8-2017

BERINGIAN BIOGEOGRAPHY: PATTERNS OF INTERCONTINENTAL DISPERSAL AND THE STRUCTURING OF A HOLARCTIC TAPEWORM GENUS

Genevieve M. Haas
Northern Michigan University, ghaas@nmu.edu

BERINGIAN BIOGEOGRAPHY: PATTERNS OF INTERCONTINENTAL DISPERSAL AND THE STRUCTURING OF A HOLARCTIC TAPEWORM GENUS

By

Genevieve M. S. Haas

THESIS

Submitted to
Northern Michigan University
In partial fulfillment of the requirements
For the degree of

MASTER OF SCIENCE

Office of Graduate Education and Research

August 2017

SIGNATURE APPROVAL FORM

BERINGIAN BIOGEOGRAPHY: PATTERNS OF INTERCONTINENTAL DISPERSAL AND THE STRUCTURING OF A HOLARCTIC TAPEWORM GENUS

This thesis by Genevieve Haas is recommended for approval by the student's Thesis Committee and Department Head in the Department of Biology and by the Interim Director of Graduate Education and Research.

_____

Committee Chair: Dr. Kurt Galbreath Date

_____

First Reader: Dr. Eric Hoberg Date

_____

Second Reader: Dr. Alec Lindsay Date

_____

Third Reader: Dr. Neil Cumberlidge Date

_____

Department Head: Dr. John Rebers Date

_____

Dr. Lisa Eckert Date
Interim Director of Graduate Education and Research

ABSTRACT

The opening and closing of the Bering Land Bridge due to Pleistocene climate fluctuations facilitated the exchange of taxa between the Palearctic and Nearctic. While many studies have worked toward elucidating the role of Beringia in assembling northern faunas, relatively little work has focused on parasites. Here I examine the number and direction of transberingian colonization events within the Holarctic tapeworm genus, *Arostrilepis* Mas-Coma & Tenora, 1997. I performed maximum likelihood and multi-locus coalescent phylogenetic reconstructions using mitochondrial and nuclear DNA sequences. Biogeographic ancestral range estimations were conducted on the resulting species phylogeny. My systematic reconstructions reveal as many as 16 *Arostrilepis* lineages that could represent previously undescribed species-level diversity. Biogeographic estimates strongly indicate that *Arostrilepis* experienced at least four eastward transberingian dispersals associated with *Microtus, Myodes,* and *Lemmus* hosts. Comparing the *Arostrilepis* colonization history with the pattern of its host associations shows that host-switching is prevalent in its history particularly following the Nearctic colonization associated with *Microtus* voles. Evidence also suggests that during the colonization event associated with lemmings, the direction of *Arostrilepis* colonization may have been counter to that of its hosts. These results highlight the complex history of faunal assembly associated with Beringian mammal-parasite assemblages.

ACKNOWLEDGEMENTS

I would like to express my deepest thanks to my advisor Dr. Kurt Galbreath for guidance and support throughout this project, not to mention being the person to introduce me to parasitology. To my committee members Dr. Eric Hoberg, Dr. Alec Lindsay, and Dr. Neil Cumberlidge, thank you for providing additional insights and feedback.

I am forever thankful to Neil Poglese, my parents, family, and friends for their love and support throughout all of my pursuits. My fellow Biology graduate students provided both support and insightful conversation during this project; for that, I am especially grateful to my labmates Heather Toman and Sarah Gallagher.

I would also like to thank Sigma Xi, Northern Michigan University, the American Society of Mammalogists, and the American Society of Parasitologists for funding in support of my research and travel to professional conferences. Finally, I would like to acknowledge the many people who took part in the collection efforts that made my project possible.

This thesis is formatted following the guidelines of the Journal of Biogeography.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

North America and Eurasia are currently separated by the Bering Strait, but up until about the Miocene-Pliocene Transition (4.8-5.5 Million Years Ago; MYA) there existed a land bridge between the two continents (Marincovich & Gladenkov, 1999; Gladenkov & Gladenkov, 2004). Since that initial separation, this region (Beringia) has alternated between allowing intercontinental exchange of terrestrial species, and presenting a barrier to such dispersal. The status of this ephemeral connection is tightly tied to Earth's climate cycles. During interglacial periods, warmer temperatures result in higher sea levels, inundating the Bering Land Bridge. Glacial periods are marked by lower sea levels due to water being locked up in continental ice sheets, which exposed the Bering Land Bridge. Because of these processes, the region that spans the strait has played a key role in the structuring of high latitude species assemblages. As the land bridge was exposed during glacial periods, some taxa expanded into the newly available region. Other species retracted into Beringia as conditions in their previous ranges became inhospitable due to the cooling climate (Stewart et al., 2010). In this regard Beringia was an ice free refugium, made up of a habitat mosaic and hosting a variety of taxa (Guthrie, 1968; Hoffmann, 1981; Guthrie, 1984). Beringia was a center of diversification for many northern species, as they adapted to the conditions of the region (Sher, 1999). Beringia also served as a dispersal corridor between the northern continents. The timing of the presence of appropriate conditions and the locations of the continental ice sheets mediated which flora and fauna were exchanged and when. The ebb and flow of environmental conditions within the region added additional constraints to this ecological filter (Meiri et al., 2014). In its capacity as a joint refugium and dispersal corridor, Beringia had profound and lasting effects on the biological structuring

of high-latitude species assemblages (Waltari et al., 2007; Hoberg et al., 2012; Hope et al., 2013). The repeated episodes of species and population expansion and retraction have shaped the mosaic of diversity observed today.

Studies elucidating the history and structure of these faunal assemblages have begun to reveal the timing, directionality, and role of transberingian dispersal events, particularly for several northern mammals across the Holarctic. These include voles (Conroy & Cook, 1999, 2000; Brunhoff et al., 2003; Galbreath & Cook, 2004; Kohli et al., 2014b, 2014a), lemmings (Fedorov, 1999, 1999; Fedorov & Goropashnaya, 1999; Fedorov et al., 2003), shrews (Hope et al., 2013), ground squirrels (Galbreath et al., 2011), and others (Repenning, 2001; Cook et al., 2005; Waltari et al., 2007). Many taxa exhibit histories involving colonization of Beringia from Asia with subsequent differentiation promoted by isolation created by glacial barriers (e.g., *Myodes* voles and *Microtus* voles). Although colonization out of the Beringian refugium has been a common theme, for some taxa (e.g., *Lemmus sibiricus* and *Lemmus trimucronatus*) vicariance between Palearctic and Nearctic sister species was driven by the inundation of the Bering Strait during interglacial periods and subsequent isolation in refugia other than Beringia led to the population structure seen today (Fedorov et al., 1999, 2003).

Transberingian patterns revealed in the studies of northern mammal biogeographic histories provide a framework for investigations of the processes shaping those of their parasites (Cook et al., 2005). Host organisms are typically the primary facilitator of parasite dispersal. Because of this, much emphasis has been placed on detecting instances of host-tracking/co-speciation, where the parasite phylogeny tightly follows that of the host; when the host speciates, so does the parasite (Hafner & Nadler,

1988, 1990). But parasites, both with complex and direct lifecycles, do not always strictly follow their hosts. When the host expands geographically the parasite may be left behind or fail to establish in the new region resulting in what has been termed "missing the boat". Also, parasites are known to speciate when their definitive host does not (duplication) or to carry on as a single species despite diversification by their hosts (inertia) (Paterson & Banks, 2001). Another deviation from host-tracking is host-switching where the parasite successfully colonizes a new host species. Host-switching is often linked to the oscillation between periods of environmental perturbation and stability (Hoberg & Brooks, 2010). Disturbance creates the opportunity for taxon pulses, episodes of faunal mixing brought on as species distributions shift in response to the environmental stimulus. Parasites may seem to be specialists when they are only found to be associated with a single host species, but this can be a consequence of not having an opportunity to colonize other potential hosts. Taxon pulses can bring parasites into contact with novel susceptible hosts (Hoberg & Brooks, 2008). If a potential host and its environment provide the resources (physiological, intermediate hosts, etc.) necessary for the parasite to complete its life cycle, the parasite may colonize this new host and over time, as conditions stabilize, diverge from the original population or species in a process called ecological fitting. These cycles can facilitate radiation and lead to reticular patterns of diversity across geographic landscapes.

Parasite biogeography can reveal otherwise cryptic movements of host species (Galbreath & Hoberg, 2012), but also can provide a unique lens through which to examine the broader assembly of complex Holarctic biological systems (Cook et al., 2005; Hoberg et al., 2012). Comparing host and parasite phylogenies allows inferences

regarding aspects of history such as geographic sources for past colonization events, identities of ancestral hosts, and timing of parasite speciation events relative to host or environmental histories. Host-parasite perspectives facilitate the linking of host ecology and history with the changing environment (Brooks & McLennan, 2002; Hoberg & Brooks, 2010; Hoberg et al., 2012; Brooks et al., 2015; Hoberg & Brooks, 2015).

Here I investigate the processes that shaped the transberingian biogeographic history of a Holarctic genus of tapeworms, *Arostrilepis* (Mas-Coma & Tenora, 1997). Definitive hosts of these tapeworms are most often members of the subfamily Arvicolinae, a group of rodents with a well-studied history of multiple transberingian dispersals. Once considered to be a single morphologically variable and geographically widespread species, *Arostrilepis* now includes 13 nominal species as well as other discrete genetic lineages that may represent additional unnamed species (Makarikov et al., 2011; Makarikov & Kontrimavichus, 2011; Makarikov et al., 2012, 2013; Makarikov & Hoberg, 2016). This study builds upon recent morphological and molecular studies that revealed *Arostrilepis* to be a species assemblage and has three primary objectives. First, to determine the scope of *Arostrilepis* species diversity and geography through genetic identification of newly sampled specimens collected from across the range of the species complex. Second, to generate a robust *Arostrilepis* phylogeny. Third, to use the improved resolution of *Arostrilepis* geographic distributions and the phylogeny to address two biogeographic questions: (1) How many times did *Arostrilepis* cross the Bering Land Bridge? (2) Does *Arostrilepis* follow the general trend of eastward colonization (from Eurasia to the Nearctic), or does its biogeographic history include both eastward and westward colonization events?

METHODS

Study System

Arostrilepis is found broadly across northern Eurasia and North America. Of the 13 described species, five are restricted to the Palearctic, five to the Nearctic, and three have Holarctic distributions (Makarikov et al., 2013; Makarikov & Hoberg, 2016). Most nominal species of *Arostrilepis* are associated with arvicoline rodents (i.e., voles and lemmings). Exceptions include associations with *Neotominae*, *Heteromyidae*, and *Geomyidae*, and incidental infections of *Sciuridae*. Each parasite species is generally associated with a specific host genus rather than a host species, though in some cases there does appear to be a dominant association with a particular species (e.g. *Arostrilepis macrocirrosa* Makarikov, Galbreath & Hoberg, 2013 to *Myodes rutilus*) (Makarikov et al., 2013; Makarikov & Hoberg, 2016). Because of its close association with arvicoline rodents, which arose in Eurasia during the Pliocene, *Arostrilepis* is hypothesized to have also originally diversified in Eurasia (Chaline & Graf, 1988; Repenning, 2001; Hoberg et al., 2012; Makarikov et al., 2013).

Specimens

To increase resolution of geographic distributions for *Arostrilepis* species and to clarify host-parasite associations, I acquired 70 *Arostrilepis* specimens collected through the Beringian Coevolution Project (BCP; Cook et al., In Press, 2005; Table 1). The BCP is a long-running effort to build a specimen-based infrastructure for the study of high-latitude mammal-parasite assemblages. Additionally, 12 specimens from field collections by Antti Lavikainen (University of Helsinki, Finland), and 10 specimens from field

collections by Nikolai Dokuchaev (Institute of Biological Problems of the North, Magadan, Russia), were included in my dataset (Table 1). Specimens in this dataset were collected throughout the Holarctic region (Figures 1, 2, 3). Relative to past examinations of *Arostrilepis* diversity (Makarikov et al., 2013; Makarikov & Hoberg, 2016), my dataset has higher geographic representation from the temperate Nearctic, and central and western Palearctic. Hosts of the specimens in my study belong to seven genera: *Microtus, Myodes, Lemmus, Peromyscus, Synaptomys, Thomomys,* and *Cricetulus* (Table 1). This is the first molecular systematic study to include *Arostrilepis* specimens from *Thomomys* and *Cricetulus* hosts.

Molecular data collection

From each individual tapeworm, 3-10 posterior proglottids were sampled and whole genomic DNA was extracted from these tissues using a Qiagen™ DNeasy Tissue Kit®. I used both mitochondrial and nuclear markers to provide multiple independent perspectives of *Arostrilepis* evolution. Specifically, part of the mitochondrial gene cytochrome *b* (cyt-*b*; ~570 base pairs), as well as nuclear 28S ribosomal RNA (28S; ~1340 base pairs) and the second internal transcribed spacer of nuclear ribosomal DNA (ITS2; ~740 base pairs) were PCR amplified.

I amplified all three loci using the following primers: HYM01, HYM08, and HYMLEM02 (Makarikov et al., 2013) were used to amplify cyt-*b,* LSU5 and 1200R (Littlewood, 2000; Lockyer et al., 2003; Haukisalmi et al., 2010) were used to amplify 28S, and 3S and A28 (Okamoto et al., 1997) were used to amplify ITS2. Annealing temperatures were set to 50ºC for cyt-*b* and ITS2, and 52ºC for 28S. PCR products were sequenced in both directions on an ABI 3100 genetic analyzer (Applied Biosystems Inc.,

Foster City, CA) using ABI PRISM® BigDyeTM sequencing chemistry and sequences were checked by eye and assembled using GENEious v6 (Kearse et al., 2012).

Two species of *Arostrilepis*, *Arostrilepis rauschorum* Makarikov, Galbreath, & Hoberg, 2013 and *Arostrilepis microtis* Gulyaev & Chechulin, 1997, contain nuclear copies of the cyt-*b* marker, which co-amplifies with the target when using primers HYM01 and HYM08 (Makarikov et al., 2013). DNA from individuals whose sequence electropherograms contained double peaks characteristic of co-amplification was re-amplified using either primers HYM29 (5'TGATTAATATTATACGACGT) and HYM30 (5'TGTGCAAATAAAATAAATGT) if Nearctic (i.e., potentially *A. rauschorum*) or primers HYM32 (5'AATGTAAAAACATTAAGCCC) and HYM33 (5'TACGACGTAATTTAATTGAT) if Palearctic (i.e., potentially *A. microtis*). These primers bind to the target portion of cyt-*b* slightly offset from the other primer set in order to avoid co-amplification and obtain clean mitochondrial cyt-*b* sequences. Additionally, 14 cyt-*b* sequences (Makarikov et al., 2013), four 28S sequences (Haukisalmi et al., 2010), and two ITS2 sequences (Galbreath et al., 2013) were obtained from GenBank. These sequences are associated with specimens whose species identities have been morphologically confirmed, allowing me to compare the new sequences I generated to those with definitive identifications (Table 1). Four species, *Arostrilepis horrida* (Linstow, 1901), *Arostrilepis kontrimavichusi* Makarikov & Hoberg, 2016, *Arostrilepis mariettavogeae* Makarikov, Gardner, & Hoberg, 2012, and *Arostrilepis schilleri* Makarikov, Gardner, & Hoberg, 2012 (Makarikov et al., 2012; Makarikov & Hoberg, 2016), do not currently have tissue material or sequence data available and thus were not included in this molecular evaluation.

