4-2017

COMPARATIVE BIOGEOGRAPHY OF NORTH AMERICAN PIKA PARASITES: UNRAVELING A HISTORY DRIVEN BY CLIMATE CHANGE

Heather M. Toman
htoman@nmu.edu

Recommended Citation

Toman, Heather M., "COMPARATIVE BIOGEOGRAPHY OF NORTH AMERICAN PIKA PARASITES: UNRAVELING A HISTORY DRIVEN BY CLIMATE CHANGE" (2017). *All NMU Master's Theses*. 140.
https://commons.nmu.edu/theses/140

COMPARATIVE BIOGEOGRAPHY OF NORTH AMERICAN PIKA PARASITES: UNRAVELING A HISTORY DRIVEN BY CLIMATE CHANGE

By

Heather M. Toman

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

April 2017

SIGNATURE APPROVAL FORM

COMPARATIVE BIOGEOGRAPHY OF PIKA PARASITES: UNRAVELLING A HISTORY DRIVEN BY CLIMATE CHANGE

This thesis by Heather Toman is recommended for approval by the student's Thesis Committee and Department Head in the Department of Biology and by the Assistant Provost 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. John Bruggink Date

Department Head: Dr. John Rebers Date

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

ABSTRACT

COMPARATIVE BIOGEOGRAPHY OF NORTH AMERICAN PIKA PARASITES: UNRAVELING A HISTORY DRIVEN BY CLIMATE CHANGE

By

Heather M. Toman

Parasites are an understudied group that can provide important information on ecosystem dynamics and climate change as well as host biogeographic history. I performed a comparative biogeographic study on two endoparasitic helminth lineages associated with pikas (*Ochotona*) -the tapeworm genus *Schizorchis* and the pinworm subgenus *Labiostomum (Eugenuris)*. Colonization history across Beringia was assessed using analyses of historical range estimation and tests for simultaneous divergence. Analysis results allow me reject the hypothesis that these parasites colonized the Nearctic during a single glacial cycle, as evidenced by multiple Nearctic and Palearctic lineage pairs that did not diverge simultaneously. This evidence indicates a pre-Pleistocene (>2.5Mya) history of multiple, independent colonizations of pikas and their parasites across Beringia primarily in an eastward direction. Comparison of these two major parasite lineages indicates semi-independent histories, with trans-Beringian parasite lineage formation likely driven by taxon pulses, with expansion and vicariance cycles reflecting periods of environmental stability and perturbation, leading to complex histories of colonization and host-switching within pikas.

ACKNOWLEDGEMENTS

I greatly appreciate the contribution and support of the biology faculty at Northern Michigan University, including the Department Head Dr. John Rebers who always had the answers. I thank my committee members, Dr. John Bruggink, Dr. Alec Lindsay, and Dr. Eric Hoberg for their support in analyzing and reporting this work. Warm thanks to my labmates Gena Haas and Sarah Gallagher, and to Mary Martin for their friendship and collaboration. Most of all, I thank my advisor, Dr. Kurt Galbreath, who provided unquantifiable support during all aspects of my time as a masters student. He invested innumerable hours explaining and discussing related concepts as well as provided emotional support. Every office visit left me feeling much more optimistic and uplifted than when I walked in. I also thank him for the financial support to conduct research and participate in fieldwork during the summers. I feel privileged to have been a part of three wonderful field crews in Canada and Mongolia, who all worked tirelessly while developing great friendships with each other.

Deep respect and thanks to the National Science Foundation for granting funding to specimen collection efforts, such as the Beringian Coevolution Project, that has been gathering baseline data for mammals and their parasites across the Beringian region for ten years. This NSF grant, as well as ones awarded to Dr. Kurt Galbreath to collect pikas and their parasites across the American Intermountain West have made research possible for many investigators, including myself (DEB 0196095, DEB 0415668, DEB 1258010, DEB 0506042 and DEB 1256943).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Over half of all known species can be classified as parasites, and these numbers are undoubtedly underestimated due to the continual discovery of many cryptic species within known groups (Cook *et al.*, 2005; Dobson *et al.*, 2008). Parasites have been understudied despite their potential to provide important information regarding ecosystem dynamics and the effects of historical climate and geological changes (Bermingham & Moritz, 1998). Traditionally, the evolution of parasites has been widely viewed as directly linked to the evolution of their host. This model of cospeciation predicts host and parasite phylogenies to be perfectly congruent (Brooks, 1979; Brooks & McLennan, 1993). However, the accumulation of co-phylogenetic studies partnered with a more detailed understanding of parasite taxonomy has revealed that perfect cospeciation is rare (e.g. Brant & Garder, 2000; Brooks & Ferrao, 2005; Huyse & Volckaert, 2005; Hoberg et al., 2012; Hoberg & Brooks, 2013; Brooks et al., 2015).

