 7 L 1.000**, 12 L 0.997**, 20 L 0.997**, 23 T 0.984*, 28 V 0.958*, 29 A 0.996**, 35 W 0.996**, 42 A 0.996**, 49 F 0.964*, 50 L 0.970*, 58 T 0.992**, 61 T 0.999**, 69 T 1.000**, 85 N 0.994**, 87 L 0.990*, 95 H 1.000**, 96 K 0.995**, 97 I 0.956*, 100 S 0.996**, 101 G 0.982*, 102 T 0.956*, 103 E 0.995**, 106 T 0.977*, 129 S 1.000**, 133 S 0.986*
					f	0.47135	0.07130	0.39725	0.06009			
					ω0	0.06479	1.00000	0.06479	1.00000			
					ω1	0.06479	1.00000	61.36780	61.36780			
		Model A null	71	-11106.46	1							Not Allowed
	cox1	Model A	72	-24421.90	Site class	0	1	2a	2b	Model A vs.Model A null	0.0385	155 E 0.972*
					f	0.90038	0.02854	0.06889	0.00218			
					ω0	0.01865	1.00000	0.01865	1.00000			
					ω1	0.01865	1.00000	74.69725	74.69725			
		Model A null	71	-24424.04	1							Not Allowed
	cytb	Model A	72	-23302.46	Site class	0	1	2a	2b	Model A vs.Model A null	0.0000	4 S 1.000**, 24 L 0.961*, 50 S 0.999**, 52 T 0.985*, 53 L 0.997**, 81 C 1.000**, 89 G 0.999**, 96 L 0.997**, 97 Q 1.000**, 99 H 0.999**, 111 T 0.997**, 147 D 0.997**, 150 K 0.958*, 179 A 0.999**, 181 A 0.966*, 197 L 0.998**, 237 L 0.998**, 244 T 0.984*, 250 T 0.979*, 253 I 0.999**, 273 N 0.984*, 279 L 0.968*, 282 L 0.993**, 283 L 0.998**, 294 T 0.998**, 298 K 0.999**, 307 N 1.000**, 324 L 0.967*, 335 F 0.997**, 356 T 0.998**
					f	0.00003	0.00000	0.96156	0.03841			
					ω0	0.03404	1.00000	0.03404	1.00000			
					ω1	0.03404	1.00000	1.00000	1.00000			
		Model A null	71	-23269.10	1							Not Allowed

P. daviui	*nad4*	Model A	72	-29345.91	Site class	0	1	2a	2b	Model A vs.Model A null	0.0134	23 S 0.951*, 37 S 0.970*, 48 S 0.974*, 60 A 0.989*, 84 L 0.983*, 260 V 0.971*, 284 S 0.996**, 319 W 0.959*, 379 E 0.973*, 390 G 0.990**, 411 S 0.958*, 413 I 0.956*, 414 F 0.989*
					f	0.78376	0.04107	0.16645	0.00872			
					ω0	0.04528	1.00000	0.04528	1.00000			
					ω1	0.04528	1.00000	13.69295	13.69295			
		Model A null	71	-29348.97	1							Not Allowed
	nad5	Model A	72	-42222.23	Site class	0	1	2a	2b	Model A vs.Model A null	0.0362	6 A 0.970*, 20 S 0.955*, 23 V 0.982*, 132 Q 0.992**, 148 G 0.953*, 157 S 0.953*, 232 S 0.951*, 282 I 0.976*, 285 A 0.970*, 332 I 0.956*, 461 T 0.958*, 486 K 0.990*
					f	0.73648	0.11827	0.12515	0.02010			
					ω0	0.07092	1.00000	0.07092	1.00000			
					ω1	0.07092	1.00000	43.91410	43.91410			
		Model A null	71	-42224.43	1							Not Allowed
B. jaraguensis	*nad2*	Model A	72	-19779.55	Site class	0	1	2a	2b	Model A vs.Model A null	0.0001	162 M 0.959*, 233 F 0.973*
					f	0.79491	0.03611	0.16164	0.00734			
					ω0	0.04900	1.00000	0.04900	1.00000			
					ω1	0.04900	1.00000	1.28390	1.28390			
		Model A null	71	-19771.49	1							Not Allowed
M. dominicanus	*nad3*	Model A	72	-7323.42	Site class	0	1	2a	2b	Model A vs.Model A null	0.0190	81 N 0.995**
					f	0.85321	0.12507	0.01895	0.00278			
					ω0	0.03891	1.00000	0.03891	1.00000			
					ω1	0.03891	1.00000	161.98168	161.98168			
		Model A null	71	-7326.17	1							Not Allowed

B. brachycaudus	*atp6*	Model A	72	-15165.74	Site class	0	1	2a	2b	Model A vs.Model A null	0.0206	10 S 0.992**, 66 S 0.969*, 79 S 0.987*, 119 N 0.971*, 128 I 0.990**, 147 N 0.974*, 177 Y 0.982*, 184 G 0.990*
					f	0.70794	0.06884	0.20344	0.01978			
					ω0	0.05148	1.00000	0.05148	1.00000			
					ω1	0.05148	1.00000	44.53036	44.53036			
		Model A null	71	-15168.42	1							Not Allowed
	nad5	Model A	72	-42197.28	Site class	0	1	2a	2b	Model A vs.Model A null	0.0001	1 S 0.977*, 12 G 0.959*, 13 S 0.991**, 45 S 0.980*, 73 S 0.972*, 74 S 0.953*, 97 G 0.956*, 130 S 0.972*, 132 Q 0.999**, 163 G 0.984*, 229 S 0.987*, 312 I 0.986*, 330 L 0.992**, 336 C 0.986*, 365 S 0.981*, 375 S 0.982*, 388 S 0.983*, 391 V 0.976*, 409 S 0.966*, 411 G 0.990**, 423 L 0.988*, 448 L 0.985*, 469 Y 0.983*, 472 G 0.991**, 479 W 0.991**, 485 Y 0.957*, 493 S 0.989*, 497 E 0.975*, 498 V 0.993**, 500 P 0.991**, 519 S 0.990*
					f	0.65345	0.10879	0.20383	0.03393			
					ω0	0.07113	1.00000	0.07113	1.00000			
					ω1	0.07113	1.00000	54.77953	54.77953			
		Model A null	71	-42204.57	1							Not Allowed
S. allegheniensis	*nad3*	Model A	72	-7325.19	Site class	0	1	2a	2b	Model A vs.Model A null	0.0375	77 V 0.990*
					f	0.85787	0.12736	0.01286	0.00191			
					ω0	0.03868	1.00000	0.03868	1.00000			
					ω1	0.03868	1.00000	108.89135	108.89135			
		Model A null	71	-7327.36	1							Not Allowed

S. indentatus	*atp8*	Model A	72	-4020.49	Site class	0	1	2a	2b	Model A vs.Model A null	0.0028	12 F 0.994**, 14 I 0.981*
					f	0.30650	0.45936	0.09370	0.14043			
					ω0	0.09877	1.00000	0.09877	1.00000			
					ω1	0.09877	1.00000	999.00000	999.00000			
		Model A null	71	-4024.97	1							Not Allowed
	nad5	Model A	72	-42219.00	Site class	0	1	2a	2b	Model A vs.Model A null	0.0001	3 S 0.974*, 59 S 0.991**, 171 S 0.967*, 310 S 0.989*, 321 V 0.970*, 419 L 0.982*, 468 A 0.985*, 469 Y 0.968*, 486 K 0.986*
					f	0.79435	0.13501	0.06038	0.01026			
					ω0	0.07100	1.00000	0.07100	1.00000			
					ω1	0.07100	1.00000	13.75919	13.75919			
		Model A null	71	-42227.16	1							Not Allowed
	cox3	Model A	72	-16038.31	Site class	0	1	2a	2b	Model A vs.Model A null	0.0150	118 T 0.989*
					f	0.92090	0.06266	0.01539	0.00105			
					ω0	0.03997	1.00000	0.03997	1.00000			
					ω1	0.03997	1.00000	29.92053	29.92053			
		Model A null	71	-16041.27	1							Not Allowed