Phylogenetic Analyses

The sequences from each marker were aligned using MUSCLE (Edgar, 2004), implemented in MEGA version 7 (Kumar et al., 2015). All sequence alignments were checked by eye. I confirmed that newly sequenced individuals belong to the ingroup by generating Neighbor-joining trees in PAUP* version 4.0b10 (Swofford, 2003) with *Hymenolepis diminuta* included as an outgroup.

I ran separate Maximum Likelihood (ML) analyses using RAxML version 8 (Stamatakis, 2014) on all three loci to evaluate diversity within each locus and to determine where the newly collected *Arostrilepis* sequences cluster in relation to confirmed species. Outgroups were not used in these analyses because the only outgroup taxa available are distantly related and poorly sampled, factors which have been shown to negatively affect the accuracy of phylogenetic rooting (Wheeler, 1990; Graham et al., 2002; Huelsenbeck et al., 2002; Gatesy et al., 2007). Partition Finder (Lanfear et al., 2012) was used to determine the model of nucleotide substitution to apply for each locus and to determine the best partitioning scheme for cyt-*b*. Because cyt-*b* is coding the different codon positions have different mutation rates sometimes warranting partitioning with different models of nucleotide substitution for each. Partition Finder was set to make evaluations specifically for analyses using RAxML. Based on the Bayesian Information Criterion (BIC), Partition Finder determined that the best partitioning scheme for cyt-*b* in RAxML is two partitions, one in which codon positions 1 and 2 are combined, and another that includes only codon position 3. The GTRGAMMA+I model was selected for application to both cyt-*b* partitions as well as ITS2, and the GTRGAMMA model was selected for 28S. Using these parameters, I ran the RAxML analysis five times for each

locus from different random starting seeds. Nodal support was assessed by conducting 1000 rapid bootstrap searches within each run.

Putative taxon identities were assigned to new *Arostrilepis* sequences based on where they clustered in relation to the GenBank sequences with confirmed species identities in the ML analysis. I then used MEGA7 (Kumar et al., 2015) to calculate mean uncorrected pairwise genetic distances for each locus both between and within groups for described species and suspected lineages (Table 2 and Supplemental Material).

The *BEAST method (Heled & Drummond, 2010), implemented in BEAST v2.4.4.0 (Bouckaert et al., 2014) was used to generate a phylogeny integrating all 3 loci. This program uses a coalescent approach in reconstructing species phylogenies, allowing it to account for stochastic variation in locus-specific genealogies to recover a joint estimate of the species tree. Using this integrated process can yield a more robust result than concatenating sequence data from multiple loci, which fails to account for phylogenetic variation in the histories of unlinked markers. Being a coalescent-based method, *BEAST reconstructs phylogenetic relationships by estimating where the branches within the tree coalesce. Like the position of other nodes in the tree, *BEAST also samples the root position, therefore rooting of the tree using this approach can be achieved without including outgroup taxa. Because of the problems associated with identifying informative outgroups for this genus, and because *BEAST produces a rooted tree without outgroups, I did not include any in this analysis.

In the multi-species coalescent analysis, *A. macrocirrosa* was split into central Asian and eastern (Beringian) populations in recognition of phylogeographic structure detected within the species (S. Gallagher and K. Galbreath, unpublished data). This was

important for subsequent biogeographic analyses, adding necessary resolution to examine specific colonization hypotheses.

For the *BEAST analysis, Partition Finder (Lanfear et al., 2012) determined the best partitioning scheme for cyt-*b*, this time optimized for a BEAST analysis. Results indicated that in *BEAST, cyt-*b* should be partitioned by each codon position. Appropriate models of nucleotide substitution for these subsets were selected for cyt-*b* codon positions 1 to 3, 28S, and ITS2 alignments under Akaike's Information Criterion (AIC) in jModelTest v2.1.10 (Guindon & Gascuel, 2003; Darriba et al., 2012). Transition rates provided by jModelTest were used as starting points in the *BEAST run to help the analysis converge faster. I performed likelihood ratio rests (LRT) using PAUP* to check for clock-like evolution (Felsenstein, 1988). LRTs failed to reject the strict molecular clock for every alignment except cyt-*b* codon position 3. Based on these results, a relaxed log normal molecular clock was applied to the cyt-*b* position 3 partition and a strict molecular clock was applied to all other loci and partitions. I applied the linear with constant root demographic model and the Yule species tree prior. The *BEAST input file was prepared with these parameter specifications using BEAUti v2.1.3.0 (Bouckaert et al., 2014). The *BEAST analysis was run for 200 million generations and the first 10% of the run was discarded as burn-in. The quality of the run was assessed using TRACER v1.6 (Rambaut et al., 2014) to evaluate effective sample size (ESS) values for model parameters, all of which were greater than 200, and to examine the parameter trend plots for stationarity. The analysis was repeated three times with different random starting seeds to ensure convergence of the parameters in probability space between runs.

Biogeographic Analysis

To examine the biogeographic history of *Arostrilepis*, ancestral species distributions were estimated using the R package BioGeoBEARS (Matzke, 2013a). BioGeoBEARS provides a supermodel where parameters controlling various biogeographic processes can be set, estimated, or excluded, and the fit of the data to the resulting model evaluated. I used the package to run and statistically compare the fit of the data to three biogeographic models: Dispersal Extinction Cladogenesis (DEC) from Lagrange (Ree & Smith, 2008); DIVALIKE, a maximum likelihood version of Dispersal-Vicariance (DIVA) (Ronquist, 1997); and BAYAREA, a likelihood model that mimics the model used in the program BayArea (Landis et al., 2013) and the Bayesian Binary Model of RASP (Yu et al., 2013). These models differ in their underlying assumptions regarding the processes that shape geographic range such as whether sympatry and vicariance are allowed to occur narrowly (at a splitting event one daughter inhabits a single subset of a multi-region ancestral range, while the other inhabits the rest) or widespread (multiple subranges of a wide ancestral range are split between the two daughters) and whether modeled range shifts occur at cladogenesis (DEC and DIVALIKE) or along the branch lengths (BAYAREA) (Matzke, 2013b). BioGeoBEARS offers the option for each of these models to be run with the incorporation of the J parameter, which accounts for founder event speciation. The standard versions of these models allow for ancestral species to transition into a null range that only exists within the model. The null range has been shown to inflate the estimated value of the extinction parameter in runs using these models, and comparisons show that excluding the null range can improve the accuracy of estimations (Massana et al., 2015). Because of this, I

chose to perform my range estimations without the null range. The exclusion of the null range is designated by an asterisk following the names of the models. The analysis was performed using the *BEAST phylogeny. I assigned a geographic range to each tip in the tree. The possible ranges were defined based on major biogeographic breaks described for *Arostrilepis* definitive hosts across the Holarctic (Fedorov et al., 1999; Galbreath & Cook, 2004; Hewitt, 2004; Runck & Cook, 2005). The five possible ranges were: western Palearctic (WP), central Palearctic (CP), eastern Palearctic (EP), northern Nearctic (NN), and southern Nearctic (SN). Each species or lineage was coded as being present or absent in each region. The maximum range size was set equal to three, which is the maximum known number of ranges occupied by an extant species or lineage. I also compared models with and without the inclusion of dispersal multipliers, which alter the probability of specific range shifts, in this instance to inform the model of the connectivity between regions (e.g., an ancestor in the EP would be more likely to disperse into the CP or NN than any of the other regions because it is adjacent to these two). Using the AIC feature of this package, the fit of the data to the DEC*, DEC*+J, DIVAlike*, DIVAlike*+J, BAYAREAlike*, and BAYAREAlike*+J models were compared.

RESULTS

My final dataset included 173 sequences from 95 *Arostrilepis* individuals. Newly generated sequences total 52 cyt-*b*, 51 28S, and 52 ITS2. Eight of the nine morphologically described *Arostrilepis* species for which molecular material is available were successfully sequenced at all three markers. Attempts to amplify *Arostrilepis intermedia* Makarikov & Kontrimavichus, 2011 at 28S consistently failed. In addition to the nine described species, I identified 16 deeply divergent genetic lineages of

Arostrilepis that do not correspond with known species (Figures 4, 5, 6). I assigned these individuals to numbered lineage identifiers that include whether the lineage is Nearctic (Ne) or Palearctic (Pa) and the first three letters of the primary host genus. Most of these lineages were sequenced for at least two markers, except for Lineage 11-Pa-Lem and Lineage 16-Pa-Mic for which only cyt-*b* was successfully amplified.

Independent Gene Genealogies

Of the three single-locus phylogenies, the cyt-*b* phylogeny was most resolved. Each lineage that is represented by more than one individual has strong bootstrap values in support of its monophyly (Figure 4). This support was not present for all lineages in the reconstructions from the other two loci (Figures 5, 6). However, structure across the three loci is generally consistent, and topological disagreements between the loci are associated with relationships that have low bootstrap support.

In general, strong bootstrap support is observed towards the tips of the gene trees with poorer resolution of interior relationships. The lack of resolution at deep nodes weakens assessments of the number and timing of transberingian dispersal events, but the complex distribution of geographic associations across the tips of these gene trees implies a history requiring multiple Beringian crossings.

One notable instance of discontinuity between markers is the placement of Lineage 3-Ne-Lem in the cyt-*b* tree versus the 28S tree. For cyt-*b*, Lineage 3 is shown to be closely related to Lineage 16-Pa-Mic; however, in the 28S tree Lineage 3 is placed deeper in the tree within the cluster of *Arostrilepis beringiensis* (Kontrimavichus & Smirnova, 1991) individuals. In another instance, an individual considered to be part of

Lineage 10 has a long branch length in the 28S phylogeny indicating divergence from the rest of the lineage, but this same individual displays no such structure in the ITS2 topology.

Interspecific Distances

The distinctiveness between lineages is further supported by cyt-*b* mean uncorrected pairwise distance values; *Arostrilepis* cyt-*b* sequences are on average 1.04% divergent within-species and 11.6% divergent between-species (Table 2, Supplemental Table 1). ITS2 pairwise distances also support the distinction of most undescribed lineages (0.0924% average within-species and 1.75% between-species divergence). However, in the case of *Arostrilepis janickii* Makarikov & Kontrimavichus, 2011 and Lineage1-Pa-Mic, structure was not detected at ITS2, despite genetic divergence at cyt-*b* being greater than that between the two most closely related described species, *Arostrilepis cooki* Makarikov, Galbreath & Hoberg, 2013 and *A. macrocirrosa.* (Supplemental Table 2). The pairwise distance values for 28S indicate that this marker is more conserved than the other two, but despite this they do show a difference in within-group distance (on average 0.0515%) versus between-group distance (on average 0.755%) (Supplemental Table 3).

Multi-locus Species Tree Estimation

The coalescent tree produced using *BEAST is largely consistent with the topologies recovered in the ML gene trees. Relationships between the tips of the *Arostrilepis* species phylogeny are generally well-resolved, and relative to the single-locus analyses, internal relationships are better resolved.

The phylogenetic analysis revealed four primary subdivisions. I identify these as the Central Palearctic clade, *Myodes*-associated clade, Western Palearctic clade, and Temperate Nearctic clade (Figure 7). Additional structure may be present, splitting the Temperate Nearctic clade into those with more northern Nearctic distributions and those that are strictly temperate; however current support for this distinction is limited. The three geographically distinguished clades are predominantly associated with *Microtus* voles, though not exclusively. The three species that make up the *Myodes*-associated clade have differing geographic distributions, with *A. intermedia* being restricted to the Palearctic, *A. cooki* being restricted to the Nearctic, and *A. macrocirrosa* being Holarctic.

When host associations are considered across the species tree, or any of the gene trees, it is clear that the taxa associated with particular host genera do not form a monophyletic group. The M*yodes* associated clade is the only strongly supported clade where all members are associated with the same host genus, but this group does not include all of the *Myodes*-associated *Arostrilepis* taxa.

Biogeography

Both raw and corrected AIC scores found the data to best fit the DEC*+J model with dispersal multipliers applied (Table 6 and Table 7). The resulting range estimations of ancestral *Arostrilepis* distributions across the Holarctic derived from this model are provided in Figure 8 and were used for biogeographic interpretation. The biogeographic ancestral range estimation indicates the most probable distribution of the earliest *Arostrilepis* ancestor is within the Central/Eastern Palearctic. The Central Palearctic is the predicted origin for the majority of the basal *Arostrilepis* nodes. Instances in the ancestral estimation where a node or corner of the tree in the Palearctic is followed by a node or

corner in the Nearctic, or vice versa, were considered to be signatures of transberingian dispersal. This estimation identified four or five separate colonization events, all eastward, from Eurasia into North America. Estimated transberingian range shifts for *Arostrilepis* by host association are: two with *Myodes* hosts (II and III Figure 8), one with *Microtus* hosts (IV Figure 8), and possibly two with *Lemmus* hosts (I and ? Figure 8). The questionable *Lemmus* colonization event involves the split between Lineage 3 and Lineage 16, which is unclear due to ambiguity regarding phylogenetic placement of the individual that represents Lineage 3-Ne-Lem. Node height 95% Highest Posterior Density (HPD) ranges of the nodes in the coalescent phylogeny associated with the strongly supported transberingian colonization events indicate that these dispersal events probably occurred during at least two distinct episodes (Figure 7). This is indicated by the deeper *Microtus* associated node height range not overlapping with those of the two *Myodes* associated node heights which occur at more shallow positions in the tree. It is ambiguous whether the *Lemmus* associated colonization occurred during a relatively shallow or a relatively deep episode due to its wide node height 95% HPD.

Of the three next best models (Table 6 and Table 7), DEC*, BAYAREALIKE*+J, and DEC*+J without dispersal multipliers, only the latter differed from greatly from the best model. DEC*+J without dispersal multipliers estimated that rather than the ancestor of the Temperate Nearctic Clade having a widespread (EPNNSN) distribution, the ancestor had a Central Palearctic range and there were two colonization events one by the strictly Southern Nearctic diversity and another by the ancestor of *A. rauschorum* and Lineage 8-Ne-Mic.

Lineage 9-Ne-Tho did not group with any other lineages or species. The Bayesian coalescent analysis placed this lineage as descending from the basal split within the genus. The inclusion of Lineage 9-Ne-Tho in the BioGeoBEARS estimation caused the geographic range estimate for deepest node in the *Arostrilepis* phylogeny to be a combined Central Palearctic and Southern Nearctic distribution, but otherwise did not alter the estimation. Reconstructions using a dataset with lower taxon sampling but a greater number of loci place Lineage 9-Ne-Tho interior rather than basal in the *Arostrilepis* phylogeny, but are otherwise consistent with the mutli-locus phylogeny in this study (Haas, Galbreath, Yuan, and Li, unpublished data). Because of this Lineage 9-Ne-Tho was excluded from the tree for further biogeographic analysis.

DISCUSSION

Diversity within *Arostrilepis*

Divergent genetic lineages identified in this study may represent heretofore undescribed species. The lineages display depths of divergence consistent with interspecific distances between described *Arostrilepis* species (Table 2), though morphological assessment is needed for confirmation of species identities. The continental distributions of described species remain mostly unchanged in light of my new sampling, with six species restricted to the Palearctic, five restricted to the Nearctic, and two with Holarctic distributions. My results indicate that *A. janickii* is restricted to the Palearctic, and that three specimens from Alaska, previously considered to be referable to *A. janickii* (Makarikov et al., 2013) probably represent a distinct species. If morphological assessments confirm the numerous genetically divergent lineages described here as species, Palearctic and Nearctic endemic diversity could increase to 12

species and 14 species respectively. Three of the described Nearctic species, *A. kontrimavichusi* associated with *Myodes californicus, A. mariettavogeae* found primarily in *Peromyscus californicus* (as well as other *Peromyscus* and *Perognathus* species), and *A. schilleri* found in pocket gophers *(Thomomys bulbivorous*), could not be included in this study due to a lack of genetic material. In my dataset, these hosts are associated with Lineages 6-Ne-Myo, 4-Ne-Per, and 7-Ne-Tho and 9-Ne-Tho respectively, suggesting the possibility that the species are represented here, but yet to be confirmed based on morphological criteria. Additional fresh voucher-linked tissues will be necessary to acquire DNA extracts that could be used to confirm whether any of these four lineages represent described *Arostrilepis* species.