Events that lead to differences between host and parasite phylogeny include lineage extinction or duplication, but are primarily driven by host-switching events that correlate with extrinsic historical periods of strong environmental perturbation (such as the episodic glacial cycling of the Quaternary) (Hafner & Page, 1995; Johnson *et al.*, 2002; Zarlenga *et al.*, 2006; Hoberg *et al.*, 2012; Hoberg & Brooks, 2013). During these periods, hosts are most likely to be moving across the landscape in search of suitable habitat, which creates opportunity for contact between species that did not previously overlap. Parasites in this situation are much more likely to successfully switch hosts into species that offer similar resources. Additionally, vicariance events that split or bottleneck host populations and expose them to different environmental pressures can have discordant phylogenetic outcomes for their parasites (Hoberg and Brooks, 2008; Hoberg *et al.*, 2012). However, when environmental conditions are stable, we may see co-divergence between tightly associated organisms (e.g. Wickström *et al.*, 2003; Haukisalmi *et al*., 2015).

Studies on parasite biogeography are important in order to gain a more thorough understanding of the processes that lead to parasite diversification. They can also offer insight into future distributions of potential emerging infectious diseases through the use of predictive modeling (Brooks *et al*., 2014; Hoberg & Brooks, 2015). Another advantage arises from the close association parasites have with their hosts, where a comparative phylogeographic approach can be used to fill in gaps present in host history and illuminate a shared history that may be unavailable by looking at the host alone (Criscione *et al*., 2005; Nieberding & Olivieri, 2007). In order to distinguish shared historical events from parasite specific processes it is important to include multiple co-distributed, independent lineages of parasites whenever possible (e.g., Galbreath & Hoberg, 2012). In doing this, parasite-specific patterns can also be uncovered, and we can begin to parse out the extrinsic historical factors that could have influenced the evolution of these independent parasite lineages similarly, from differences between parasite and host history that are specific to a single lineage (Galbreath & Hoberg, 2015). In uncovering these patterns of parasite biogeography, and linking them to what is known about the history of the parasite's environment, we also begin to learn important details about past conditions of a region and how they have influenced assemblages of fauna over time (Cook *et al*., 2005; Hoberg *et al*., 2012; Hoberg & Brooks, 2013).

Beringia, which encompasses the area surrounding the Bering Strait from eastern Siberia to northwestern North America (Figure 1), is central to our understanding of Holarctic species distributions and genetic diversity. Since the separation of the continents at the beginning of the Neogene (~25 Myr), Beringia has become an intermittent terrestrial corridor and refugium during glacial maxima, and a barrier to gene flow during warmer periods (Hopkins, 1959; Sher, 1999; Cook *et* al., 2005; DeChaine 2008). Until the late Pliocene (~3.6 Myr), climate was not a limiting factor for biota during land bridge exposure, as northern conditions were significantly warmer than seen currently. Therefore, major exchanges of land vertebrates were bi-directional and symmetrical. But during the Pleistocene (2.58 Myr – 11.7 Kyr), temperature changes became more extreme and ice sheets advanced and receded with longer periods of more extreme cooling emphasized (Dyke and Prest, 1987; Williams *et al.*, 1998). During this time, Asia and Beringia remained largely ice-free with relatively mild conditions, which allowed Asian species to move into the exposed Beringian land mass. Species from North America were cut off from Beringia due to the Laurentide and Cordilleran ice sheets that covered the majority of Canada and the northern United States (Dyke and Prest, 1987; Roberts, 1991; Williams *et al.*, 1998). As temperatures warmed during interglacials, a corridor between the ice sheets opened, allowing the flora and fauna established in Beringia to move east into North America, thus creating a pattern of mostly west to east colonization (Waltari *et al.*, 2007). However, recent studies are beginning to reveal evidence for multiple colonizations of Holarctic species occurring in both directions across the Bering land bridge during the Pleistocene (e.g. Dawson *et al.*, 2013; Kohli *et al.*, 2015).