Appendix D. Chapter 3 Figures and Tables

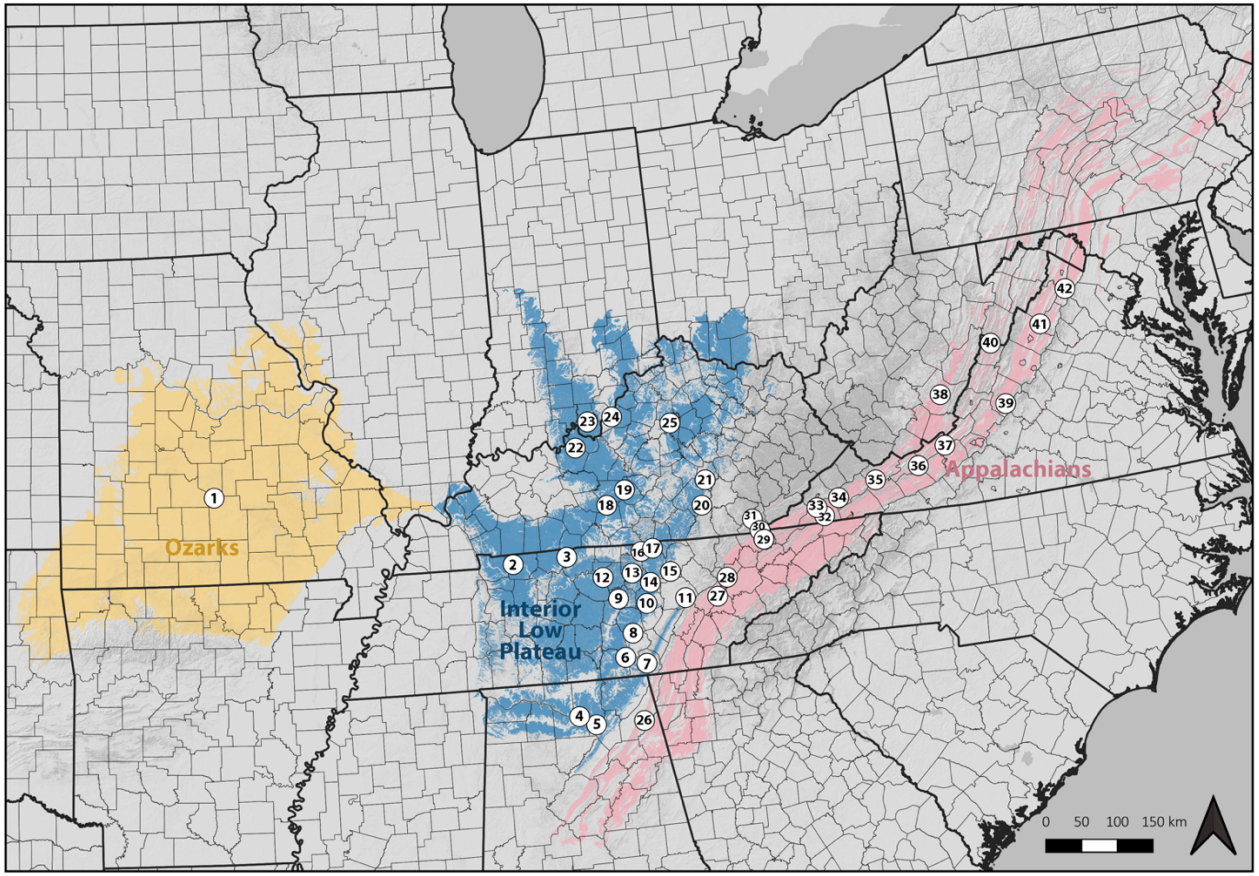


Figure D.1 Map of the localities sampled for the studied species in the cave trechine group are indicated in white circles enclosing specimen number. Geographical area ranges including the Appalachians (APP), the Interior Low Plateau (ILP), and the Ozarks (OZK) are represented in pink, blue, and yellow colors respectively.

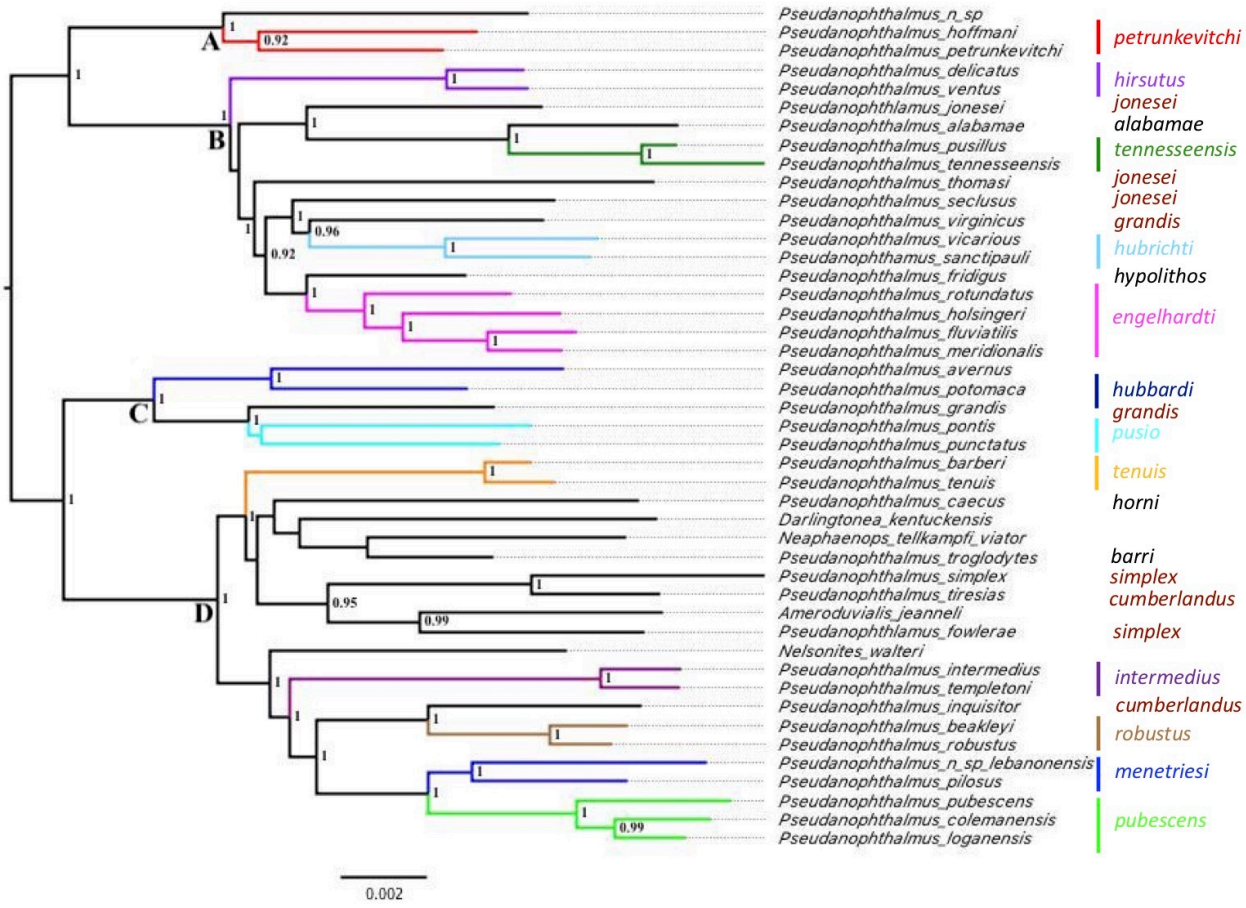


Figure D.2 Bayesian phylogeny of cave trechine beetles from eastern North America inferred from 75% complete concatenated UCE matrix. Numbers indicate support values (Bayesian posterior probability) for nodes greater than 0.90. Outgroup taxa not shown, and four primary clades are labeled A, B, C, and D.