Arostrilepis **Systematics**

The *Arostrilepis* phylogeny generated by the coalescent analysis represents the most robust reconstruction of the relationships within this genus to date. This study also includes more *Arostrilepis* specimens from temperate latitudes in North America in its molecular analysis than any previous study. Many of the new lineages in the temperate Nearctic are associated with *Microtus* voles. Prior to my detection of additional *Arostrilepis* diversity associated with *Microtus* hosts, greater *Arostrilepis* diversity was observed in association with *Myodes* hosts (Makarikov & Hoberg, 2016). As more of the geographic gaps in *Arostrilepis* sampling are filled perhaps latitudinal trends in host association will emerge (e.g. more prevalent associations with *Microtus* at temperate latitudes and with *Myodes* at higher latitudes transitioning to *Lemmus* associations at the most northern extent of *Arostrilepis* distribution).

The detection of new *Arostrilepis* diversity generates numerous questions. Many lineages are currently only represented by one individual (e.g., Lineage 9-Ne-Tho), leaving considerable uncertainty regarding geographic distributions and breadth of host associations. In addition to limited sampling for specific lineages, large geographic gaps in sampling of *Arostrilepis* across Eurasia and central/eastern North America suggest that our understanding of *Arostrilepis* diversity is incomplete, warranting further collection efforts and the archiving of specimens.

Biogeographic history of *Arostrilepis*

My data show that *Arostrilepis* likely arose in the Palearctic and that it colonized eastward into the Nearctic at least four times. No evidence for westward *Arostrilepis* colonization was detected. My findings are consistent with previous suggestions that *Arostrilepis* crossed the Bering Land Bridge in association with *Microtus, Myodes* and *Lemmus* hosts (Hoberg et al., 2012; Makarikov et al., 2013). The two distinct episodes of Nearctic colonization by *Arostrilepis* (Figure 7) presumably correlate to two separate openings of the Bering Land Bridge, though the details of timing are difficult to infer in the absence of a robust molecular clock. Mammal-focused studies have revealed that each of the three host taxa which apparently facilitated transberingian *Arostrilepis* colonization moved through Beringia in two temporally-disjunct expansion events (Fedorov et al., 1999; Conroy & Cook, 2000; Fedorov et al., 2003; Galbreath & Cook, 2004; Cook et al., 2004; Kohli et al., 2014a), and it is possible that the distinct episodes of dispersal by *Arostrilepis* accompanied these sequential movements across the land bridge by the rodent hosts.

In the first transberingian expansion by *Arostrilepis* detected in this study, a primarily Nearctic clade of *Arostrilepis* is found to have arisen from a single eastward colonization in association with *Microtus* voles (Figures 7 and 8). Previously it was suggested that *Arostrilepis* colonized the Nearctic twice in association with this host genus (Makarikov et al., 2013). This hypothesis was based on *A. rauschorum* ranging from temperate North American latitudes well into the Arctic, indicating a prolonged presence in the region and contrasting with the detection of specimens in Alaska being morphologically referable to *A. janickii*, a species otherwise found only in Europe. Had these individuals been conspecific with *A. janickii*, they would indicate a disjunct Holarctic distribution acquired separately from the Nearctic *Microtus*-associated assemblage. However, I have shown that those individuals are not conspecific with *A. janickii*, but instead belong to Lineage 8-Ne-Mic, which is nested within Nearctic *Arostrilepis* diversity. These findings suggest that Nearctic *Microtus*-associated *Arostrilepis* diversity arose from a single ancestral colonization event from Eurasia into North America.

The placement of Lineage 15-Pa-Lem within the temperate Nearctic clade is likely the result of multiple splitting events in the history of this clade. The common ancestor of this clade is estimated to have had a wide distribution spanning the eastern Palearctic, northern Nearctic, and southern Nearctic. Phylogenetic structure within this clade suggests that after achieving this distribution, the ancestral population experienced two splitting events that subdivided it geographically and initiated divergence of regionally isolated populations (Figure 8). First the southern Nearctic population split from the combined eastern Palearctic and northern Nearctic population. Then, the

Palearctic population diverged from the Nearctic, distinguishing Lineage 15-Pa-Lem from the northern Nearctic taxa, *A. rauschorum* and Lineage 8-Ne-Mic.

Most host associations within this temperate Nearctic clade are with *Microtus*, which suggests a single ancestral colonization of this host genus that was retained through multiple parasite speciation events. Host associations in this clade that fall outside of the genus *Microtus,* involving *Lemmus*, *Myodes, Peromyscus,* and *Thomomys*, are therefore parsimoniously inferred to be the result of host-switching events.

The diversity of this Nearctic *Microtus*-associated clade may reflect a history in which the parasites underwent a post-colonization radiation similar to that of their hosts. Excluding the Holarctic *Microtus oeconomus*, Nearctic endemic *Microtus* form a monophyletic group that may have initially colonized from Eurasia about 1.5-2.1 MYA (Repenning, 1990; Conroy & Cook, 2000). Across their range, *Microtus* radiated rapidly, making their systematic relationships challenging to discern (Conroy & Cook, 2000; Jaarola et al., 2004). One hypothesis explaining the rapid radiation of *Microtus* diversity is an abrupt isolation of populations within separate refugia (Conroy & Cook, 2000). *Arostrilepis* radiation in this clade may reflect simultaneous isolation with these fragmented host populations, though evidence for host-switching (Figure 7. e.g., *Myodes, Peromyscus,* and *Thomomys*) suggests that refugial isolation alone is unlikely to be the sole driver of diversification (Hoberg & Brooks, 2008).

Biogeographic range estimation results indicate independent, sequential transberingian dispersal events for the ancestor of *A. cooki* and for Holarctic *A. macrocirrosa*. These findings are consistent with the hypothesis that these parasites dispersed eastward into North America with the ancestor of endemic Nearctic *Myodes*

species and with the last glacial maximum (LGM) colonizer *M. rutilus*, respectively (Makarikov et al., 2013). The two host colonization events are considered to be sequential, occurring during an opening of the Bering Land Bridge in the early Pleistocene and a second opening in the late Pleistocene (Cook et al., 2004; Runck & Cook, 2005; Kohli et al., 2014a). While narrow overlap of node height ranges associated with the origins of *A. cooki* and Nearctic *A. macrocirrosa* does not reject the possibility of simultaneous dispersal, the history of the primary hosts, geographic distributions of the hosts and parasites, and patterns of intraspecific diversity within the parasites all are consistent with two separate sequential colonization events. Population structure within *M. gapperi* indicates that the ancestor of the species persisted in multiple Nearctic refugia south of the continental ice sheets during the LGM (Runck & Cook, 2005). *Myodes rutilus* arose in Eurasia, eventually forming three primary subclades spanning western Asia, central Asia, and Beringia (Kohli et al., 2014a). Phylogeographic evidence indicates that the biogeographic break between the central and eastern clades of *M. rutilus,* originally located at the Kolyma River, shifted around the time of the LGM, resulting in their present division being at the Bering Strait. The eastern clade of *M. rutilus,* apart from a few isolated populations in Kamchatka, is now restricted to the Nearctic (Kohli et al., 2014a). Following the recession of the glaciers in the Nearctic, *M. gapperi* expanded northward where it came into contact with the LGM colonizer *M. rutilus.* These two *Myodes* species form a narrow contact zone running east to west from southeast Alaska through northern Canada (Runck & Cook, 2005). *Arostrilepis cooki*, like its host *M. gapperi*, has a range that extends from temperate latitudes northward to where it comes into contact with *A. macrocirrosa* in *M. rutilus*, but the degree to which the geographic

distributions of the two parasite species overlap is yet to be determined. Well-supported intraspecific structure within *A. cooki* contrasts with a general lack of structure in Nearctic *A. macrocirrosa* (S. Gallagher and K. Galbreath, unpublished data), providing further evidence that *A. cooki* has a deeper history in the continent (Figures 4 and 5). The Eurasian origin and then later eastward expansion into North America by *A. macrocirrosa* is supported by the finding of higher haplotype diversity in Eurasia (S. Gallagher and K. Galbreath, unpublished data).

The broad range of *A. beringiensis* implies that the species would have had to cross the Bering Land Bridge at least once in its history. My biogeographic range estimation and phylogenetic reconstructions are consistent with that implication, and suggest that the species arose in the Palearctic and subsequently colonized North America. While the resolution of the estimation does not preclude the possibility of *A. beringiensis* being an endemic Beringian species which radiated both eastward and westward into the northern continents, investigations into the history of *Lemmus* biogeography indicate that Beringia was not an important source of postglacial colonization for *Lemmus* populations in Asia and North America (Fedorov et al., 2003). The range estimation shows that at some point after its relatively early split from Lineage 11-Pa-Lem, the range of *A. beringiensis* expanded from a narrow eastern Palearctic distribution to also encompass the central Palearctic and the northern Nearctic. The ancestors of *Lemmus sibiricus* and *Lemmus trimucronatus* are hypothesized to have split from each other prior to the penultimate glacial maximum due to separation caused by the periodic inundation of Beringia (Fedorov et al., 1999). Fossil evidence and population structure indicate that the progenitor of *L. trimucronatus* occupied southern Nearctic

periglacial refugia before moving back north during a prior interglacial period and eventually expanding its range westward into Eurasian Beringia during the LGM (Fedorov et al., 2003). In Eurasia, the range of *L. trimucronatus* followed the receding distribution of *L. sibiricus* to their current population break at the Kolyma River (Fedorov et al., 1999, 2003). Because *A. beringiensis* has an extensive range in Eurasia, it is unlikely that it arose in the Nearctic with *L. trimucronatus*, which would require it to have penetrated deeply into the range of *L. sibiricus* after the two hosts came into contact. For comparison, another *Lemmus* parasite, *Anoplocephaloides lemmi*, which is known to have shifted westward with *L. trimucronatus*, does not appear to have dispersed as far beyond the break between *L. sibiricus* and *L. trimucronatus* as is seen in *A. beringiensis* (Haukisalmi et al., 2016). Perturbation and range-expansion events brought on by climate change can result in parasites being brought into contact with susceptible, naïve hosts resulting in host-switching events (Hoberg & Brooks, 2010). Though *L. sibiricus* and *L. trimucronatus* are currently allopatric, I propose that historically they were in contact within Beringia, allowing *A. beringiensis* to colonize *L. trimucronatus* and expand eastward through this host into North America eventually expanding into *Synaptomys* hosts as well.

The cyt-*b* placement of Lineage 3-Ne-Lem suggests either a second Nearctic *Arostrilepis* colonization in association with lemmings or, considering the lineage is nested within *Microtus* associated diversity, a colonization with *Microtus* and subsequent host-switch to lemmings. However, the 28S phylogeny suggests there was only one *Lemmus* associated *Arostrilepis* colonization because Lineage 3-Ne-Lem is placed within *A. beringiensis.* A possible explanation for the phylogenetic incongruence between the

cyt-*b* (Figure 4) and 28S (Figure 6) ML placements of Lineage 3-Ne-Lem is a history of ancient hybridization, which resulted in capture of Lineage 16-Pa-Mic mitochondrial DNA by *A. beringiensis*. Such a history has not been demonstrated previously for cestodes to my knowledge, but it has been widely documented in other taxa (Good et al., 2008, 2015; Bronstein et al., 2016; Shipham et al., 2017). Evidence of hybridization has also been found in the rodent hosts of *Arostrilepis.* For example, ancient hybridization between *M. rutilus* and *M. gapperi* has been shown have occurred in populations from southeast Alaska (Runck et al., 2009). Hybridization would support the hypothesis of a single *Lemmus* associated *Arostrilepis* colonization.

The possibility of a recent *A. beringiensis* colonization event suggests the potential for simultaneous colonization with *A. macrocirrosa*. This colonization event was probably independent from the colonization by the ancestor of *A. cooki* with the common ancestor of Nearctic *Myodes* species. In my coalescent phylogeny, the error bars for the node heights concerning *Myodes*-associated colonization do not overlap with that of the *Microtus*-associated colonization, suggesting that the *Microtus*-associated event may represent the first of three sequential colonization events (Figure 7). Evidence provided by host histories for the distinct nature of the *Microtus* associated colonization from the earlier *Myodes* associated colonization is weakly in agreement with that of the *Arostrilepis* phylogeny. The earliest fossil evidence places *Microtus* in North America about 1.5-2.1 MYA (Repenning, 1990, 2001). *Myodes* fossil evidence is less conclusive (Cook et al., 2004). Some estimates put *Myodes* in North America around 0.85 MYA (Repenning, 2001), and the results of a recent assessment of the tribe Myodini (Kohli et al., 2014b) suggest that the Nearctic common ancestor of *Myodes* coalesces with

Palearctic *Myodes* about 1.5 MYA (± 0.5 MYA). All *Arostrilepis* colonization events are inferred to be eastward, even in the unresolved case of Lineage 3-Ne-Lem. Whether Lineage 3 is an independent lineage, still conspecific with Lineage 16-Pa-Mic, or a hybrid of *A. beringiensis* and Lineage 16, eastward movement is indicated. The eastward colonization of *A. beringiensis* is consistent with the majority of transberingian colonizations, but is counter to the flow of its host movement. This highlights the fact that some parasites have their own histories and do not necessarily track their respective host histories with perfect fidelity (Hoberg & Brooks, 2015). Differences like this have the potential to elucidate otherwise cryptic nuances of biogeographic movement, contributing to the greater understanding of transberingian host-parasite dynamics.

REFERENCES

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., & Drummond, A.J. (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology*, **10**, e1003537.

Bronstein, O., Kroh, A., & Haring, E. (2016) Do genes lie? Mitochondrial capture masks the Red Sea collector urchin's true identity (Echinodermata: Echinoidea: *Tripneustes*). *Molecular Phylogenetics and Evolution*, **104**, 1–13.

Brooks, D.R., Hoberg, E.P., & Boeger, W.A. (2015) In the Eye of the Cyclops: The Classic Case of Cospeciation and Why Paradigms are Important. *Comparative Parasitology*, **82**, 1–8.

Brooks, D.R. & McLennan, D.A. (2002) *The Nature of Diversity: an evolutionary voyage of discovery*. University of Chicago Press, Chicago.

Brunhoff, C., Galbreath, K.E., Fedorov, V.B., Cook, J.A., & Jaarola, M. (2003) Holarctic phylogeography of the root vole (*Microtus oeconomus*): implications for late Quaternary biogeography of high latitudes. *Molecular Ecology*, **12**, 957–968.

Chaline, J. & Graf, J.-D. (1988) Phylogeny of the Arvicolidae (Rodentia): Biochemical and Paleontological Evidence. *Journal of Mammalogy*, **69**, 22.

Conroy, C.J. & Cook, J.A. (1999) MtDNA Evidence for Repeated Pulses of Speciation Within Arvicoline and Murid Rodents. *Journal of Mammalian Evolution*, **6**, 221–245.

Conroy, C.J. & Cook, J.A. (2000) Molecular systematics of a Holarctic rodent (*Microtus*: Muridae). *Journal of Mammalogy*, **81**, 344–359.