The endoparasitic helminths of pikas (genus *Ochotona*) provide an excellent comparative system for investigating colonization history across Beringia and diversification in North America, as they are diverse, widespread, and host-specific to pikas (Grundmann & Lombardi, 1976; Rausch & Smirnova, 1984). Pikas are small lagomorphs with 30 extant species within the single *Ochotona* genus. Most species occur in the Palearctic (Eurasia). The only two that occur in North America are sister species, with *O. collaris* found in Alaska and northern Canada, and *O. princeps* occurring farther south in the American West (Hoffmann & Smith, 2005). The ancestor of the *Ochotona* genus arose in Eurasia, and North American pika species presumably descended from a single ancestral colonization eastward across the Bering land bridge (Rausch & Ritter, 1973; Niu *et al*., 2004; Formozov *et al*., 2006; Lissovsky *et al*., 2007; Lanier & Olson, 2009). Similar to their hosts, the closest relatives to parasites of North American pikas occur in Eurasia. They presumably arrived in the Nearctic with pikas, and diversified after colonization, leading to the species now found in North America (Hoberg *et al*., 2012; Galbreath & Hoberg, 2015). Recent work has hinted that this group likely underwent a more complex history of movement, with some limited evidence for a recent westward movement into the Palearctic (Galbreath & Hoberg, 2012).

Following the simple west-to-east model for Beringian colonization, *O. collaris* was originally thought to be directly descended from Beringian colonizers from Asia, and *O. princeps* originated from individuals who later colonized southward along mountain corridors (Guthrie, 1973). However, phylogenetic evidence from multiple lineages of pika parasites indicates that *O. collaris* likely arose from a southern ancestor (Galbreath and Hoberg 2012). Parasites from *O. princeps* were consistently found to be ancestral to parasites found in *O. collaris*, indicating that *O. princeps* is the descendent lineage from Asian pikas that originally came across the Bering land bridge. After moving south into the Intermountain West, pikas subsequently moved back north during a more recent glacial cycle, with glacial vicariance leading to the speciation of *O. collaris.* This revised and more detailed history was inferred based on evidence from a single mitochondrial DNA locus, and suffered from a limited pool of Eurasian parasite specimens. Individual gene genealogies often do not match true evolutionary history due to stochastic processes such as incomplete lineage sorting, and maternally inherited mtDNA may only reflect the history of females (Ballard & Whitlock, 2004; McCormack *et al*., 2009; Leaché & Rannala, 2011). The addition of specimens from a wider geographic area, specifically the Palearctic, as well as sequencing loci from both nuclear and mitochondrial genomic regions, is likely to recover a more complete species history.

Here I provide a multi-locus comparative biogeographical analysis of two major groups of parasites found in pikas, the pinworm subgenus *Labiostomum (Eugenuris*) Schultz, 1948, and the tapeworm genus *Schizorchis* Hansen, 1948. Previous assessments provide a basis to test alternative colonization and diversification hypotheses for this assemblage.

Hypothesis 1 (H1): Amphi-beringian colonization and diversification history

H1a: Pika parasite lineages colonized the Nearctic via the Bering Land Bridge during a single glacial cycle. Subsequent diversification occurred within North America; therefore all Nearctic pika diversity should be monopyletic with all Nearctic parasite species restricted to the two host species, *O. princeps* and *O. collaris.* Extended sampling of parasites in the Palearctic will only reveal lineages that are highly divergent from

Nearctic parasites. **H1b**: A large single colonization occurred over the Bering land bridge into the Nearctic that brought multiple parasite species for each major lineage. Nearctic pika parasite species could have maintained sister relationships with populations in the Palearctic under this scenario and are therefore not restricted to *O. princeps* and *O. collaris.* Tests should indicate simultaneous divergence between all taxon pairs found in the Palearctic and Nearctic. **H1c:** Multiple colonizations across the Bering land bridge have occurred during independent glacial cycles for one or both major parasite lineages. Nearctic parasites are predicted to be polyphyletic with respect to Palearctic parasites and tests for simultaneous divergence for independent taxon pairs across the Bering Strait will be rejected.