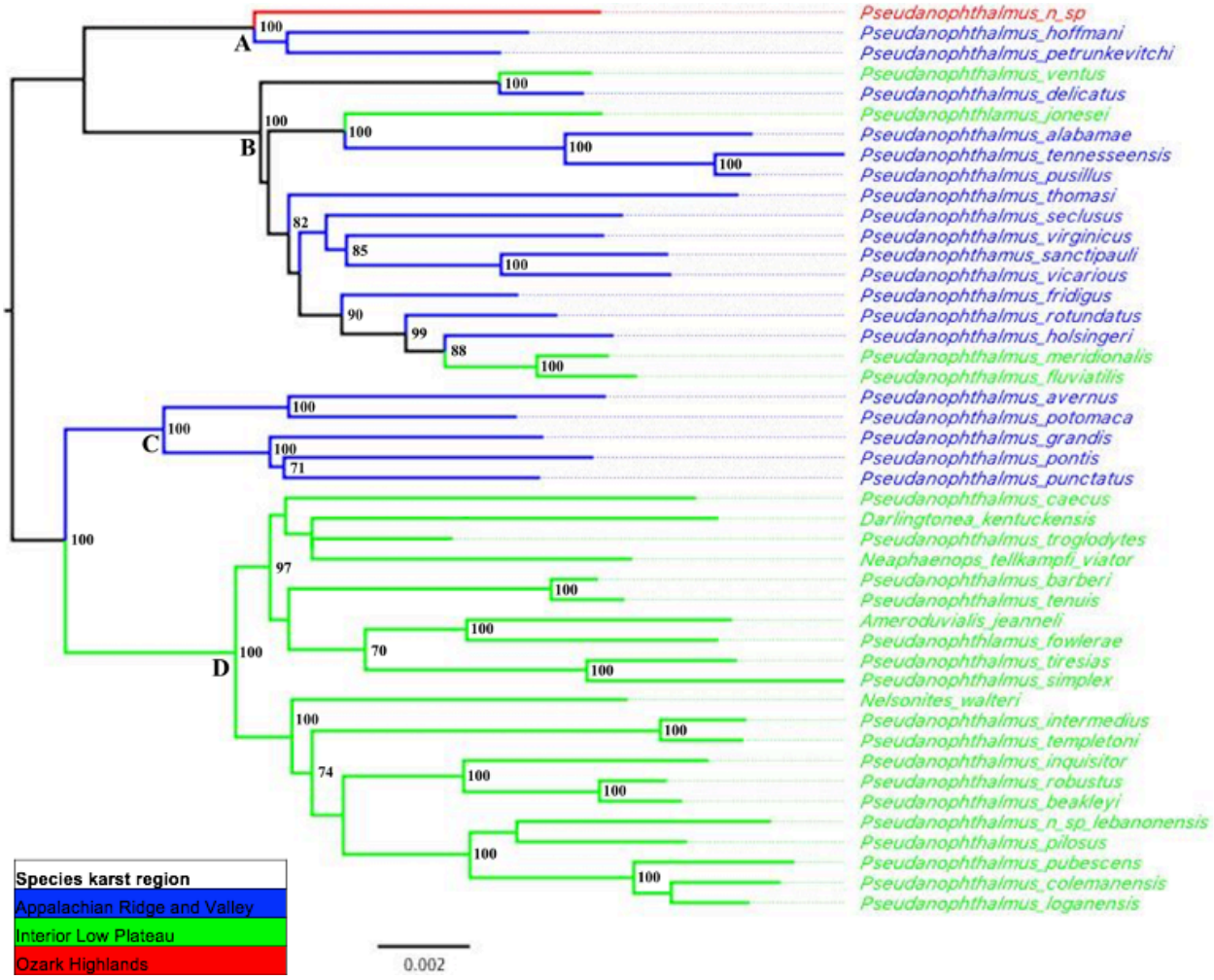


Figure D.3 Maximum-likelihood phylogeny of cave trechine beetles from eastern North America inferred from 75% complete concatenated UCE matrix. Numbers indicate bootstrap support values (maximum-likelihood bootstrap) for nodes greater than 70. Outgroup taxa not shown, and four primary clades are labeled A, B, C, D.

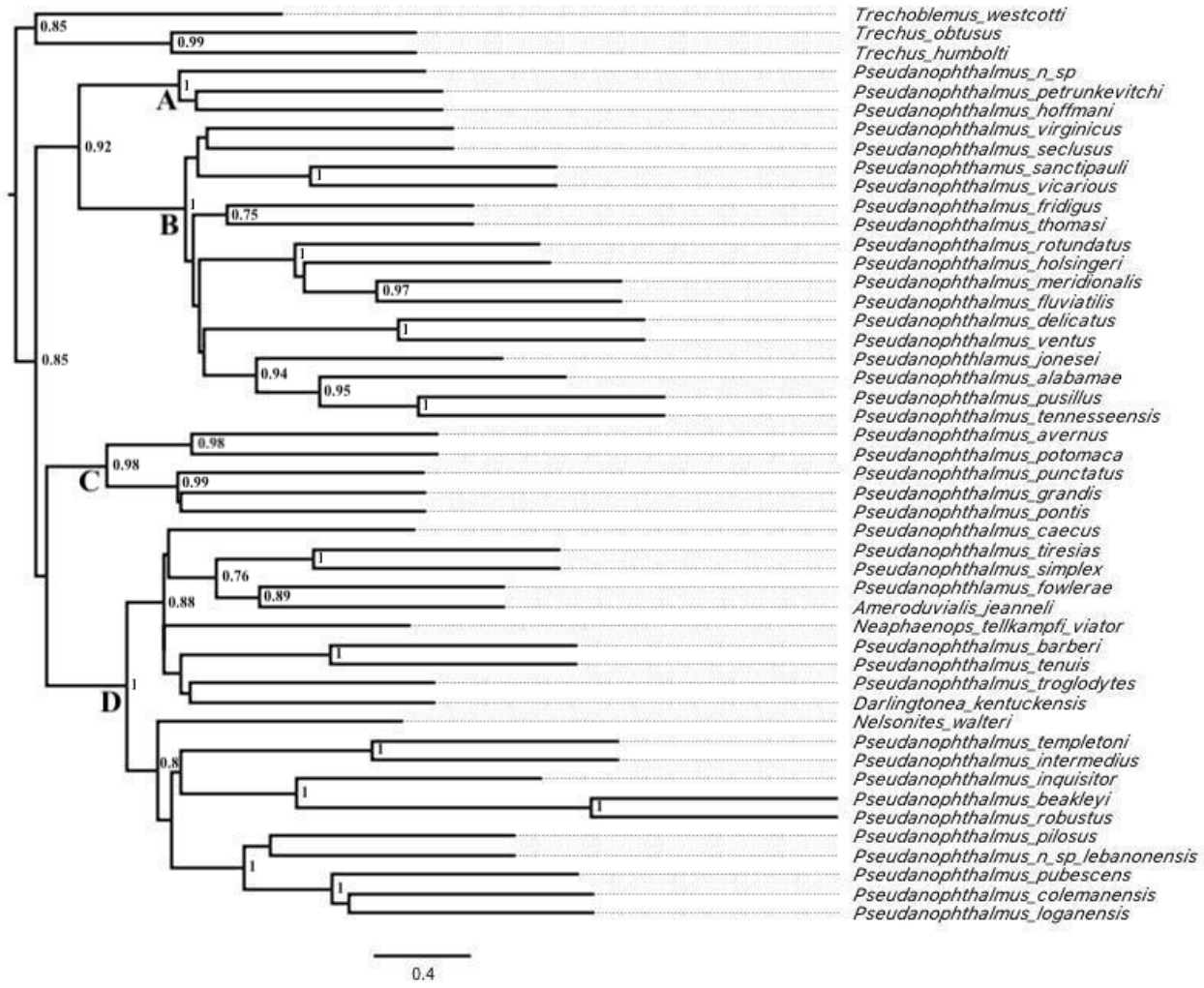


Figure D.4 ASTRAL coalescent species tree, input trees derived from multi-partitioned IQTree analyses of individual gene trees. Lower posterior probability support values greater than 0.90 are displayed.