Cook, J., Galbreath, K., Bell, K., et al. (In Press) The Beringian Coevolution Project: Holistic Collections of Mammals and Associated Parasites Reveal Novel Perspectives on Evolutionary and Environmental Change in the North. *Arctic Science*, .

Cook, J.A., Hoberg, E.P., Koehler, A., Henttonen, H., Wickström, L.M., Haukisalmi, V., Galbreath, K.E., Chernyavski, F., Dokuchaev, N., Lahzuhtkin, A., MacDonald, S.O., Hope, Andrew G., Waltari, E., Runck, A., Veitch, A., Popko, R., Jenkins, E., Kutz, S.J., & Eckerlin, R. (2005) Beringia: intercontinental exchange and diversification of high latitude mammals and their parasites during the Pliocene and Quaternary. *Mammal Study*, **30**, S33–S44.

Cook, J.A., Runck, A.M., & Conroy, C.J. (2004) Historical biogeography at the crossroads of the northern continents: molecular phylogenetics of red-backed voles (Rodentia: Arvicolinae). *Molecular Phylogenetics and Evolution*, **30**, 767–777.

Darriba, D., Taboada, G.L., Doallo, R., & Posada, D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**, 1792–1797.

Fedorov, V., Goropashnaya, A., Jarrell, G.H., & Fredga, K. (1999) Phylogeographic structure and mitochondrial DNA variation in true lemmings (*Lemmus*) from the Eurasian Arctic. *Biological Journal of the Linnean Society*, **66**, 357–371.

Fedorov, V.B. (1999) Contrasting mitochondrial DNA diversity estimates in two sympatric genera of Arctic lemmings (*Dicrostonyx*: *Lemmus*) indicate different responses to Quaternary environmental fluctuations. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **266**, 621–626.

Fedorov, V.B. & Goropashnaya, A.V. (1999) The importance of ice ages in diversification of Arctic collared lemmings (*Dicrostonyx*): evidence from the mitochondrial cytochrome b region. *Hereditas*, **130**, 301–307.

Fedorov, V.B., Goropashnaya, A.V., Jaarola, M., & Cook, J.A. (2003) Phylogeography of lemmings (*Lemmus*): no evidence for postglacial colonization of Arctic from the Beringian refugium. *Molecular Ecology*, **12**, 725–731.

Felsenstein, J. (1988) Phylogenies from molecular sequences: inference and reliability. *Annual Review of Genetics*, **22**, 521–565.

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

Galbreath, K.E., Cook, J.A., Eddingsaas, A.A., & DeChaine, E.G. (2011) Diversity and demography in Beringia: multilocus tests of paleodistribution models reveal the complex history of arctic ground squirrels. *Evolution*, **65**, 1879–1896.

Galbreath, K.E. & Hoberg, E.P. (2012) Return to Beringia: parasites reveal cryptic biogeographic history of North American pikas. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 371–378.

Galbreath, K.E., Ragaliauskaite, K., Kontrimavichus, L., Makarikov, A.A., & Hoberg, E.P. (2013) A widespread distribution for *Arostrilepis tenuicirrosa* (Eucestoda: Hymenolepididae) in *Myodes* voles (Cricetidae: Arvicolinae) from the Palearctic based on molecular and morphological evidence: historical and biogeographic implications. *Acta Parasitologica*, **58**, 441–452.

Gatesy, J., DeSalle, R., & Wahlberg, N. (2007) How Many Genes Should a Systematist Sample? Conflicting Insights from a Phylogenomic Matrix Characterized by Replicated Incongruence. *Systematic Biology*, **56**, 355–363.

Gladenkov, A.Y. & Gladenkov, Y.B. (2004) Onset of connections between the Pacific and Arctic Oceans through the Bering Strait in the Neogene. *Stratigr. Geol. Correl*, **12**, 175–178.

Good, J.M., Hird, S., Reid, N., Demboski, J.R., Steppan, S.J., Martin-Nims, T.R., & Sullivan, J. (2008) Ancient hybridization and mitochondrial capture between two species of chipmunks. *Molecular Ecology*, **17**, 1313–1327.

Good, J.M., Vanderpool, D., Keeble, S., & Bi, K. (2015) Negligible nuclear introgression despite complete mitochondrial capture between two species of chipmunks: INTROGRESSION IN CHIPMUNKS. *Evolution*, **69**, 1961–1972.

Graham, S.W., Olmstead, R.G., & Barrett, S.C.H. (2002) Rooting Phylogenetic Trees with Distant Outgroups: A Case Study from the Commelinoid Monocots. *Molecular Biology and Evolution*, **19**, 1769–1781.

Guindon, S. & Gascuel, O. (2003) A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology*, **52**, 696–704.

Guthrie, R.D. (1968) Paleoecology of the large mammal community in interior Alaska during the late Pleistocene. *American Midland Naturalist*, **79**, 346–363.

Guthrie, R.D. (1984) Mosaics, Allelochemics, and Nutrients: An Ecological Theory of Late Pleistocene Megafaunal Extinctions. *Quaternary Extinctions A Prehistoric Revolution* pp. 259–298. The University of Arizona Press, Tucson, Arizona.

Hafner, M.S. & Nadler, S.A. (1988) Phylogenetic trees support the coevolution of parasites and their hosts. *Nature*, **332**, 258–259.

Hafner, M.S. & Nadler, S.A. (1990) Cospeciation in Host-Parasite Assemblages: Comparative Analysis of Rates of Evolution and Timing of Cospeciation Events. *Systematic Zoology*, **39**, 192.

Haukisalmi, V., Hardman, L.M., Fedorov, V.B., Hoberg, E.P., & Henttonen, H. (2016) Molecular systematics and Holarctic phylogeography of cestodes of the genus *Anoplocephaloides* Baer, 1923 s. s. (Cyclophyllidea, Anoplocephalidae) in lemmings (*Lemmus*, *Synaptomys*). *Zoologica Scripta*, **45**, 88–102.

Haukisalmi, V., Hardman, L.M., Foronda, P., Feliu, C., Laakkonen, J., Niemimaa, J., Lehtonen, J.T., & Henttonen, H. (2010) Systematic relationships of hymenolepidid cestodes of rodents and shrews inferred from sequences of 28S ribosomal RNA: Systematic relationships of hymenolepidid cestodes. *Zoologica Scripta*, **39**, 631–641.

Heled, J. & Drummond, A.J. (2010) Bayesian Inference of Species Trees from Multilocus Data. *Molecular Biology and Evolution*, **27**, 570–580.

Hewitt, G.M. (2004) The structure of biodiversity–insights from molecular phylogeography. *Frontiers in Zoology*, **1**, 1–16.

Hoberg, E.P. & Brooks, D.R. (2008) A macroevolutionary mosaic: episodic host-switching, geographical colonization and diversification in complex host-parasite systems. *Journal of Biogeography*, **35**, 1533–1550.

Hoberg, E.P. & Brooks, D.R. (2010) Beyond vicariance: integrating taxon pulses, ecological fitting, and oscillation in evolution and historical biogeography. *The biogeography of host-parasite interactions* pp. 7–20. Oxford University Press.

Hoberg, E.P. & Brooks, D.R. (2015) Evolution in action: climate change, biodiversity dynamics and emerging infectious disease. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **370**, 20130553–20130553.

Hoberg, E.P., Galbreath, K.E., Cook, J.A., Kutz, S.J., & Polley, L. (2012) Northern Host-Parasite Assemblages: History and Biogeography on the Borderlands of Episodic Climate and Environmental Transition. *Advances in Parasitology*, **79**, 1–97.

Hoffmann, R.S. (1981) Different Voles for Different Holes: Environmental Restrictions on Refugial Survival of Mammals. *Evolution Today: proceedings of the Second International Congress of Systematic and Evolutionary Biology (ed. by G.G.E. Scudder and J.L. Reveal)* pp. 22–45. Carnegie Mellon University, Pittsburgh, PA.

Hope, A.G., Takebayashi, N., Galbreath, K.E., Talbot, S.L., & Cook, J.A. (2013) Temporal, spatial and ecological dynamics of speciation among amphi-Beringian small mammals. *Journal of Biogeography*, **40**, 415–429.

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

Jaarola, M., Martínková, N., Gündüz, İ., Brunhoff, C., Zima, J., Nadachowski, A., Amori, G., Bulatova, N.S., Chondropoulos, B., Fraguedakis-Tsolis, S., González-Esteban, J., José López-Fuster, M., Kandaurov, A.S., Kefelioğlu, H., da Luz Mathias, M., Villate, I., & Searle, J.B. (2004) Molecular phylogeny of the speciose vole genus *Microtus* (Arvicolinae, Rodentia) inferred from mitochondrial DNA sequences. *Molecular Phylogenetics and Evolution*, **33**, 647–663.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., & Drummond, A. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, **28**.

Kohli, B.A., Fedorov, V.B., Waltari, E., & Cook, J.A. (2014a) Phylogeography of a Holarctic rodent (*Myodes rutilus*): testing high-latitude biogeographical hypotheses and the dynamics of range shifts. *Journal of Biogeography*, **42**, 377–389.

Kohli, B.A., Speer, K.A., Kilpatrick, C.W., Batsaikhan, N., Damdinbaza, D., & Cook, J.A. (2014b) Multilocus systematics and non-punctuated evolution of Holarctic Myodini (Rodentia: Arvicolinae). *Molecular Phylogenetics and Evolution*, **76**, 18–29.

Kumar, S., Stecher, G., & Tamura, K. (2015) MEGA7: Molecular Evolutionary Genetics Analysis version 7.0. *Molecular Biology and Evolution*.

Landis, M.J., Matzke, N.J., Moore, B.R., & Huelsenbeck, J.P. (2013) Bayesian Analysis of Biogeography when the Number of Areas is Large. *Systematic Biology*, **62**, 789–804.

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

Littlewood, D. (2000) The Interrelationships of Proseriata (Platyhelminthes: Seriata) Tested with Molecules and Morphology. *Molecular Phylogenetics and Evolution*, **16**, 449–466.

Lockyer, A.E., Olson, P.D., & Littlewood, D.T.J. (2003) Utility of complete large and small subunit rRNA genes in resolving the phylogeny of the Neodermata (Platyhelminthes): implications and a review of the cercomer theory. *Biological Journal of the Linnean Society*, **78**, 155–171.

Makarikov, A.A., Galbreath, K.E., & Hoberg, E.P. (2013) Parasite diversity at the Holarctic nexus: species of *Arostrilepis* (Eucestoda: Hymenolepididae) in voles and lemmings (Cricetidae: Arvicolinae) from greater Beringia. *Zootaxa*, **3608**, 401–439.

Makarikov, A.A., Gardner, S.L., & Hoberg, E.P. (2012) New Species of *Arostrilepis* (Eucestoda: Hymenolepididae) In Members of Cricetidae and Geomyidae (Rodentia) from the Western Nearctic. *Journal of Parasitology*, **98**, 617–626.

Makarikov, A.A., Gulyaev, V.D., & Kontrimavichus, V.L. (2011) A redescription of *Arostrilepis horrida* (Linstow, 1901) and descriptions of two new species from Palaearctic microtine rodents, *Arostrilepis macrocirrosa* sp. n. and *A. tenuicirrosa* sp. n. (Cyclophyllidea: Hymenolepididae). *Folia Parasitologica*, **58**, 108–120.

Makarikov, A.A. & Hoberg, E.P. (2016) Broadening diversity in the *Arostrilepis horrida* complex: *Arostrilepis kontrimavichusi* n. sp. (Cyclophyllidea: Hymenolepididae) in the western red-backed vole *Myodes californicus* (Merriam) (Cricetidae: Arvicolinae) from temperate latitudes of the Pacific Northwest, North America. *Systematic Parasitology*, **93**, 467–477.

Makarikov, A.A. & Kontrimavichus, V.L. (2011) A redescription of *Arostrilepis beringiensis* (Kontrimavichus et Smirnova, 1991) and descriptions of two new

species from Palaearctic microtine rodents, *Arostrilepis intermedia* sp. n. and *A. janickii* sp. n. (Cestoda: Hymenolepididae). *Folia Parasitologica*, **58**, 289–301.

Marincovich, L.J. & Gladenkov, A.Y. (1999) Evidence for an early opening of theBeringStrait. *Nature*, **397**, 146–149.

Mas-Coma, S. & Tenora, F. (1997) Proposal of *Arostrilepis* n. gen. (Cestoda: Hymenolepidae). *Research and Reviews in Parasitology*, .

Massana, K.A., Beaulieu, J.M., Matzke, N.J., & O'Meara, B.C. (2015) Non-null Effects of the Null Range in Biogeographic Models: Exploring Parameter Estimation in the DEC Model. .

Matzke, N.J. (2013a) *BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis in R Scripts.*

Matzke, N.J. (2013b) Probabilistic historical biogeography: new models for founder-event specieation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography*, **5**, 242–248.

Meiri, M., Lister, A.M., Collins, M.J., Tuross, N., Goebel, T., Blockley, S., Zazula, G.D., van Doorn, N., Dale Guthrie, R., Boeskorov, G.G., Baryshnikov, G.F., Sher, A., & Barnes, I. (2014) Faunal record identifies Bering isthmus conditions as constraint to end-Pleistocene migration to the New World. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20132167–20132167.

Okamoto, M., Agatsuma, T., Kurosawa, T., & Ito, A. (1997) Phylogenetic relationships of three hymenolepidid species inferred from nuclear ribosomal and mitochondrial DNA sequences. *Parasitology*, **115**, 661–666.

Paterson, A.M. & Banks, J. (2001) Analytical approaches to measuring cospeciation of host and parasites: through a glass, darkly. *International Journal for Parasitology*, **31**, 1012–1022.

Rambaut, A., Suchard, M.A., Xie, D., & Drummond, A.J. (2014) *Tracer v1.6.*

Ree, R.H. & Smith, S.A. (2008) Maximum Likelihood Inference of Geographic Range Evolution by Dispersal, Local Extinction, and Cladogenesis. *Systematic Biology*, **57**, 4–14.

Repenning, C.A. (1990) Arvicolid rodent biochronology of the Northern Hemisphere. *International symposium on evolution, phylogeny, and biostratigraphy of arvicolids (Rodentia, Mammalia)* (ed. by O. Fejfar and Heinrich), pp. 385–418. Pfiel-Verlag, Munich.

Repenning, C.A. (2001) Beringian climate during intercontinental dispersal: a mouse eye view. *Quaternary Science Reviews*, **20**, 25–40.

Ronquist, F. (1997) Dispersal-Vicariance Analysis: A New Approach to the Quantification of Historical Biogeography. *Systematic Biology*, **46**, 195.

Runck, A.M. & Cook, J.A. (2005) Postglacial expansion of the southern red-backed vole (Clethrionomys gapperi) in North America. *Molecular Ecology*, **14**, 1445–1456.

Runck, A.M., Matocq, M.D., & Cook, J.A. (2009) Historic hybridization and persistence of a novel mito-nuclear combination in red-backed voles (genus *Myodes*). *BMC Evolutionary Biology*, **9**, 114.

Sher, A. (1999) Traffic lights at the Beringian crossroads. *Nature*, **397**, 103–104.

Shipham, A., Schmidt, D.J., Joseph, L., & Hughes, J.M. (2017) A genomic approach reinforces a hypothesis of mitochondrial capture in eastern Australian rosellas. *The Auk*, **134**, 181–192.

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

Stewart, J.R., Lister, A.M., Barnes, I., & Dalen, L. (2010) Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 661–671.

Swofford, D.L. (2003) *Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer Associates, Sunderland, Massachusetts.