Hypothesis 2 (H2): Nearctic colonization history

H2a: *Ochotona collaris* colonized northward from a southern ancestor, as hypothesized by Galbreath and Hoberg (2012). The diversity of *O. collaris* parasites is predicted to be phylogenetically nested within those of southern *O. princeps* parasites for both parasite lineages. **H2b**: *Ochotona princeps* is derived from a northern Nearctic ancestor. Results supporting this hypothesis will show the diversity of *O. princeps* parasites nested within northern *O. collaris* parasites for both parasite lineages. **H2c**: The two parasite lineages will show conflicting results regarding host ancestry. *Schizorchis* and *L. (Eugenuris*) will differ in their phylogenetic structure concerning their host associations and colonization history, indicating a more complicated history of movement north and south. **H2d**: For both *Schizorchis* and *L.* (*Eugenuris*), populations associated with *O. princeps* and *O. collaris* will be reciprocally monophyletic, providing no additional information regarding colonization history within North America.

METHODS

Study System

There are seven major helminth lineages (at the generic to subgeneric level) associated with Nearctic pikas, with a total of ten formally described and eight potentially new species (Galbreath & Hoberg, 2012). In order to test biogeographic hypotheses of pika parasite colonization across the Bering land bridge and species diversification history in North America, I focus on helminths associated with these Nearctic pikas, specifically the pinworm subgenus *Labiostomum (Eugenuris)* and the tapeworm genus *Schizorchis.* Both parasite lineages have a total of three described species in North America (*L. (E.) talkeetnaeuris*, *S. ochotonae, S. caballeroi*), and they each have three genetically divergent and morphologicaly distinct clades that may represent additional species-level diversity. Here they are identified as species 1, 2 and 3 for each parasite lineage, as in Galbreath & Hoberg (2012).

Data collection

The majority of the parasites used in this study were from *O. princeps* collected during a survey of 37 field localities across the full geographic range of the American pika (Galbreath and Hoberg, 2012; 2015), and a smaller proportion from *O. collaris* hosts (n = 12). In addition, parasite specimens were acquired from several pika species (*O. dauurica, O. alpina, O. hyperborea,* and *O. pallasi*) collected in Mongolia during a Beringian Coevolution Project field expedition. Parasites dissected from these pikas were identified to genus level, and those identified as *Schizorchis* (n=3) and *(L.) Eugenuris* (n=4) were added to the few Palearctic specimens (from *O. cansus* and *O. hyperborea*) used by Galbreath and Hoberg (2012). A map of all collection localities is provided in Figure A1.

I extracted or acquired previously extracted DNA from an average of two specimens per locality for each lineage. All DNA was extracted using either a Qiagen DNeasy kit or Epicentre Masterpure DNA extraction kit. For *Schizorchis*, two mitochondrial markers were sequenced, including an 836 base pair (bp) region overlapping sections of the 12s and 16s ribosomal genes (rDNA, n = 154) and a 546 bp section of the cytochrome *b* (Cytb) gene (n = 82). The *L.* (*Eugenuris*) dataset included two mitochondrial DNA regions (Cytb [566 bp, n = 111] and cytochrome oxidase I [*COI*, 369 bp, n = 104]) and one autosomal DNA region, the second internal transcribed spacer between nuclear coded rRNA genes (ITS2 [774 bp, n = 115]). In addition, I acquired sequence data from GenBank for *Schizorchis* (rDNA, n = 150, HQ189777-HQ189840, KP876383-KP876464) and *L.* (*Eugenuris*) (*COI*, n = 92 parasite individuals, Genbank #s HQ189933-HQ189987, KP876217-KP876293) from previously published studies (Galbreath & Hoberg, 2012; 2015). Additional parasite sequences were acquired using polymerase chain reaction (PCR) conditions described elsewhere (Galbreath *et al.,* 2009), except for the annealing temperature (AT), which was unique for each primer set. *Schizorchis* sequences were acquired for rDNA using primers Hym16s/Hym12sR (von Nickisch-Rosenegk *et al.*, 2001) (AT = 52˚C, n = 7) and for Cytb using primers Hym01/HymLem02 (Makarikov *et al.*, 2013) (AT = 50˚C, n = 82). *Labiostomum* (*Eugenuris*) sequences for *COI* were acquired using primers BPCOX1-F1/BPCOX1-R1 (Sato *et al.*, 2005) (AT = 50˚C, n = 18). I designed Cytb (EUCYB-F/EUCYB-R; AT =