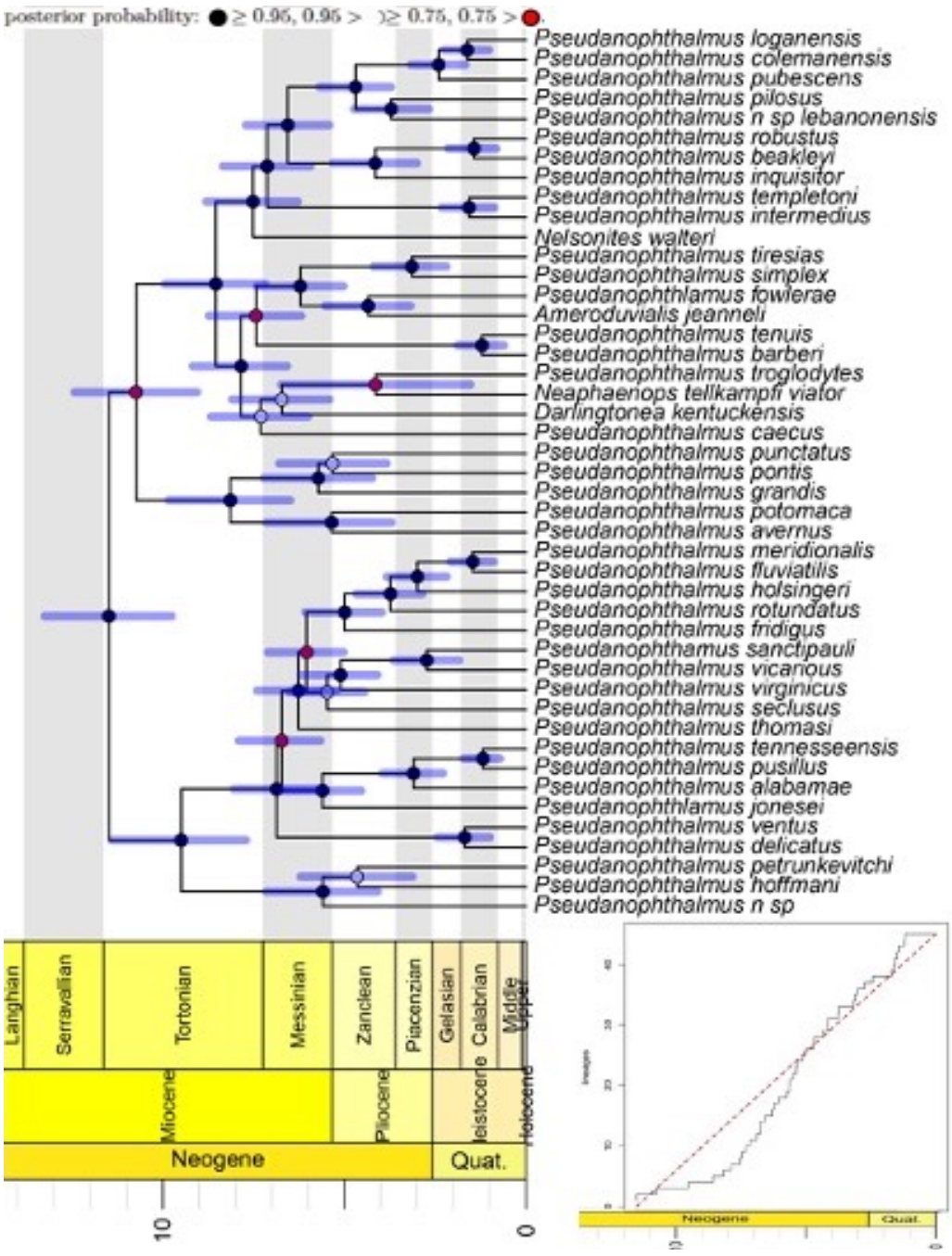


Figure D.5 Time-calibrated maximum clade credibility tree inferred from 75% concatenated UCE matrix, summarized by TreeAnnotator, and plotted with a geological time scale using the strap package in R. Phylogeny dated using a Bayesian relaxed clock method in BEAST. Branches are proportional to time in millions of years. Outgroups were pruned after analyses for an enlarged view. The 95% confidence intervals for the ages of basal branches in the tree are indicated with transparent blue bars. Lineages through time plot displayed at lower right corner. The internal nodes of the tree are indicated with circles, where circles mark nodes with posterior probability: ● ≥ 0.95, 0.95 > ○≥ 0.75, 0.75 > ●.

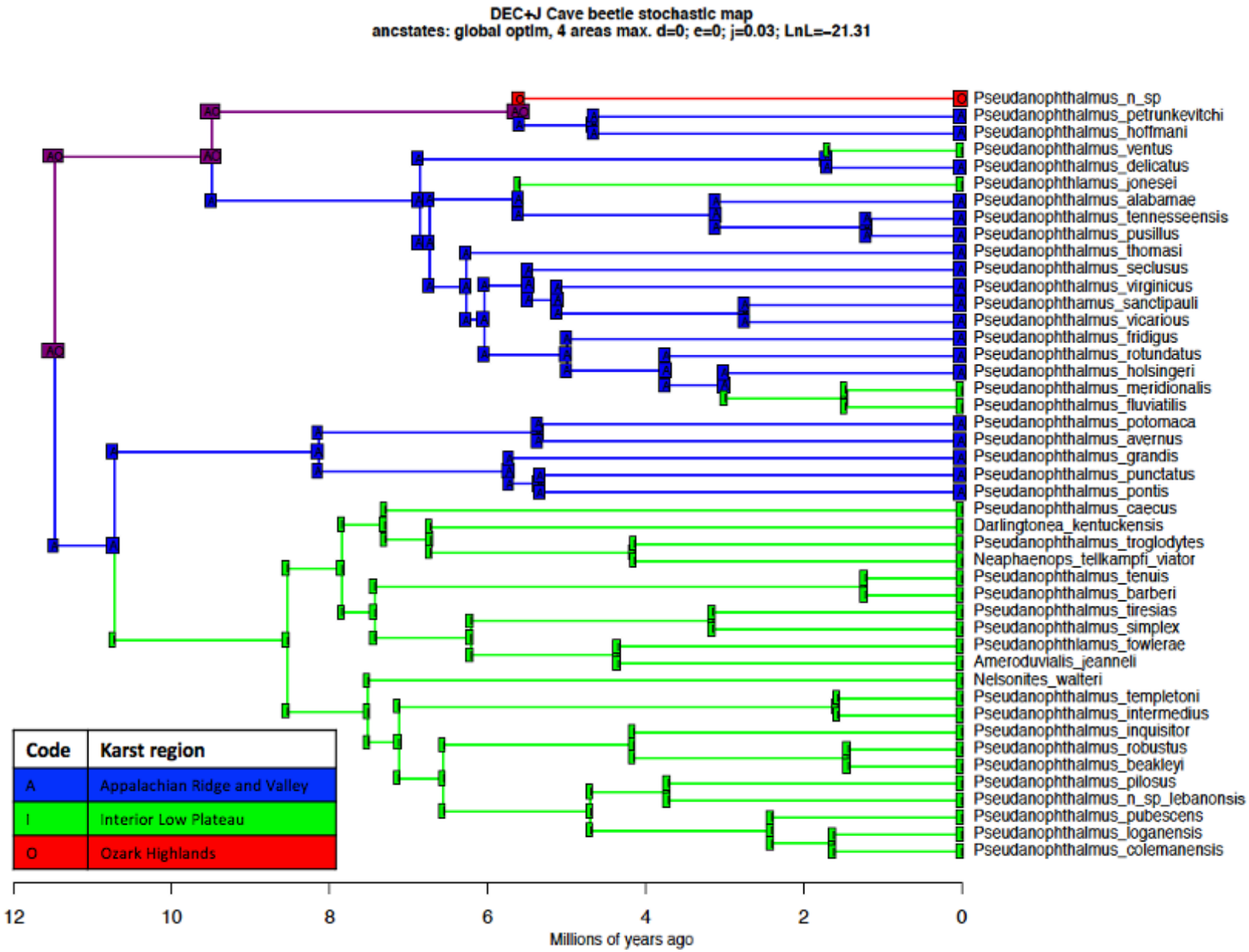


Figure D.6 Ancestral area estimation for cave trechine beetles from eastern North America based on the preferred DEC+J model. Ancestral areas were estimated across the time-calibrated phylogeny inferred from 75% complete concatenated UCE matrix. Most probable ancestral karst region range at each node shown. Corner positions represent the geographic karst region range immediately after a cladogenetic event.

Table D.1 Specimens used in the study, with species group information, locality including karst regions, and voucher reference numbers.