Waltari, E., Hoberg, E.P., Lessa, E.P., & Cook, J.A. (2007) Eastward ho: phylogeographical perspectives on colonization of hosts and parasites across the Beringian nexus. *Journal of Biogeography*, **34**, 561–574.

Wheeler, W.C. (1990) Nucleic acid sequence phylogeny and random outgroups. *Cladistics*, **6**, 363–367.

Yu, Y., Harris, & He, X.J. (2013) *RASP (Reconstruct Ancestral State in Phylogenies)*.

Table 1. *Arostrilepis* specimen information. In the columns for cyt-*b*, 28S, and ITS2, x denotes that a sequence from that individual was generated or obtained from Genbank for the respective marker.

Species/Lineage number	ID number	Host species	Regional locality	Cyt-*b*	28S	ITS2	Genbank #
A. beringiensis	52267	*Synaptomys borealis*	Yukon-Charley, AK	x	x	x	
A. beringiensis	Lem117	*Lemmus sibericus*	Taymyr, Russia	x	x	x	JX392048 (cytb)
A. beringiensis	49480	*Synaptomys borealis*	Yukon-Charley, AK	x			JX392046 (cytb)
A. beringiensis	Z55	*Lemmus sibericus*	Wrangel Island, Russia	x	x	x	GU166223 (28S)
A. beringiensis	211	*Lemmus sibericus*	Anabar River, Russia	x			
A. beringiensis	Lem2	*Lemmus sibericus*	Indigirka, Russia		x	x	
A. beringiensis	Lem786	*Lemmus trimuchronatus*	Nunavut, Canada			x	
A. cooki	IF 6750	*Myodes gapperi*	Meziadin, BC, Canada	x	x	x	JX392032 (cytb)
A. cooki	NK 231346 C1B	*Myodes*	Hooke Lake, Canada	x	x		
A. cooki	AZ0	*Peromyscus maniculatus*	Oregon	x	x	x	
A. cooki	BA1	*Myodes gapperi*	Oregon	x	x	x	
A. cooki	NK231329 c3	*Myodes*	Fort Smith, YK, Canada		x		
A. cooki	IF6827c2	*Myodes gapperi*	Bell2, BC Canada			x	
A. cooki	IF6830c6	*Myodes gapperi*	Bell2, BC Canada			x	
A. gulyaevi	IF 5657	*Myodes rufocanus*	Buynda River, Magadan, Russia	x	x	x	JX392030 (cytb)
A. gulyaevi	F24	*Myodes rufocanus*	China	x	x		
A. gulyaevi	AI7	*Myodes rufocanus*	China		x		
A. gulyaevi	T33	*Myodes rutilus*	Magadan, Russia		x		
A. gulyaevi	U34	*Myodes rufocanus*	Buryatia, Russia		x		
A. gulyaevi	187	*Myodes rufocanus*	Severo Evensk, Russia	x			
A. gulyaevi	IF5094	*Myodes rufocanus*	Anadyr, Russia			x	
A. gulyaevi	IF5079c1	*Myodes rufocanus*	Anadyr, Russia			x	
A. intermedia	IF 5519	*Myodes rutilus*	Buynda	x		x	JX392042 (cytb)
A. intermedia	38071 C1	*Myodes rufocanus*	Omolon, Russia	x		x	JX392041 (cytb)
A. intermedia	124	*Myodes rufocanus*	Bolshoy Shantar Island, Russia	x			
A. intermedia	189	*Myodes rufocanus*	Zavyalova Island, Russia	x			
A. janickii	MHNG933v	*Arvicola amphibius*	Le Lieu, Switzerland	x	x	x	JX392049 (cytb)
A. janickii	MHNG931vd	*Arvicola amphibius*	Le Lieu, Switzerland	x	x	x	JX392050 (cytb)

Species/Lineage number	ID number	Host species	Regional locality	Cyt-*b*	28S	ITS2	Genbank #
A. macrocirrosa	AF 55136	*Myodes rutilus*	Wrangel St. Elias, AK	x	x	x	
A. macrocirrosa	49374	*Myodes rutilus*	Yukon Charley, AK	x			JX392039 (cytb)
A. macrocirrosa	AF 48447 c1	*Microtus oeconomus*	Noatuk-Kobuk, AK	x	x		
A. macrocirrosa	U40	*Myodes rutilus*	Buryatia, Verhnaya Berezovka	x	x		GU166224 (28S)
A. macrocirrosa	V6	*Myodes rutilus*	Tunguska Irkutsk, Russia	x	x		
A. macrocirrosa	AF 38004 c1	*Myodes rutilus*	Omolon, Russia		x	x	
A. macrocirrosa	52554	*Myodes rutilus*	Yukon Charley, AK			x	
A. macrocirrosa	38880c1	*Myodes rutilus*	Upper Kolyma			x	
A. macrocirrosa	37107c1	*Myodes rutilus*	Noatak-Kobuk, AK			x	
A. macrocirrosa	36651c2	*Myodes rutilus*	Seward Peninsula, AK			x	
A. microtis	38351--Hym118	*Microtus oeconomus*	Omolon, Russia	x	x	x	
A. microtis	H115-1	*Microtus oeconomus*	Irkutsk, Russia	x			JX392045 (cytb)
A. microtis	38148--Hym108	*Microtus oeconomus*	Omolon, Russia	x	x	x	
A. microtis	AF 38376 c1	*Microtus oeconomus*	Omolon, Russia		x	x	JX392043 (cytb)
A. microtis	38356c1	*Microtus oeconomus*	Omolon, Russia			x	
A. rauschorum	49499--Hym118	*Microtus longicaudus*	Noatak, AK	x	x	x	JX392037 (cytb)
A. rauschorum	48555--Hym108	*Microtus miurus*	Yukon Charley, AK	x	x	x	
A. rauschorum	C96	*Microtus oeconomus*	Amundsen Gulf, Canada	x			
A. rauschorum	G42	*Microtus pennsylvanicus*	Fairbanks, AK		x		GU166226 (28S)
A. rauschorum	G18	*Microtus miurus*	Toolik, AK		x		
A. rauschorum	G1	*Microtus oeconomus*	Toolik, AK		x		
A. rauschorum	AF 37462	*Microtus pennsylvanicus*	Bonanza Creek, AK		x	x	
A. rauschorum	46399c4	*Microtus oeconomus*	Seward Peninsula, AK			x	
A. tenuicirrosa	AF 38038 C2	*Myodes rutilus*	Omolon, Russia	x	x	x	JX104768 (ITS2); JX104762 (cytb)
A. tenuicirrosa	W11	*Myodes rufocanus*	Buryatia, Russia	x	x		
A. tenuicirrosa	197	*Myodes rutilus*	Altai Mountains, Russia	x			
A. tenuicirrosa	198	*Myodes rutilus*	Altai Mountains, Russia	x			
A. tenuicirrosa	38814c2	*Myodes rutilus*	Magadan, Russia			x	JX104773 (ITS2); JX104767 (cytb)
A. tenuicirrosa	W12	*Myodes rutilus*	Buryatia, Russia			x	

Species/Lineage number	ID number	Host species	Regional locality	Cyt-*b*	28S	ITS2	Genbank #
Lineage 1-Pa-Mic	E43	*Microtus arvalis*	Italy	x	x	x	
Lineage 1-Pa-Mic	E47	*Microtus arvalis*	Italy	x	x	x	
Lineage 2-Pa-Mic	U87	*Microtus subterraneus*	Croatia	x	x	x	
Lineage 3-Ne-Lem	AL5	*Lemmus trimuchronatus*	Hope Bay, Nunavut	x	x		
Lineage 4-Ne-Per	AY9	*Peromyscus maniculatus*	Oregon	x	x	x	GU166225 (28S)
Linege 5-Ne-Mic	H49	*Microtus oregoni*	Corvallis, Oregon	x	x		
Linege 5-Ne-Mic	H50	*Microtus oregoni*	Oregon	x		x	
Lineage 6-Ne-Myo	H55	*Myodes californicus*	Oregon	x	x	x	
Lineage 7-Ne-Tho	BA0	*Thomomys talpoides*	Oregon	x	x	x	
Lineage 8-Ne-Mic	AF 36025 c2	*Microtus oeconomus*	Seward Peninsula, AK	x	x	x	
Lineage 8-Ne-Mic	AF 36738 c2	*Microtus oeconomus*	Seward Peninsula, AK	x	x	x	
Lineage 9-Ne-Tho	AX8	*Thomomys idahoensis*	Montana	x	x	x	
Lineage 10-Pa-Mic	W23	*Microtus fortis*	Buryatia, Russia	x	x		
Lineage 10-Pa-Mic	W20	*Microtus fortis*	Buryatia, Russia	x			
Lineage 10-Pa-Mic	U32	*Microtus fortis*	Buryatia, Russia	x	x	x	
Lineage 10-Pa-Mic	CD2	*Microtus oeconomus*	Nenetskiy, Russia		x		
Lineage 10-Pa-Mic	W18	*Microtus oeconomus*	Buryatia, Russia		x		
Lineage 10-Pa-Mic	U33	*Microtus oeconomus*	Buryatia, Russia		x	x	
Lineage 10-Pa-Mic	W21	*Cricetulus barabensis*	Buryatia, Russia		x	x	
Lineage 11-Pa-Lem	170	*Lemmus sibericus*	Baikalo Lenskiy, Russia	x			
Lineage 11-Pa-Lem	171	*Lemmus sibericus*	Baikalo Lenskiy, Russia	x			
Lineage 12-Ne-Mic	AY7	*Microtus richardsoni*	Wyoming	x	x	x	
Lineage 12-Ne-Mic	AY8	*Microtus richardsoni*	Wyoming	x			
Lineage 13-Pa-Mic/Myo	BE0	*Microtus middendorfi*	Yamal, Russia	x	x	x	
Lineage 13-Pa-Mic/Myo	194	*Myodes rutilus*	Altai Mountains, Russia	x			
Lineage 13-Pa-Mic/Myo	196	*Myodes rutilus*	Altai Mountains, Russia		x		
Lineage 13-Pa-Mic/Myo	DL9	*Microtus gregalis*	Yamal, Russia	x			
Lineage 13-Pa-Mic/Myo	DM0	*Microtus gregalis*	Yamal, Russia	x			
Lineage 13-Pa-Mic/Myo	DM1	*Microtus middendorfi*	Yamal, Russia	x			

Species/Lineage number	ID number	Host species	Regional locality	Cyt-*b*	28S	ITS2	Genbank #
Lineage 14-Ne-Mic	BB7	*Microtus longicaudus*	Montana	x	x	x	
Lineage 14-Ne-Mic	AX6	*Microtus pennsylvanicus*	Montana	x			
Lineage 15-Pa-Lem	Lem6	*Lemmus sibericus*	Indigirka, Russia	x	x	x	
Lineage 15-Pa-Lem	Lem102	*Lemmus sibericus*	New Siberian Islands	x	x	x	
Lineage 15-Pa-Lem	Lem 18	*Lemmus sibericus*	Yamal, Russia			x	
Lineage 15-Pa-Lem	Lem122	*Lemmus sibericus*	New Siberian Islands			x	
Lineage 16-Pa-Mic	DL8	*Microtus middendorfi*	Yamal, Russia	x			
Lineage 16-Pa-Mic	DM7	*Microtus middendorfi*	Yamal, Russia	x			
Lineage 16-Pa-Mic	NK 270705	*Microtus gregalis*	Turgen Sum, Mongolia	x			

Table 2. Results of the mean pairwise distance analysis by locus, comparing the within and between group diversity for undescribed lineages and species together versus only described *Arostrilepis* species. The average percent distance across the dataset is provided first, separated from the distance range by a comma.

		cyt-*b*	28S	ITS2
Within Group	Lineages & Species	1.04%, 0-4.33%	0.0515%, 0-0.254%	0.0924%, 0-0.656%
	Described Species	1.02%, 0-2.14%	0.0519%, 0-0.157%	0.133%, 0-0.656%
Between Group	Lineages & Species	11.6%, 2.43-21.1%	0.755%, 0-1.72%	1.75%, 0 -5.22%
	Described Species	11.6%, 3.6-20.0%	0.713%, 0.113-1.35%	1.87%, 0.328-3.85%

Table 3. Uncorrected cyt-*b* mean pairwise genetic distances between and within species and undescribed lineages of *Arostrilepis*. Values below the highlighted diagonal are mean distances in substitutions per site between species and highlighted values are mean distances among individuals within species and lineages. Highlighted boxes with no value indicate that among individual within species or lineage distance could not be calculated because only one representative sequence was available.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	A_beringiensis	0.02144																								
2	A_cooki	0.14013	0.01312																							
3	A_gulyaevi	0.10696	0.12324	0.00604																						
4	A_intermedia	0.20021	0.09906	0.17762	0.01569																					
5	A_janickii	0.13577	0.10537	0.11166	0.15692	0.00000																				
6	A_macrocirrosa	0.14080	0.03604	0.12664	0.10740	0.12033	0.01332																			
7	A_microtis	0.11867	0.09063	0.10021	0.15993	0.09006	0.09240	0.00602																		
8	A_rauschroum	0.13156	0.07562	0.11044	0.11511	0.09436	0.08029	0.07934	0.00804																	
9	A_tenuicirrosa	0.12831	0.11166	0.09504	0.16556	0.10219	0.11625	0.10829	0.11125	0.00803																
10	Lineage_1	0.14755	0.13115	0.12246	0.17849	0.04330	0.13201	0.10803	0.11720	0.11995	0.00601															
11	Lineage_2	0.13945	0.12021	0.11403	0.15500	0.04646	0.12390	0.10857	0.10883	0.11099	0.04966															
12	Lineage_3	0.11087	0.08975	0.10283	0.14580	0.10536	0.09291	0.04871	0.07957	0.09466	0.11534	0.11239														
13	Lineage_4	0.14019	0.11665	0.11640	0.14757	0.11595	0.11888	0.12269	0.07239	0.13065	0.13842	0.12742	0.13464													
14	Lineage_5	0.15106	0.11520	0.11326	0.14709	0.09237	0.11923	0.12742	0.08112	0.12352	0.12100	0.12088	0.11696	0.09116	0.01817											
15	Lineage_6	0.14636	0.13109	0.12287	0.16909	0.12246	0.11748	0.12087	0.09619	0.13524	0.12964	0.11956	0.12238	0.08707	0.09363											
16	Lineage_7	0.21098	0.17132	0.16758	0.17486	0.18072	0.17375	0.17611	0.13025	0.19027	0.18578	0.17791	0.18069	0.11307	0.14248	0.13395										
17	Lineage_8	0.15717	0.12064	0.11933	0.14216	0.10359	0.11748	0.10217	0.06946	0.12994	0.12173	0.11462	0.10480	0.11051	0.11783	0.09620	0.14399	0.00000								
18	Lineage_9	0.12773	0.12617	0.11437	0.16391	0.12350	0.12071	0.13753	0.12704	0.13381	0.14258	0.12419	0.13795	0.13876	0.13655	0.14531	0.19414	0.14443								
19	Lineage_10	0.12817	0.09174	0.10891	0.15117	0.10271	0.08759	0.04539	0.07876	0.11003	0.11031	0.10505	0.05730	0.12907	0.12765	0.11701	0.16970	0.10794	0.14079	0.04331						
20	Lineage_11	0.08368	0.11823	0.11334	0.16919	0.11309	0.12538	0.12699	0.11542	0.12273	0.12754	0.12022	0.10885	0.13915	0.13302	0.13555	0.19875	0.15317	0.10858	0.13419	0.00000					
21	Lineage_12	0.14644	0.09639	0.10780	0.14814	0.08679	0.10025	0.10012	0.05945	0.11246	0.10458	0.10100	0.09116	0.08628	0.06982	0.07916	0.13715	0.07870	0.13755	0.10564	0.12059	0.00000				
22	Lineage_13	0.13812	0.09882	0.12352	0.14605	0.07575	0.11147	0.10375	0.08186	0.12888	0.09661	0.08627	0.10649	0.11543	0.09769	0.11129	0.15874	0.08870	0.14092	0.10545	0.12192	0.06723	0.00910			
23	Lineage_14	0.13170	0.08091	0.10253	0.14438	0.08508	0.09642	0.10300	0.04977	0.10361	0.10281	0.09924	0.09636	0.07957	0.06322	0.08582	0.14075	0.08868	0.12489	0.10154	0.10807	0.02432	0.07885	0.00601		
24	Lineage_15	0.14010	0.11562	0.10193	0.13984	0.09525	0.13042	0.11585	0.06796	0.11704	0.12737	0.09923	0.10680	0.11598	0.10068	0.11192	0.14255	0.06076	0.13492	0.10998	0.13610	0.08389	0.09484	0.08388	0.00903	
25	Lineage_16	0.12420	0.08949	0.09906	0.14166	0.09121	0.09037	0.03907	0.07263	0.09382	0.10107	0.09811	0.03164	0.12940	0.11779	0.12086	0.16107	0.10107	0.13513	0.03976	0.11323	0.09668	0.10510	0.09493	0.09840	0.01416