42˚C, n = 111) and ITS2 (EUITS2-F/EUITS2-R; AT = 55˚C, n = 115) primers for *L.* (*Eugenuris*) by importing annotated sequences of *Wellcomia siamensis* and *Enterobius vermicularis* from GenBank into the program Geneious 6.1 (Kearse *et al.*, 2012), performing a global alignment using ClustalW (Thompson *et al*., 1994), and then using the design primers feature (Primer 3; Untergasser *et al*., 2012) to find the best sequence locations for optimal primer effectiveness. Primer 3 does this by avoiding repetitive elements and portions of the sequence that could be chimeric. It also uses ratios of base pair richness to determine the optimal primer length for a given melting temperature (a list of all primers provided in Table A1). Vouchers and frozen tissues for parasite specimens and hosts are archived at the United States National Parasite Collection (USNPC), Museum of Southwestern Biology (MSB), and Cornell Museum of Vertebrates (CMV) (Tables B1 & B2).

PCR products were sent to ELIM Biopharmaceuticals for sequencing in both directions. Sequences were then assembled and edited in Geneious 6.1 (Kearse *et al.*, 2012), alignments were performed using the ClustalW algorithm (Thompson *et al*., 1994) implemented under default parameters. The alignments were checked by eye and indels within ITS2 were manually removed, yielding a final length for this locus of 609 bp.

Phylogenetic analysis

To evaluate phylogenetic relationships within and between Palearctic and Nearctic parasite species, I constructed both locus-specific and multi-locus phylogenetic analyses. In order to more accurately parameterize these analyses, statistical selection of the best-fit out of 66 possible substitution models of nucleotide substitution was carried out in jModeltest (Posada, 2008) using the Akaike Information Criterion (AIC) separately for each locus and species (Table A2). *Hymenolepis diminuta* (GenBank # AF314223) served as an outgroup for *Schizorchis*, and *Labiostomum* (*Labiostomum*) *rauschi* (GenBank # HQ189993) for *L.* (*Eugenuris*) in phylogenetic analyses (Galbreath & Hoberg, 2012). I constructed both maximum likelihood (ML) and Bayesian (BA) trees for each locus to compare results from these two widely used statistical methods. Once constructed, independent gene trees were then compared to assess consistency between markers and determine consistent clades within lineages. I used Garli 2.0 (Zwickl, 2006) to conduct heuristic tree searches and nonparametric bootstrap resampling under the ML criterion. Likelihood searches were run using the default parameters for run optimization, and each bootstrap pseudoreplicate was run until –ln L values converged (changing less than 0.01). PAUP* v. 4.0b10 (Swofford, 2000) was then used to calculate nodal support by computing a consensus tree of 100 bootstrap pseudoreplicates. Gene trees were also constructed under BA methods using MrBayes 3.2 (Ronquist *et al.*, 2012). I performed runs with five chains and 10 million generations, taking samples every 1000 generations, and nodal support was given as posterior probabilities (Figures B1-B5).