Specimen	Voucher	Species	Species-group	State	County	Cave	Karst_region	Karst_subregion
1	DNA689	*Pseudanophthalmus n.sp.*	*?*	MO	Texas	Pine Hollow Cave	OZK	Ozarks
2	DNA370	*Pseudanophthalmus colemanensis*	*pubescens*	TN	Montgomery	Clarksville Lake Cave	ILP	Western Pennyroyal
3	DNA498	*Pseudanophthalmus loganensis*	*pubescens*	TN	Robertson	Bradley Hill Caverns	ILP	Western Pennyroyal
4	DNA385	*Pseudanophthalmus fluviatilis*	*engelhardti*	AL	Morgan	Talucah Cave	ILP	Cumberland Plateau
5	DNA303	*Pseudanophthalmus meridionalis*	*engelhardti*	AL	Marshall	Beech Spring Cave	ILP	Cumberland Plateau
6	DNA474	*Pseudanophthalmus intermedius*	*intermedius*	TN	Franklin	Dry Cave	ILP	Cumberland Plateau
7	DNA570	*Pseudanophthalmus ventus*	*hirsutus*	TN	Marion	Dancing Fern Cave	ILP	Sequatchie Valley
8	DNA571	*Pseudanophthalmus templetoni*	*intermedius*	TN	Grundy	Skull Cave	ILP	Cumberland Plateau
9	DNA510	*Pseudanophthalmus tiresias*	*cumberlandus*	TN	DeKalb	Indian Grave Point Cave	ILP	Highland Rim
10	DNA404	*Pseudanophthlamus jonesei*	*jonesei*	TN	Cumberland	McCullough Sump Cave	ILP	Sequatchie Valley
11	DNA315	*Pseudanophthalmus robustus*	*robustus*	TN	White	John Henry Demps Cave	ILP	Cumberland Plateau
12	DNA289	*Pseudanophthalmus n. sp. lebanonensis*	*menetriesi*	TN	Wilson	Shell Caverns	ILP	Nashville Basin
13	DNA600	*Pseudanophthalmus simplex*	*simplex*	TN	Jackson	Flatt Cave	ILP	Highland Rim
14	DNA240	*Nelsonites walteri*		TN	Putnam	Stamps Cave	ILP	Cumberland Plateau
15	DNA511	*Pseudanophthalmus beakleyi*	*robustus*	TN	Fentress	Hurricane Maze Cave	ILP	Cumberland Plateau
16	DNA410	*Pseudanophthlamus fowlerae*	*simplex*	TN	Clay	Shankey Branch Cave	ILP	Highland Rim
17	DNA574	*Pseudanophthalmus inquisitor*	*cumberlandus*	TN	Clay	JC Melton Cave	ILP	Cumberland Plateau
18	DNA257	*Pseudanophthalmus pubescens*	*pubescens*	KY	Barren	L & N Railroad Cave	ILP	Western Pennyroyal
19	DNA656	*Neaphaenops tellkampfi viator*		KY	Green	Scotts Cave	ILP	Western Pennyroyal
20	DNA296	*Ameroduvialis jeanneli*		KY	Pulaski	Drowned Rat Cave	ILP	Cumberland Plateau
21	DNA467	*Darlingtonea kentuckensis*		KY	Rockcastle	Fletcher Spring Cave	ILP	Cumberland Plateau
22	DNA618	*Pseudanophthalmus pilosus*	*menetriesi*	KY	Hardin	Cassell Cave	ILP	Western Pennyroyal
22	DNA607	*Pseudanophthalmus barberi*	*tenuis*	KY	Breckinridge	Webster Cave	ILP	Western Pennyroyal
23	DNA525	*Pseudanophthalmus tenuis*	*tenuis*	IN	Harrison	Binkley Cave	ILP	Western Pennyroyal

24	DNA544	*Pseudanophthalmus troglodytes*	*barri*	KY	Jefferson	Eleven Jones Cave	ILP	Outer Bluegrass
25	DNA554	*Pseudanophthalmus caecus*	*horni*	KY	Woodford	Richardson Spring Cave	ILP	Inner Bluegrass
26	DNA748	*Pseudanophthalmus alabamae*	*alabamae*	AL	DeKalb	Manitou Cave	APP	Wills Valley
27	DNA313	*Pseudanophthalmus tennesseensis*	*tennesseensis*	TN	Roane	Eblen Cave	APP	Ridge and Valley
28	DNA530	*Pseudanophthalmus pusillus*	*tennessensis*	TN	Anderson	Martin Cave	APP	Ridge and Valley
29	DNA593	*Pseudanophthalmus rotundatus*	*engelhardti*	TN	Claiborne	Kings Saltpeter Cave	APP	Ridge and Valley
29	DNA492	*Pseudanophthalmus delicatus*	*hirsutus*	TN	Claiborne	Kings Saltpeter Cave	APP	Ridge and Valley
30	DNA631	*Pseudanophthalmus holsingeri*	*engelhardti*	VA	Lee	Young-Fugate Cave	APP	Ridge and Valley
31	DNA541	*Pseudanophthalmus fridigus*	*hypolithos*	KY	Bell	Ice Box Cave	APP	Pine Mountain
32	DNA632	*Pseudanophthalmus thomasi*	*jonesei*	VA	Scott	Blair-Collins Cave	APP	Ridge and Valley
33	DNA670	*Pseudanophthalmus seclusus*	*jonesei*	VA	Scott	Kerns No.1 Cave	APP	Ridge and Valley
34	DNA758	*Pseudanophthamus sanctipauli*	*hubrichti*	VA	Russell	Banners Corner Cave	APP	Ridge and Valley
35	DNA685	*Pseudanophthalmus virginicus*	*grandis*	VA	Tazewell	Hugh Young Cave	APP	Ridge and Valley
35	DNA766	*Pseudanophthalmus vicarious*	*hubrichti*	VA	Tazewell	Ward Cave	APP	Ridge and Valley
36	DNA781	*Pseudanophthalmus hoffmani*	*petrunkevitchi*	VA	Bland	Banes Spring Cave	APP	Ridge and Valley
37	DNA582	*Pseudanophthalmus punctatus*	*pusio*	VA	Giles	Smokehole Cave	APP	Ridge and Valley
38	DNA538	*Pseudanophthalmus grandis*	*grandis*	WV	Greenbrier	Culverson Creek Cave	APP	Greenbrier Karst
39	DNA577	*Pseudanophthalmus pontis*	*pusio*	VA	Rockbridge	Bradys Cave	APP	Ridge and Valley
40	DNA756	*Pseudanophthalmus potomaca*	*hubbardi*	VA	Highland	Vandevander Cave	APP	Ridge and Valley
41	DNA507	*Pseudanophthalmus avernus*	*hubbardi*	VA	Rockingham	Endless Caverns	APP	Ridge and Valley
42	DNA767	*Pseudanophthalmus petrunkevitchi*	*petrunkevitchi*	VA	Warren	Brother Daves Cave	APP	Ridge and Valley
	DNA787	*Trechus obtusus*	*outgroup*	OR				
	DNA786	*Trechoblemus westcotti*	*outgroup*	OR				
	DNA788	*Trechus humbolti*	*outgroup*	OR				

Table D.2 Comparison of dispersal-extinction-cladogenesis (DEC) models with jump dispersal (+J) and without (+J) for cave trechine beetles based on their dispersal within major karst region. Abbreviations as follows: LnL, loglikelihood; numparams, number of parameters in each model; d, dispersal rate; e, extinction rate; j, founder-event speciation rate; AIC, Akaike Information Criterion; LRT, likelihood-ratio test.

Model	LnL	numparams	d	e	j	AIC	AIC_wt	LRT pval
DEC	-30.31	2	0.0085	1.00E-12	0	64.62	0.0003	
DEC+J	-21.31	3	1.00E-12	1.00E-12	0.03	48.63	1	2.20E-05