Table 4. 28S uncorrected mean pairwise genetic distances between and within species and undescribed lineages of *Arostrilepis*. Values below the highlighted diagonal are mean distances in substitutions per site between species and highlighted values are mean distances among individuals within species and lineages. Highlighted boxes with no value indicate that among individual within species or lineage distance could not be calculated because only one representative sequence was available.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	A_beringiensis	0.00157																					
2	A_cooki	0.01353	0.00038																				
3	A_gulyaevi	0.00835	0.00950	0.00113																			
4	A_janickii	0.00854	0.00492	0.00454	0.00000																		
5	A_macrocirrosa	0.01238	0.00113	0.00835	0.00378	0.00000																	
6	A_microtis	0.00949	0.00778	0.00549	0.00283	0.00663	0.00000																
7	A_rauschorum	0.01223	0.00572	0.00820	0.00458	0.00458	0.00648	0.00107															
8	A_tenuicirrosa	0.00853	0.00970	0.00549	0.00473	0.00855	0.00568	0.00839	0.00000														
9	Lineage_1	0.00854	0.00492	0.00454	0.00000	0.00378	0.00283	0.00458	0.00473	0.00000													
10	Lineage_2	0.00806	0.00588	0.00549	0.00094	0.00473	0.00378	0.00553	0.00568	0.00094													
11	Lineage_3	0.00189	0.01161	0.00739	0.00663	0.01046	0.00758	0.01031	0.00663	0.00663	0.00758												
12	Lineage_4	0.01335	0.00969	0.00931	0.00664	0.00854	0.00759	0.00839	0.00951	0.00664	0.00759	0.01143											
13	Lineage_5	0.01237	0.00873	0.00835	0.00568	0.00758	0.00473	0.00647	0.00854	0.00568	0.00663	0.01046	0.00854										
14	Lineage_6	0.01334	0.00968	0.00930	0.00663	0.00853	0.00568	0.00838	0.00950	0.00663	0.00759	0.01141	0.00950	0.00473									
15	Lineage_7	0.01721	0.01161	0.01315	0.01046	0.01046	0.01142	0.00935	0.01335	0.01046	0.01142	0.01528	0.00758	0.01045	0.01333								
16	Lineage_8	0.01143	0.00493	0.00740	0.00379	0.00378	0.00569	0.00079	0.00760	0.00379	0.00474	0.00951	0.00759	0.00663	0.00759	0.00951	0.00000						
17	Lineage_9	0.01622	0.01065	0.01217	0.00949	0.00950	0.01045	0.01127	0.01045	0.00949	0.01045	0.01429	0.01335	0.01333	0.01430	0.01528	0.01047						
18	Lineage_10	0.01079	0.00907	0.00677	0.00411	0.00792	0.00127	0.00777	0.00696	0.00411	0.00506	0.00887	0.00888	0.00601	0.00697	0.01272	0.00697	0.01174	0.00254				
19	Lineage_12	0.01143	0.00588	0.00740	0.00474	0.00473	0.00569	0.00458	0.00760	0.00474	0.00569	0.00951	0.00378	0.00663	0.00759	0.00568	0.00379	0.00951	0.00697				
20	Lineage_13	0.01046	0.00683	0.00644	0.00189	0.00569	0.00284	0.00648	0.00664	0.00189	0.00284	0.00854	0.00855	0.00569	0.00664	0.01239	0.00569	0.01141	0.00411	0.00664	0.00000		
21	Lineage_14	0.01238	0.00683	0.00835	0.00569	0.00568	0.00664	0.00553	0.00855	0.00569	0.00664	0.01046	0.00473	0.00758	0.00854	0.00663	0.00474	0.01047	0.00792	0.00094	0.00760		
22	Lineage15	0.01142	0.00777	0.00930	0.00663	0.00663	0.00759	0.00553	0.00950	0.00663	0.00759	0.00950	0.00759	0.00663	0.00949	0.00759	0.00568	0.01238	0.00887	0.00378	0.00854	0.00473	0.00000

Table 5. ITS2 uncorrected mean pairwise genetic distances between and within species and undescribed lineages of *Arostrilepis*. Values below the highlighted diagonal are mean distances in substitutions per site between species and highlighted values are mean distances among individuals within species and lineages. Highlighted boxes with no value indicate that among individual within species or lineage distance could not be calculated because only one representative sequence was available.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	A_beringiensis	0.00000																					
2	A_cooki	0.02100	0.00656																				
3	A_gulyaevi	0.03257	0.02410	0.00000																			
4	A_intermedia	0.02686	0.01124	0.03391	0.00000																		
5	A_janickii	0.01995	0.00459	0.02686	0.00658	0.00000																	
6	A_macrocirrosa	0.02345	0.00792	0.03044	0.00994	0.00328	0.00000																
7	A_microtis	0.02767	0.01205	0.03472	0.00743	0.00739	0.01076	0.00163															
8	A_rauschorum	0.02420	0.00872	0.03118	0.01073	0.00409	0.00741	0.01155	0.00164														
9	A_tenuicirrosa	0.03849	0.01842	0.02454	0.02463	0.01774	0.02122	0.02544	0.02198	0.00218													
10	Lineage_1	0.01995	0.00459	0.02686	0.00658	0.00000	0.00328	0.00739	0.00409	0.01774	0.00000												
11	Lineage_2	0.03368	0.01790	0.03385	0.01999	0.01320	0.01661	0.02080	0.01739	0.02458	0.01320												
12	Lineage_4	0.02333	0.00788	0.03029	0.00989	0.00327	0.00658	0.01070	0.00738	0.01886	0.00327	0.01654											
13	Lineage_5	0.03368	0.01790	0.03385	0.01999	0.01320	0.01661	0.02080	0.01739	0.02231	0.01320	0.01995	0.00987										
14	Lineage_6	0.03357	0.01785	0.03374	0.01993	0.01316	0.01656	0.02073	0.01734	0.02224	0.01316	0.01320	0.00984	0.01320									
15	Lineage_7	0.05192	0.03540	0.05219	0.03765	0.03050	0.02691	0.03846	0.03487	0.04012	0.03050	0.03059	0.02704	0.03059	0.03050								
16	Lineage_8	0.02673	0.01119	0.02686	0.01322	0.00655	0.00989	0.01403	0.00409	0.01774	0.00655	0.01320	0.00984	0.01320	0.01316	0.03050	0.00000						
17	Lineage_9	0.03024	0.02126	0.03039	0.01661	0.01654	0.01999	0.01743	0.02075	0.02800	0.01654	0.02333	0.01989	0.02333	0.02326	0.03408	0.01654						
18	Lineage_10	0.02681	0.01122	0.02694	0.01326	0.00657	0.00992	0.00737	0.01071	0.01779	0.00657	0.01324	0.00987	0.01324	0.01320	0.03059	0.00657	0.01659	0.00000				
19	Lineage_12	0.02333	0.00788	0.02345	0.00989	0.00327	0.00658	0.01070	0.00738	0.01438	0.00327	0.00987	0.00655	0.00987	0.00984	0.02704	0.00327	0.01320	0.00328				
20	Lineage_13	0.03033	0.01460	0.03048	0.01667	0.00990	0.01330	0.01748	0.01409	0.02125	0.00990	0.01664	0.01324	0.01664	0.01659	0.03418	0.00990	0.02001	0.00993	0.00659			
21	Lineage_14	0.02673	0.01119	0.02686	0.01322	0.00655	0.00989	0.01403	0.01068	0.01774	0.00655	0.01320	0.00984	0.01320	0.01316	0.03050	0.00655	0.01654	0.00657	0.00327	0.00990		
22	Lineage_15	0.02681	0.01122	0.03385	0.01326	0.00657	0.00992	0.01407	0.01071	0.02458	0.00657	0.01995	0.00987	0.01995	0.01989	0.03758	0.01320	0.02333	0.01324	0.00987	0.01664	0.01320	0.00000

Table 6. BioGeoBEARS results table AIC scores.

Model	LnL	Number of Free Parameters	d	e	j	AIC	AIC weight	Relative likelihood AIC
DEC	-59.97992736	2	101.1373689	161.6580106	0	123.9598547	0.000313743	0.000313743
DEC+J	-50.9133253	3	367.3932439	172.0654179	0.046954224	107.8266506	0.999667043	0.999667043
DECnoMult	-77.95835019	2	8.138517334	29.52100871	0	159.9167004	4.88E-12	4.88E-12
DEC+JnoMult	-62.76692522	3	23.7015664	38.22866805	0.039313214	131.5338504	7.11E-06	7.11E-06
DIVALIKE	-85.43282359	2	35.23132138	422.4464794	0	174.8656472	2.77E-15	2.77E-15
DIVALIKE+J	-66.25351174	3	71.25592341	422.0182445	0.21997399	138.5070235	2.18E-07	2.18E-07
BAYAREALIKE	-192.1077597	2	0.674023589	0.564996794	0	388.2155194	1.30E-61	1.30E-61
BAYAREALIKE+J	-62.2531777	3	27.29832023	31.16372721	0.191918695	130.5063554	1.19E-05	1.19E-05

Table 7. BioGeoBEARS results table corrected AIC scores

Model	LnL	Number of Free Parameters	d	e	j	AICc	AICc Weight_versus Best	Relative likelihood AICc
DEC	-59.97992736	2	101.1373689	161.6580106	0	124.5053093	0.000422913	0.000422913
DEC+J	-50.9133253	3	367.3932439	172.0654179	0.046954224	108.9695077	0.999557875	0.999557875
DECnoMult	-77.95835019	2	8.138517334	29.52100871	0	160.4621549	6.58E-12	6.58E-12
DEC+JnoMult	-62.76692522	3	23.7015664	38.22866805	0.039313214	132.6767076	7.11E-06	7.11E-06
DIVALIKE	-85.43282359	2	35.23132138	422.4464794	0	175.4111017	3.73E-15	3.73E-15
DIVALIKE+J	-66.25351174	3	71.25592341	422.0182445	0.21997399	139.6498806	2.18E-07	2.18E-07
BAYAREALIKE	-192.1077597	2	0.674023589	0.564996794	0	388.7609739	1.75E-61	1.75E-61
BAYAREALIKE+J	-62.2531777	3	27.29832023	31.16372721	0.191918695	131.6492126	1.19E-05	1.19E-05

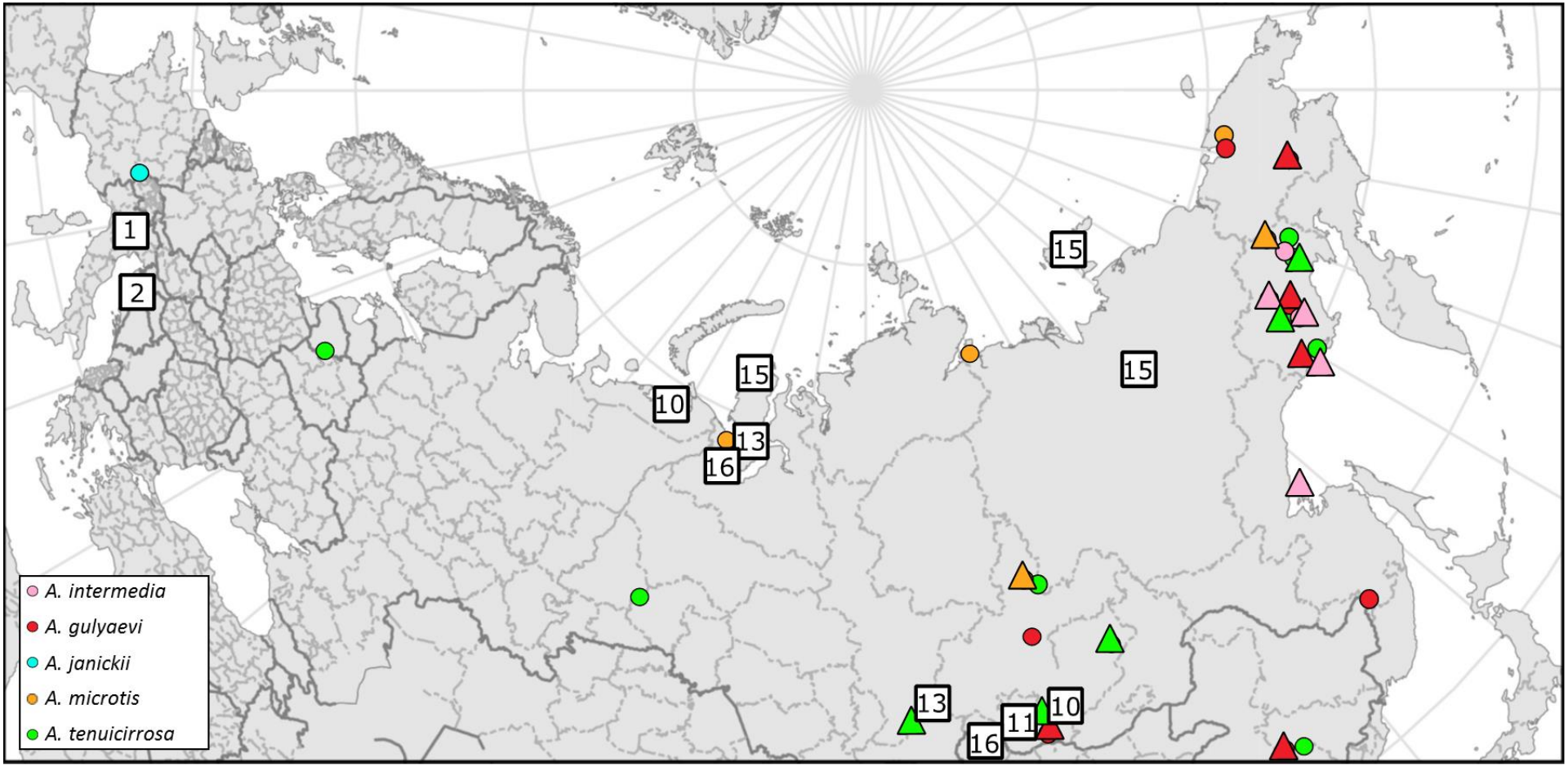

Figure 1. Map showing approximate sampling localities for *Arostrilepis* specimens collected in Eurasia. Markers of different colors signify particular species (see inset key). White boxes denote undescribed genetic lineages, the number associated with each lineage appears in the box. Multiple records for a single species or lineage from a small geographic area are denoted by a single marker for clarity. Markers have been offset for clarity at single localities where multiple species or lineages were collected from the same locality. Holarctic species are omitted on this map. Triangles indicate that molecular data from at least one individual of that species, from that locality was included in the analysis. Molecular data were used from each undescribed lineage.