Differences can exist between trees generated from a single marker due to stochastic variation and non-neutral evolution. Therefore, multi-locus estimation methods are more useful for resolving true species history (Ballard & Whitlock, 2004). Five individuals were randomly selected from within each major clade and for each genetic marker (when available) for species tree generation in BEAST 2.2 (Drummond *et al.*, 2012). This program employs the coalescent-based *BEAST (Heled & Drummond, 2010) method of phylogeny estimation which permits multi-locus estimates of species trees. All coding mitochondrial genes (Cytb and *COI*) were tested using PartitionFinder v1.1.0 (Lanfear *et al.*, 2012) implemented in Python2.7 to select the best-fit codon partitioning scheme and these partitions were implemented in BEAST. In order to determine whether to implement a strict or relaxed molecular clock in BEAST, a likelihood-ratio test of clocklike evolution (Felsenstein, 1988) was performed based on likelihood scores calculated in PAUP* under a clock constraint versus no constraint. In all cases, the null model of a molecular clock could not be rejected ($P > 0.1$); therefore a strict clock was implemented. The Yule model tree prior was used for species-level coalescent simulation because it is the simplest model with a single parameter, the birth rate. This reduced the number of parameters to be estimated, which minimized bias in the absence of additional parameter information. The specific models of evolution chosen by jModeltest for the locus-specific analyses were applied to the appropriate loci by editing the BEAST XML source code. I fixed the mean substitution rate to a value of 1 so that coalescent times would be calculated in units of substitutions per site. Three independent Markov chain Monte Carlo (MCMC) simulations of 400 million steps were run for each major lineage with different starting seeds, sampling every 4000 steps. Parameter values drawn from the posterior were viewed graphically in Tracer v1.6 (Rambaut *et al.*, 2014) to ensure that parameter estimates reached stationarity, and to confirm that the three independent runs converged on similar parameter estimates. Effective sample size (ESS) values for all parameters were confirmed to exceed 200. Runs were combined using LogCombiner, after discarding the first 10% of each run as burnin, and trees were then summarized using TreeAnnotator (both part of the BEAST package). FigTree v1.4.2 (Rambaut & Drummond, 2009) was used to display the inferred species trees.

Biogeographic analysis

Historical biogeographic estimations were reconstructed for both major parasite lineages using statistical model testing with the R package BioGeoBEARS (BioGeography with Bayesian [and likelihood] Evolutionary Analysis in R Scripts) (Matzke, 2013). This package requires a well-supported phylogeny with branch lengths, and branch tips representing entire species, as opposed to individuals. Consensus trees from the multi-locus BEAST analysis provided this input. The BioGeoBEARS program compares six different models using AIC to identify which has the best statistical fit to the data. There are three base models implemented under a maximum likelihood (ML) framework, each with an additional jump dispersal parameter added (+J). The three models are, DEC (dispersal-extinction-cladogenesis), DIVALIKE (dispersal-vicariance analysis under ML), and BayAreaLIKE (Bayesian approach with a large number of areas under ML). Each of these models have different free parameters (such as dispersal, vicariance, or extinction), allowing exploration of the relative roles these biogeographic phenomena might have had in structuring the distribution of lineages over time.

BioGeoBEARS uses current ranges of species to infer biogeographic history. Each species was assigned to one of four geographic regions based on major breaks in extant parasite distributions (Palearctic [P], Alaska/Canada [C], Intermountain West [W], and American southwest [S]). Justification for splitting up regions for *O. princeps* parasites into just W and S, rather than groups representing the five host population phylogroups, is based on species distribution patterns in both lineages (Galbreath & Hoberg, 2015). In both *Schizorchis* and *Eugenuris*, a single species is associated with the southwest portion of the host’s range, and remaining species are widely distributed along the Cascade Range and Rocky Mountain cordilleras (See Figures A2 & A3 for *O. princeps* parasite distributions). Additionally, species distribution modeling of least-cost corridors has shown that species found primarily in the American southwest have likely been geographically separated from species found widely distributed along the Cascade Range and Rocky Mountain cordilleras since before the Last Glacial Maximum (LGM, ~21 thousand years ago) (Galbreath & Hoberg, 2015). BioGeoBEARS allows ancestral ranges to be distributed across multiple regions, and the allowed maximum range size is set by the user. Without additional information to inform restrictions, one less than the total number of regions is recommended, so ancestral range size at a given node was set to a maximum of three (Matzke, 2013). Scripts were run with and without a distance dispersal multiplier, where the probability of dispersal is dependent on the relative distances between range centers. The distance dispersal multiplier matrix I used is provided as Table B3, and an example R script for one parasite lineage is available as Table B4.