Appendix E. Chapter 3 Supplemental Figures and Tables

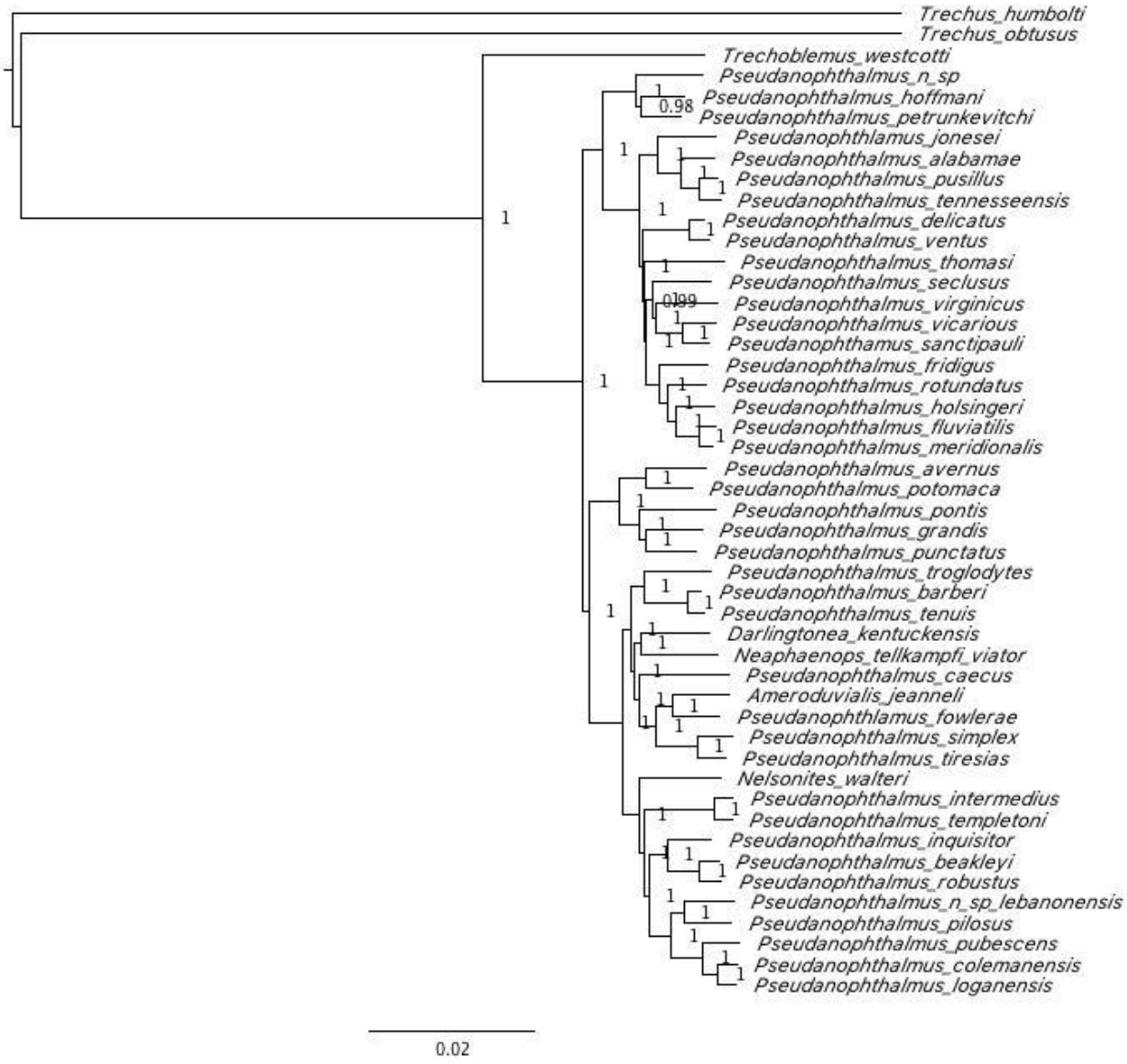


Figure E.1 Bayesian phylogeny of cave trechine beetles from eastern North America inferred from 50% complete concatenated UCE matrix. Numbers indicate support values (Bayesian posterior probability) for all nodes.

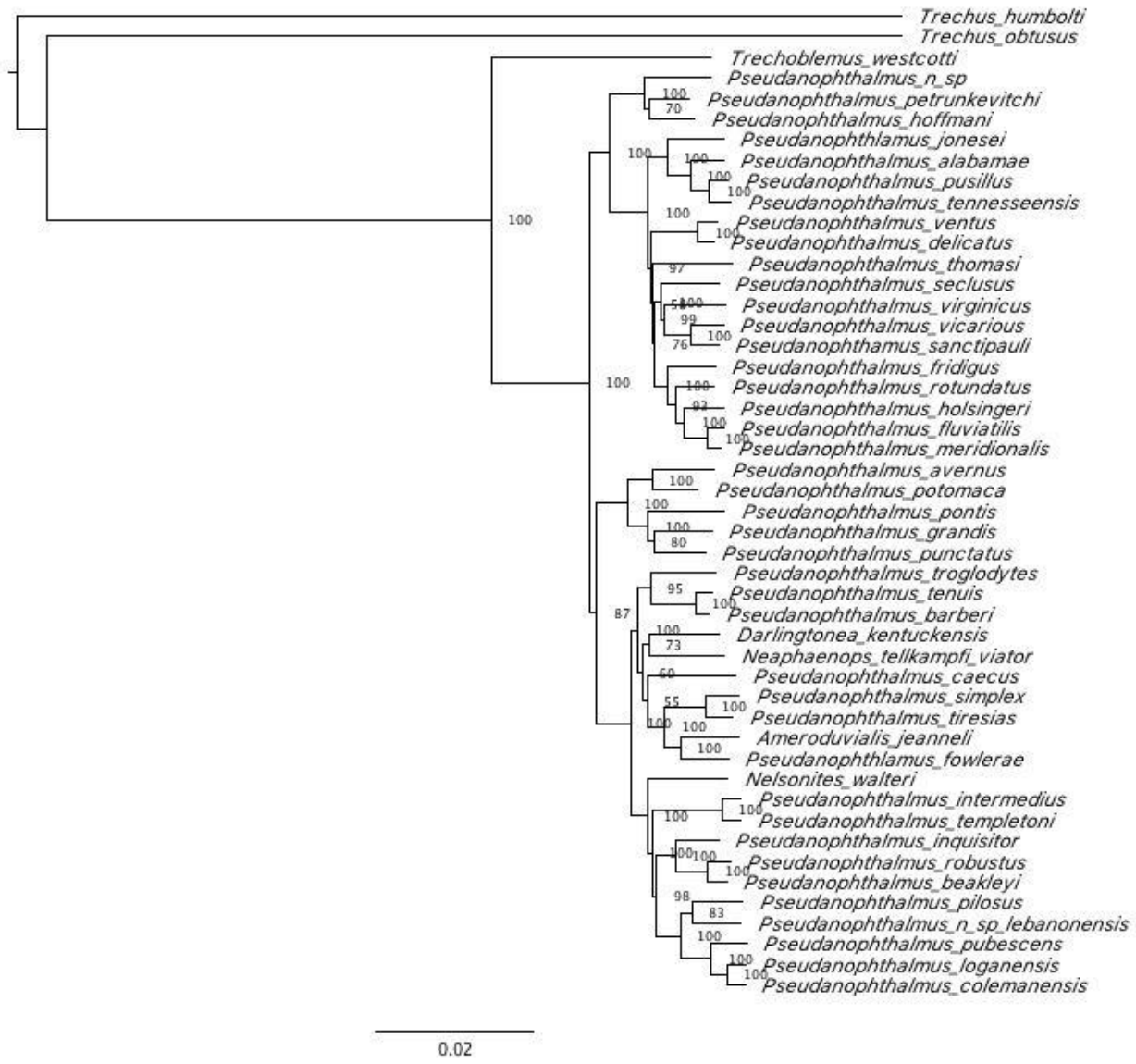


Figure E.2 Maximum-likelihood phylogeny of cave trechine beetles from eastern North America inferred from 50% complete concatenated UCE matrix. Numbers indicate support values (maximum-likelihood) for all nodes.