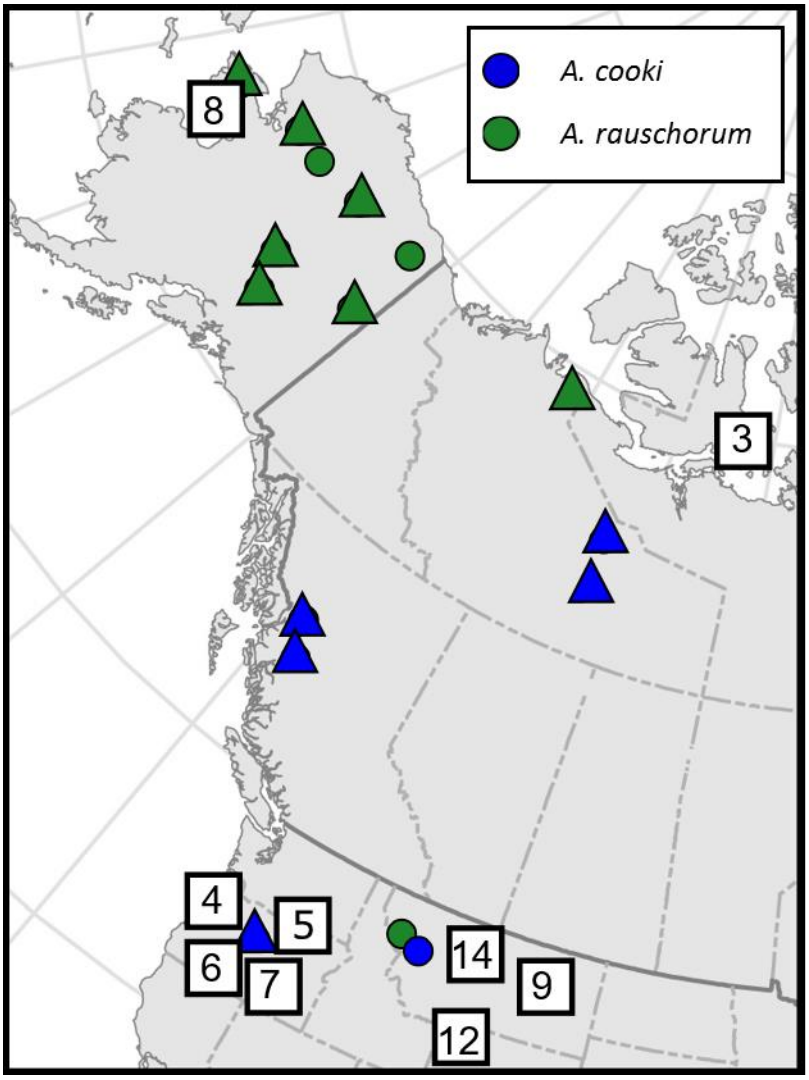

Figure 2. Map showing approximate sampling localities for *Arostrilepis* specimens collected in North America. Markers of different colors signify particular species (see inset key). White boxes denote undescribed genetic lineages, the number associated with each lineage appears in the box. Multiple records for a single species or lineage from a small geographic area are denoted by a single marker for clarity. Markers have been offset for clarity at single localities where multiple species or lineages were collected. Holarctic species are omitted on this map. Triangles indicate that molecular data from at least one individual of that species, from that locality was included in the analysis. Molecular data were used from each undescribed lineage.

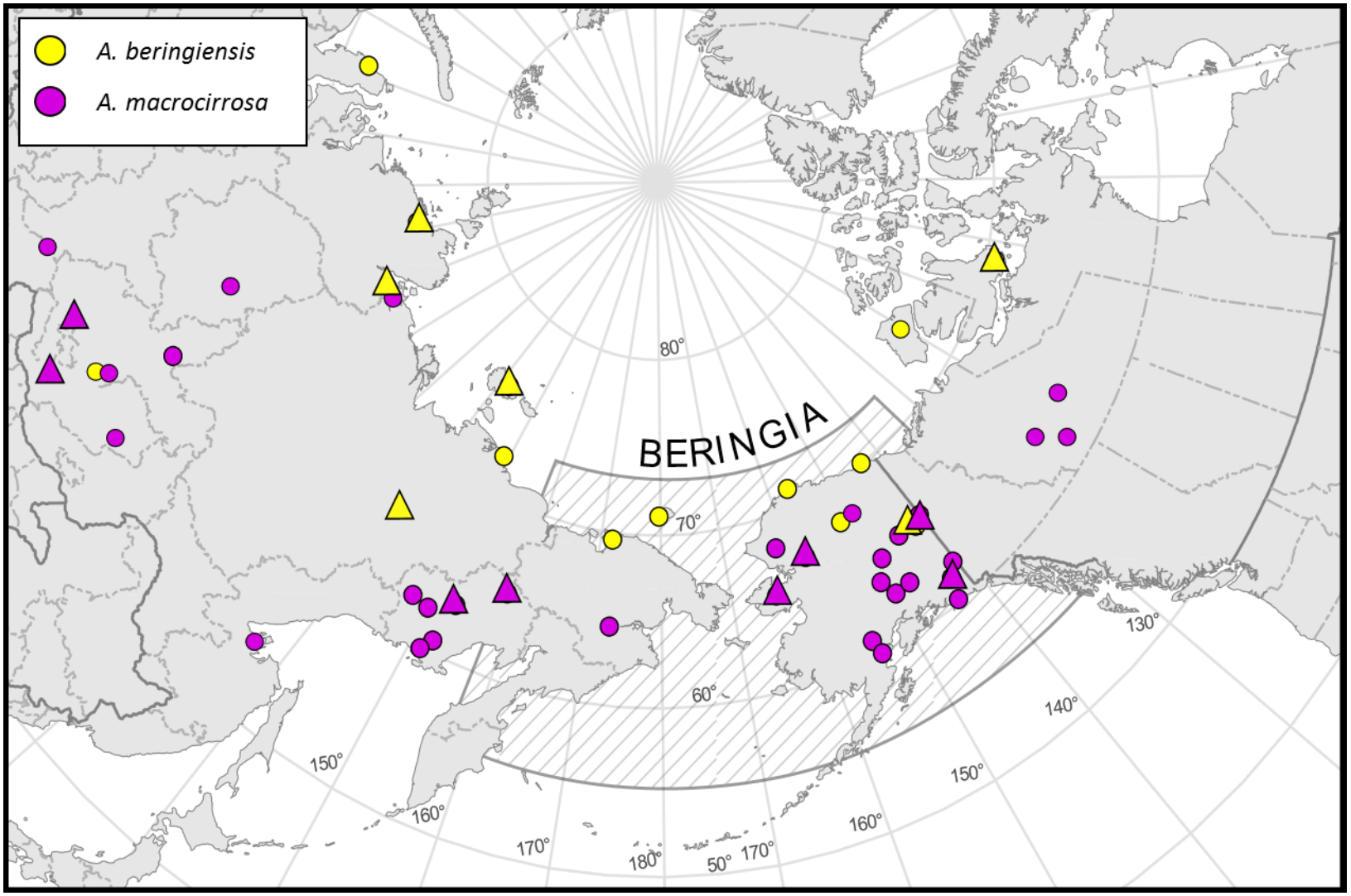

Figure 3. Map showing approximate sampling localities for Holarctic *Arostrilepis* specimens collected in Eurasia and North America. The different colors signify particular species (see inset key). Multiple records for a single species from a small geographic area are denoted by a single marker for clarity. Markers have been offset for clarity at single localities where multiple species were collected. Triangles indicate that molecular data from at least one individual of that species, from that locality were included in the analysis.

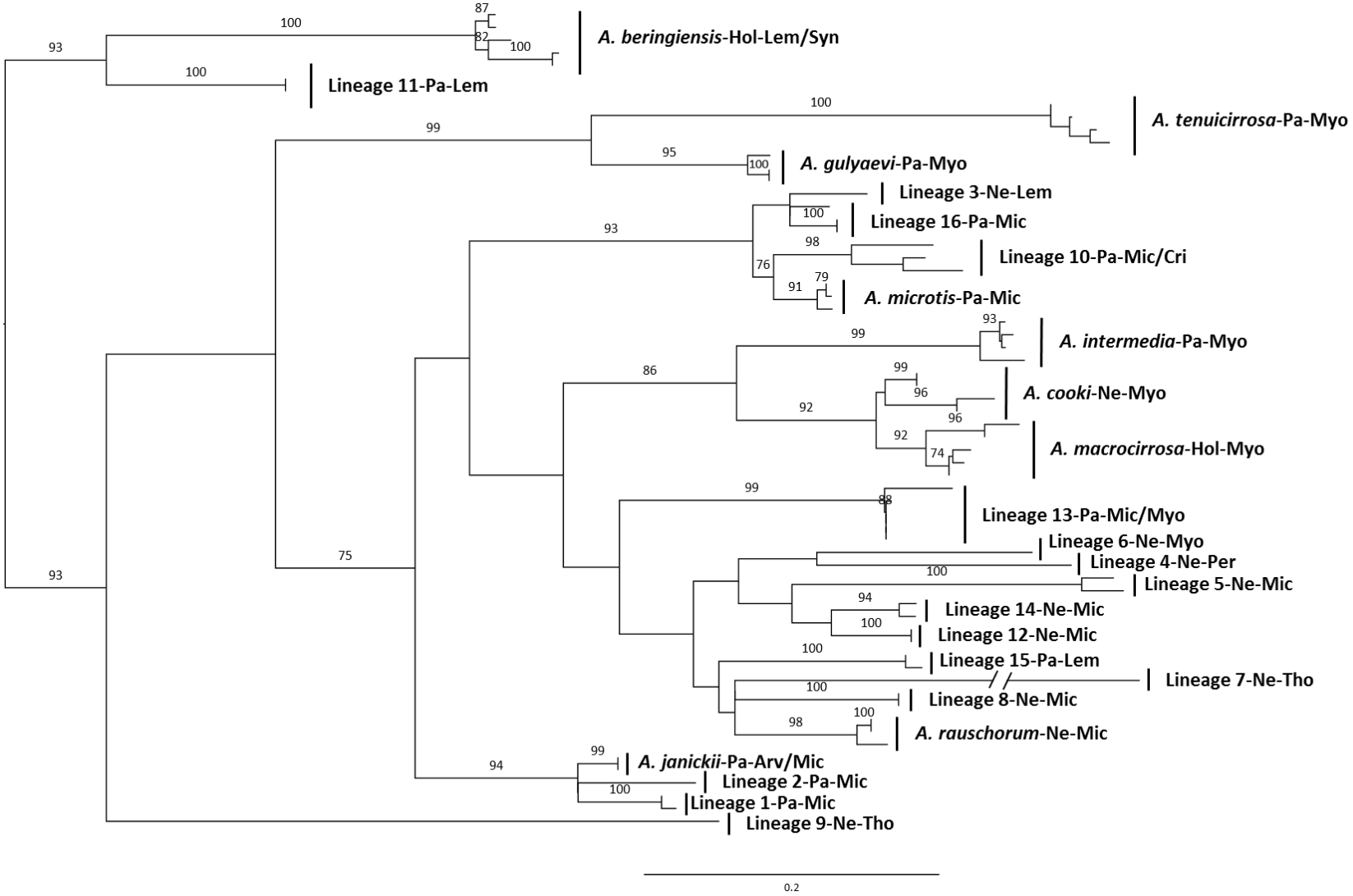

Figure 4. Maximum Likelihood reconstruction of partial cyt-*b* (mtDNA) sequences of *Arostrilepis* species and undescribed genetic lineages. Values on branches show bootstrap support based on 1000 rapid RAxML replicates. Broad geographic distribution is indicated after the species name or lineage number by the abreviations Hol-Holarctic, Pa-Palearctic, or Ne-Nearctic. Primary host association is designated by the first three letters of the host genus: Arv-*Arvicola*, Cri-*Cricetulus*, Mic-*Microtus*, Myo-*Myodes*, Lem-*Lemmus*, Per-*Peromyscus*, Syn-*Synaptomys*, Tho-*Thomomys.* Multiple codes are given if the parasite is commonly found in multiple host genera. The branch length for Lineage 7-Ne-Tho is exceptionally long, and was truncated here for clarity.

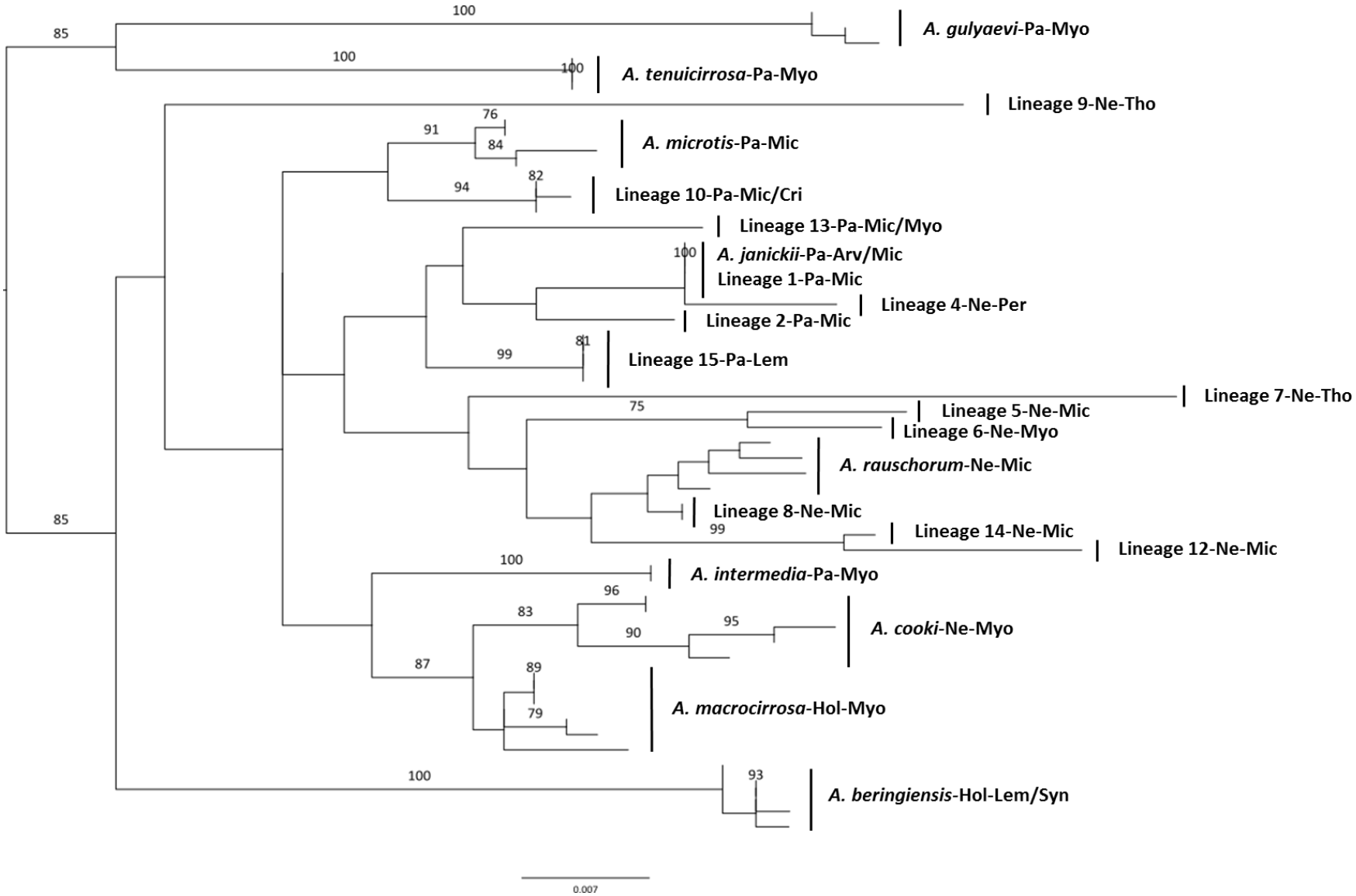

Figure 5. Maximum Likelihood reconstruction of ITS2 (rDNA) sequences of *Arostrilepis* species and undescribed genetic lineages. Values on branches show bootstrap support based on 1000 rapid RAxML replicates. Broad geographic distribution is indicated after the species name or lineage number by the abreviations Hol-Holarctic, Pa-Palearctic, or Ne-Nearctic. Primary host association is designated by the first three letters of the host genus: Arv-*Arvicola,* Cri-*Cricetulus,* Mic-*Microtus,* Myo-*Myodes*, Lem-*Lemmus,* Per-*Peromyscus,* Syn-*Synaptomys,* Tho-*Thomomys.* Multiple codes are given if the parasite is commonly found in multiple host genera.