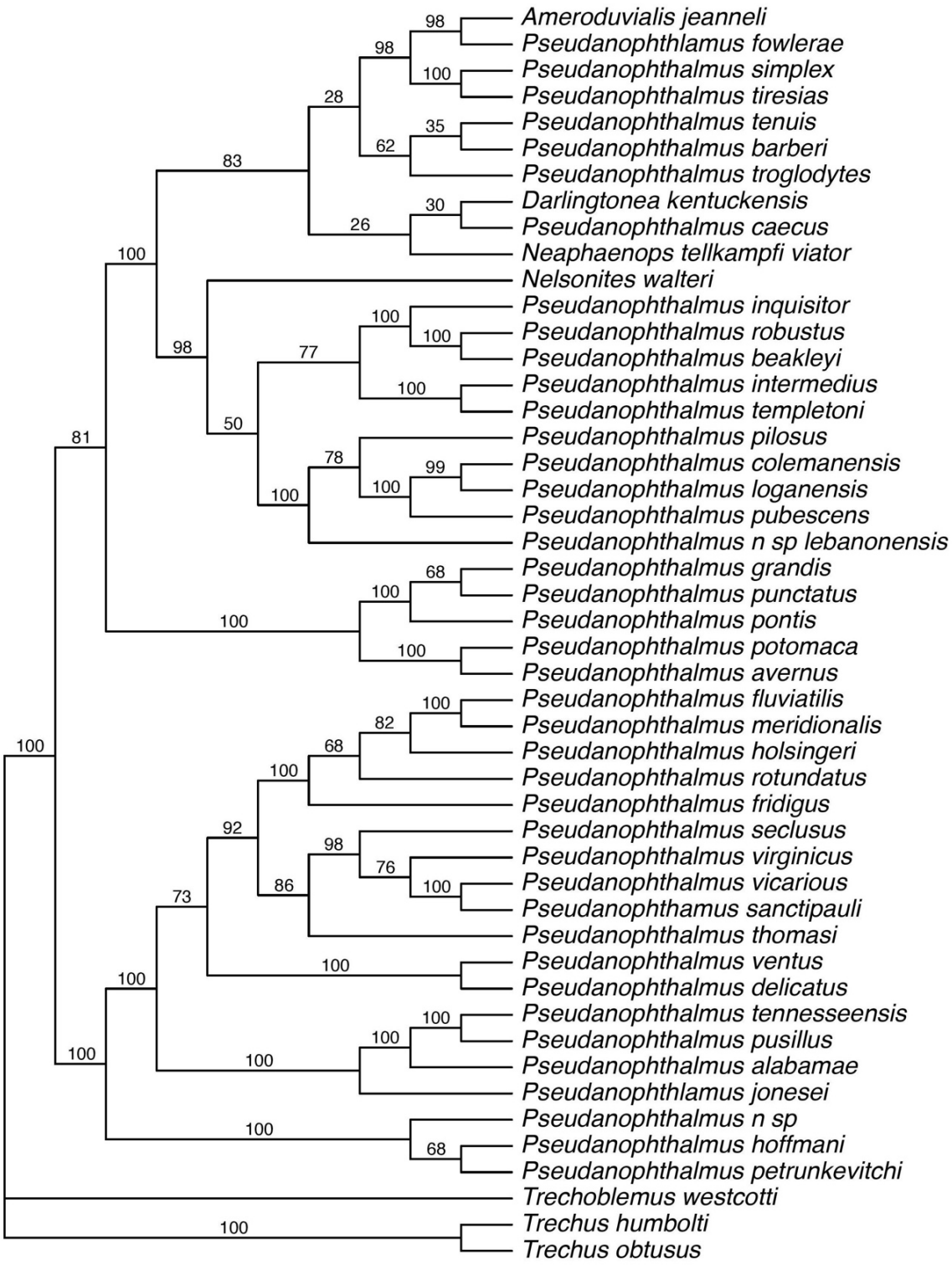


Figure E.3 Phylogenetic relationships among the cave trechine beetles based on SVDQuartets coalescent species trees with 50% majority rule consensus for SVDQuartets. Node values indicate bootstrap support values.

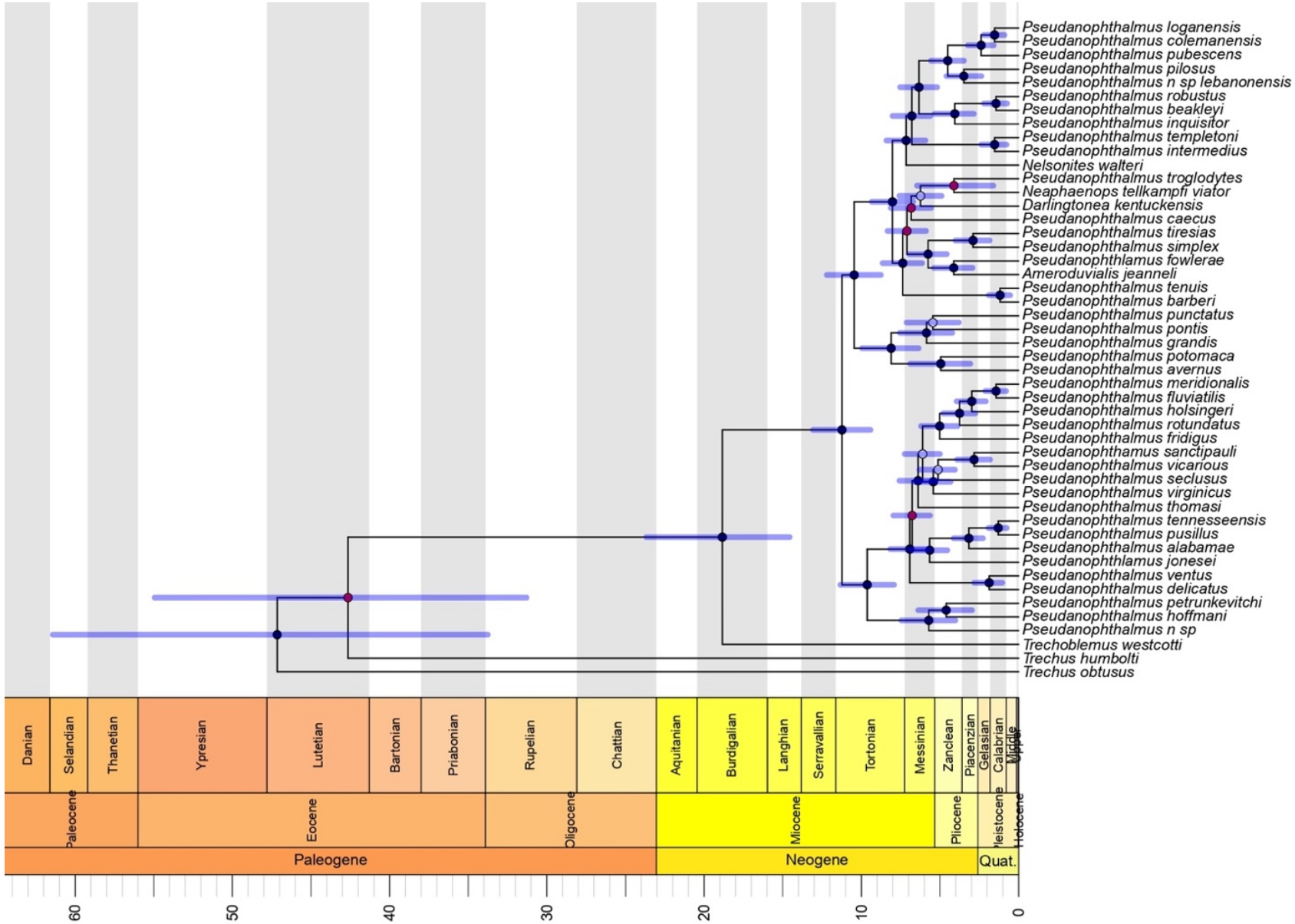


Figure E.4 Time-calibrated maximum clade credibility tree with *Trechus* outgroup inferred from 75% complete concatenated UCE matrix, summarized by TreeAnnotator, and plotted with a geological time scale using the strap package in R. Phylogeny dated using a Bayesian relaxed clock method in BEAST. Branches are proportional to time in millions of years. The 95% confidence intervals for the ages of basal branches in the tree are indicated with transparent blue bars. The internal nodes of the tree are indicated with circles, where circles mark nodes with posterior probability: ● ≥ 0.95, 0.95 > ○ ≥ 0.75, 0.75 > ●.

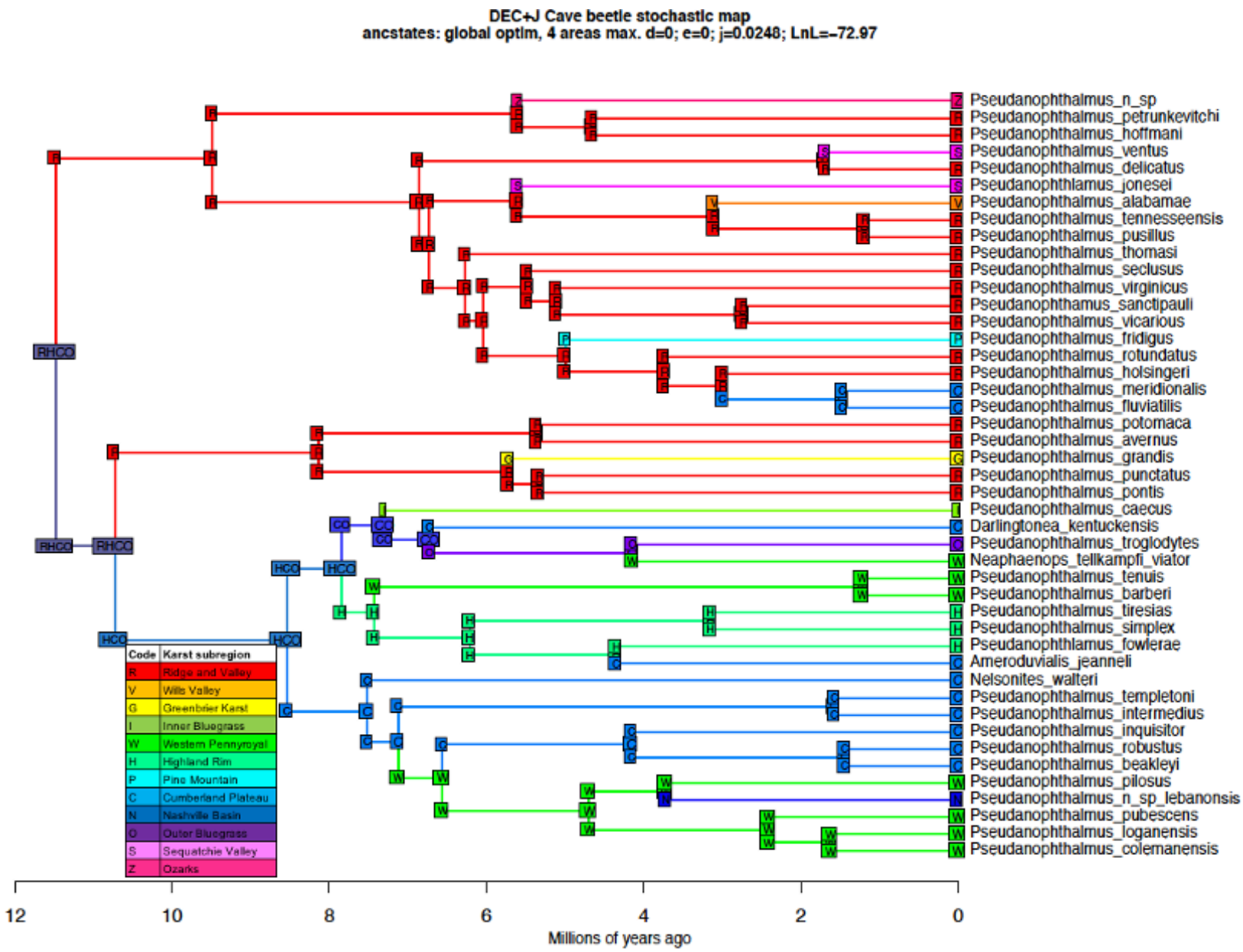


Figure E.5 Ancestral area estimation for cave trechine beetles from eastern North America based on the preferred DEC+J model. Ancestral areas were estimated across the time-calibrated phylogeny inferred from 75% complete concatenated UCE matrix. Most probable ancestral karst sub region range at each node shown. Corner positions represent the geographic karst sub region range immediately after a cladogenetic event.

Table E.1 Information on specimen vouchers, sample, data yield (Mb), raw Illumina reads before and after quality filtering and trimming, and SRA accession numbers.

Voucher	Species	Yield (Mb)	Raw Reads	Raw Reads after QC	SRA accession
DNA240	*Nelsonites walteri*	389	1290679	1264267	SAMN31468728
DNA257	*Pseudanophthalmus pubescens*	751	2489047	2433261	SAMN31468729
DNA289	*Pseudanophthalmus n. sp. lebanonensis*	1014	3359680	3284768	SAMN31468730
DNA296	*Ameroduvialis jeanneli*	1433	4746236	4661544	SAMN31468731
DNA303	*Pseudanophthalmus meridionalis*	1185	3926710	3849578	SAMN31468732
DNA313	*Pseudanophthalmus tennesseensis*	2648	8770180	8649185	SAMN31468733
DNA315	*Pseudanophthalmus robustus*	1218	4034493	3961570	SAMN31468734
DNA370	*Pseudanophthalmus colemanensis*	2234	7400154	7291619	SAMN31468735
DNA385	*Pseudanophthalmus fluviatilis*	963	3190448	3123659	SAMN31468736
DNA404	*Pseudanophthlamus jonesei*	1574	5213092	5133024	SAMN31468737
DNA410	*Pseudanophthlamus fowlerae*	2701	8944267	8824422	SAMN31468738
DNA467	*Darlingtonea kentuckensis*	1903	6304505	6211516	SAMN31468739
DNA474	*Pseudanophthalmus intermedius*	1545	5118836	5036087	SAMN31468740
DNA492	*Pseudanophthalmus delicatus*	1263	4185208	4120192	SAMN31468741
DNA498	*Pseudanophthalmus loganensis*	1988	6585439	6500856	SAMN31468742
DNA507	*Pseudanophthalmus avernus*	2269	7515566	7403566	SAMN31468743
DNA510	*Pseudanophthalmus tiresias*	890	2949246	2902518	SAMN31468744
DNA511	*Pseudanophthalmus beakleyi*	2132	7061910	6957591	SAMN31468745
DNA525	*Pseudanophthalmus tenuis*	1982	6564135	6469009	SAMN31468746
DNA530	*Pseudanophthalmus pusillus*	3450	11426668	11283637	SAMN31468747
DNA538	*Pseudanophthalmus grandis*	991	3283450	3236038	SAMN31468748
DNA541	*Pseudanophthalmus fridigus*	330	1093094	1071502	SAMN31468749
DNA544	*Pseudanophthalmus troglodytes*	46	154671	146144	SAMN31468750
DNA554	*Pseudanophthalmus caecus*	136	453183	441279	SAMN31468751
DNA570	*Pseudanophthalmus ventus*	1230	4074848	4022862	SAMN31468752
DNA571	*Pseudanophthalmus templetoni*	1811	5997310	5918933	SAMN31468753

DNA574	*Pseudanophthalmus inquisitor*	850	2814717	2764858	SAMN31468754
DNA577	*Pseudanophthalmus pontis*	1493	4944655	4875265	SAMN31468755
DNA582	*Pseudanophthalmus punctatus*	2048	6783111	6676826	SAMN31468756
DNA593	*Pseudanophthalmus rotundatus*	2839	9401238	9258512	SAMN31468757
DNA600	*Pseudanophthalmus simplex*	1637	5423131	5334661	SAMN31468758
DNA607	*Pseudanophthalmus barberi*	1530	5068330	4998957	SAMN31468759
DNA618	*Pseudanophthalmus pilosus*	1556	5153059	5085574	SAMN31468760
DNA631	*Pseudanophthalmus holsingeri*	2181	7222348	7103880	SAMN31468761
DNA632	*Pseudanophthalmus thomasi*	1655	5482482	5396298	SAMN31468762
DNA656	*Neaphaenops tellkampfi viator*	1400	4636586	4532101	SAMN31468763
DNA670	*Pseudanophthalmus seclusus*	1522	5040936	4962632	SAMN31468764
DNA685	*Pseudanophthalmus virginicus*	2571	8516008	8386082	SAMN31468765
DNA689	*Pseudanophthalmus n.sp.*	1387	4593760	4526457	SAMN31468766
DNA748	*Pseudanophthalmus alabamae*	1927	6383048	6290294	SAMN31468767
DNA756	*Pseudanophthalmus potomaca*	1799	5959574	5849553	SAMN31468768
DNA758	*Pseudanophthamus sanctipauli*	2085	6904911	6799444	SAMN31468769
DNA766	*Pseudanophthalmus vicarious*	1992	6598288	6498008	SAMN31468770
DNA767	*Pseudanophthalmus petrunkevitchi*	1955	6473609	6386504	SAMN31468771
DNA781	*Pseudanophthalmus hoffmani*	1858	6154080	6053781	SAMN31468772
DNA786	*Trechoblemus westcotti*	1103	3652403	3583207	SAMN31468773
DNA787	*Trechus obtusus*	108	358362	342995	SAMN31468774
DNA788	*Trechus humbolti*	591	1957105	1900383	SAMN31468775

Table E.2 Comparison of dispersal-extinction-cladogenesis (DEC) models with jump dispersal (+J) and without (+J) for cave trechine beetles based on their dispersal within karst sub region. Abbreviations as follows: LnL, loglikelihood; numparams, number of parameters in each model; d, dispersal rate; e, extinction rate; j, founder-event speciation rate; AIC, Akaike Information Criterion; LRT, likelihood-ratio test.

Model	LnL	numparams	d	e	j	AIC	AIC_wt	LRT pval
DEC	-101.8	2	0.0048	0.026	0	207.6	8.00E-13	
DEC+J	-72.97	3	1.00E-12	1.00E-12	0.025	151.9	1	3.10E-14