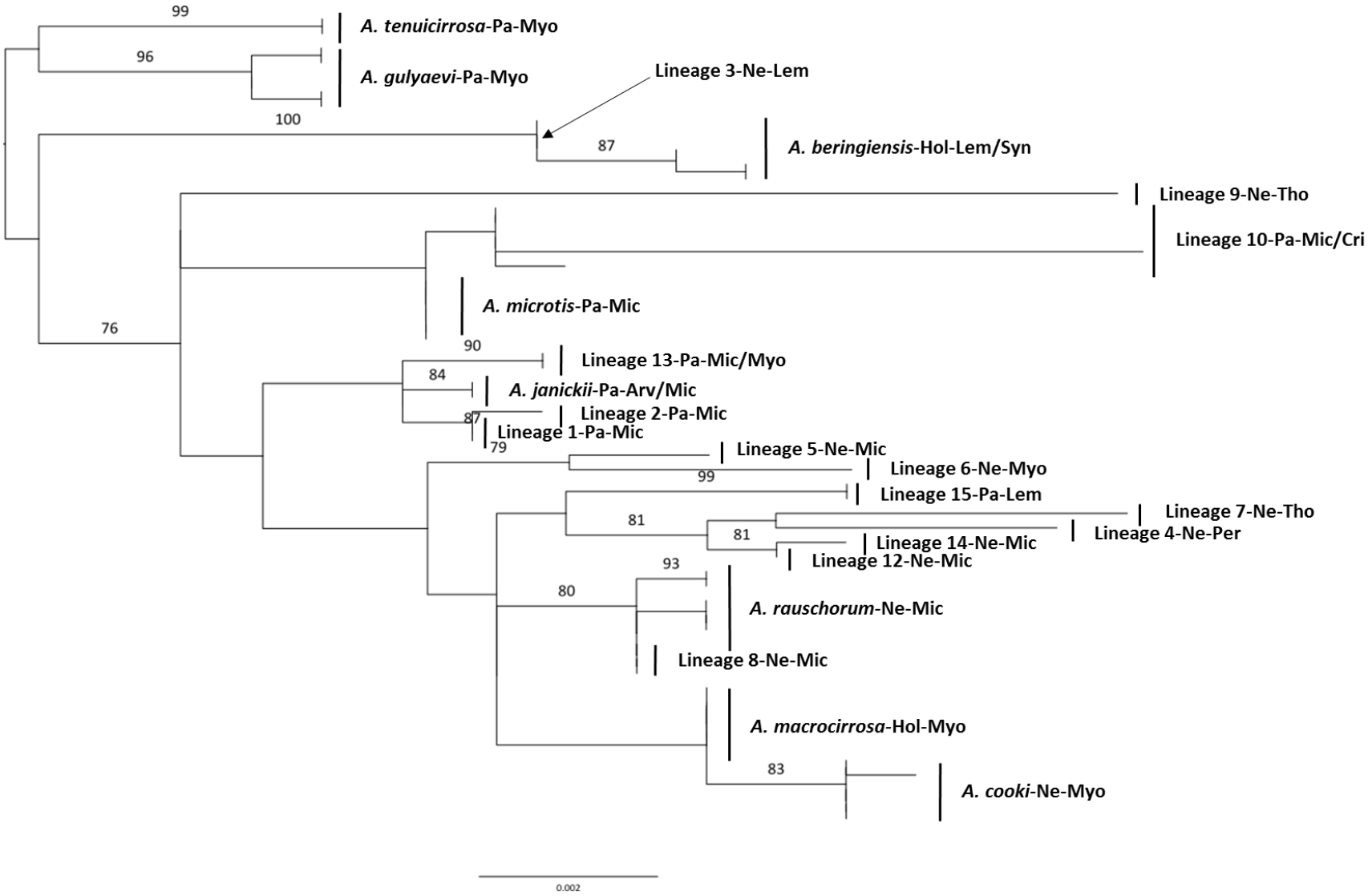

Figure 6. Maximum Likelihood reconstruction of 28S (rRNA) sequences of *Arostrilepis* species and undescribed genetic lineages. Values on branches show bootstrap support based on 1000 rapid RAxML replicates. Broad geographic distribution is indicated after the species name or lineage number by the abreviations Hol-Holarctic, Pa-Palearctic, or Ne-Nearctic. Primary host association is designated by the first three letters of the host genus: Arv-*Arvicola,* Cri-*Cricetulus,* Mic-*Microtus,* Myo-*Myodes,* Lem-*Lemmus,* Per-*Peromyscus,* Syn-*Synaptomys,* Tho-*Thomomys.* Multiple codes are given if the parasite is commonly found in multiple host genera. Lineage 3-Ne-Lem was placed within *A. beringiensis* diversity at this locus and is indicated by an arrow. This placement conflicts with what was found in the cyt-*b* reconstruction (Figure 4).

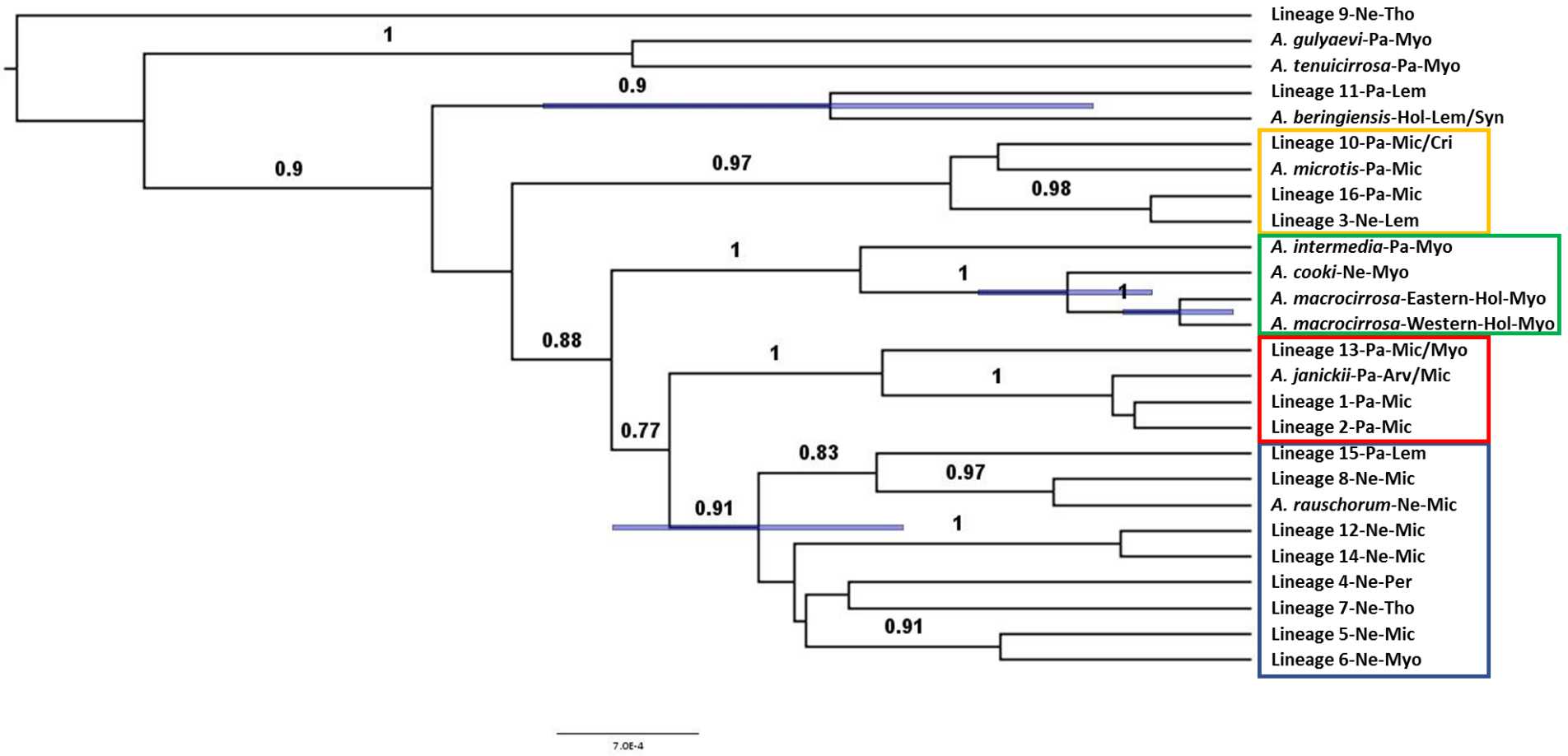

Figure 7. Multi-locus coalescent reconstruction of *Arostrilepis* species and undescribed genetic lineages using cyt-*b*, 28S, and ITS2 loci. Values on branches show Bayesian posterior probability support. Broad geographic distribution is indicated after the species name or lineage number by the abreviations Hol-Holarctic, Pa-Palearctic, or Ne-Nearctic. Primary host association is designated by the first three letters of the host genus: Arv-*Arvicola*, Cri-*Cricetulus*, Mic-*Microtus*, Myo-*Myodes*, Lem-*Lemmus*, Per-*Peromyscus*, Syn-*Synaptomys*, Tho-*Thomomys*. Multiple codes are given if the parasite is commonly found in multiple host genera. *Arostrilepis macrocirrosa* was subdivided into its Eastern and Western populations for this analysis. Colored boxes indicate well-supported discrete clades that represent notable associations to specific host groups or geographic distributions: Yellow—Central Palearctic clade, Green—*Myodes*-associated clade, Red—Western Palearctic clade, and Blue—Temperate Nearctic clade. Purple bars at nodes show the 95% High Posterior Density (HPD) interval of node height for the nodes that anchor transberingian colonization events inferred via biogeographic analysis.

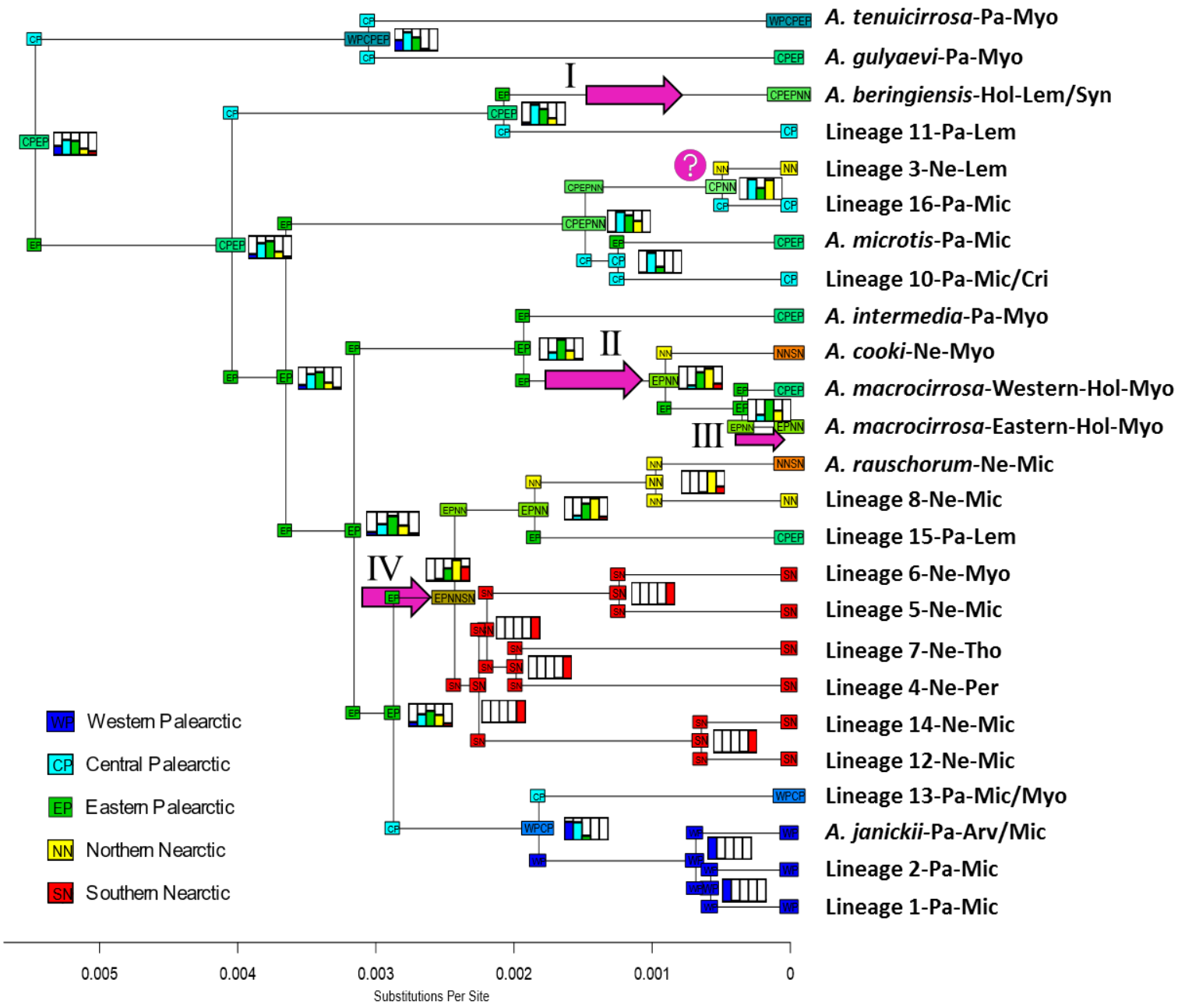

Figure 8. Ancestral state estimates for the DEC*+J model run on the three-locus coalescent *Arostrilepis* phylogeny. Present *Arostrilepis* distributions are shown on the tips of the tree. At each node the single most-probable ancestral range is shown alongside bar charts which indicate the estimated probability of occurrence in each possible geographic area for that node. Range estimates on the corners represent geographic ranges immediately following splitting events. Four inferred eastward transberingian colonization events are indicated by Roman numerals associated with purple arrows placed along the branches where the ancestral *Arostrilepis* would be undergoing a range expansion. The arrow for colonization event III has been offset for clarity; the geographic range expansion that proceeded that event is inferred to have occurred between the node where the two *A. macrocirrosa* populations coalesce and the corner that leads to the eastern population. The question mark denotes a possible fifth transberingian colonization.