Finally, I tested for simultaneous divergence and colonization of closely related taxon pairs across a common barrier (the Bering Strait) by estimating the number and timing of periods of co-divergence. This was done concurrently for both parasite lineages using coalescent-based hierarchical approximate Bayesian computation methods with multi-locus DNA sequence data through the msBayes software pipeline (Hickerson *et al.* 2007). This method accounts for demographic variation among taxa and stochasticity of the coalescent process by first estimating summary statistics from the observed sequence data. After obtaining these summary statistics I simulated a large number of data sets (1,000,000) using the prior distribution for the demographic subparameters. I set the prior for the maximum number of divergence events possible equal to six, the number of lineage pairs spanning the Bering Strait (pairs listed in Table A3). Lastly, I randomly drew 1000 samples from the simulated data sets and used the acceptance/rejection algorithm to approximate the posterior distribution for the parameters. Due to a lack of robust estimates of mutation rate for either parasite lineage, I calculated relative timing of divergences in units of substitutions per site. I summarized posterior distributions using estimates of Ω [$\mathrm{var}(\tau)/E(\tau)$; Bayesian posterior credible interval] and ψ (number of divergence events) across taxon pairs. I assessed significance for the estimated number of divergences based on 95% confidence intervals of Ω. If the confidence intervals do not encompass 0 for Ω, I can reject simultaneous divergence.

RESULTS

Taxa and phylogenies

Individual gene trees built for each locus had similar tree topologies under ML and BA methods but some relationships were inconsistent between loci for both parasite groups (see Figures B1-B5). As shown previously (Galbreath and Hoberg 2012), *S.* species 2 is more closely related to the other Nearctic species than *S. mongoliensis* based on the 12s16s locus. However, addition of Cytb data indicates that *S. mongoliensis* may be more closely related to the Nearctic group, with *S.* species 2 sister to this clade. This was supported in the species tree analysis (Figure A4). For *Labiostomum (Eugenuris)*, incongruent results were evident in the placement of *L.* (*E.*) species 3 either as separate from and sister to *L.* (*E.*) *talkeetnaeuris* (in COI and Cytb) or placed within the diversity of *L.* (*E.*) *talkeetnaeuris* (ITS2). In addition, *L.* (*E.*) *talkeetnaeuris* itself was consistently found to be paraphyletic, but specimens collected from the same host species were monophyletic. These host-specific clades were treated as separate lineages in the Bayesian multi-locus analysis (Figure A5). Species trees generated via BEAST converged on the same stable tree topology that was consistent between runs and had high ESS values, indicating a well-supported species history was evident from the combined data. Surprisingly, rather than the new Palearctic parasite specimens providing an ancestral backbone to the Nearctic parasite phylogenies, the species-tree topologies revealed multiple clades with mixed Palearctic and Nearctic diversity with high nodal support (Figures A4 & A5).

Colonization history/Biogeography

For *Schizorchis*, AIC calculations indicate that the DIVALIKE+J model best fit the data for the BioGeoBEARS biogeographic analyses (weighted AIC: 0.47), which emphasizes the processes of dispersal, extinction and vicariance, as well as founder events with the added jump parameter (as opposed to range-switching and sympatry as in other models). The *L.* (*Eugenuris*) BioGeoBEARS analysis resulted in AIC calculations supporting the DEC+J model when this was run with no dispersal multipliers compared to models that were run with this added prior (weighted AIC: 0.59). This model is similar to DIVALIKE+J except that it allows for sympatry within a subset of the regions and focuses on small vicariance events instead of widespread vicariance. Ancestral range estimations based on these best-fit models are represented as pie charts on ancestral nodes and corner nodes. Sizes of slices of the pie (color-coded to different geographic regions) show ancestral range probabilities, with >50% indicating a high probability of occurrence (Matzke, 2013).

The *Schizorchis* reconstruction of historical biogeography (Figure A4) indicates a Eurasian origin (although with some ambiguity) with a deep split between two clades that each represent at least one colonization eastward into North America. The first clade includes *S.* species 1 in North America grouped with multiple Eurasian species, which would mean an eastward colonization for *S.* species 1 with eventual movement to the western mountains of the contiguous United States. The second clade includes the rest of the North American species also mixed with Eurasian diversity. Two possible biogeographic histories are possible for this group, due to some ambiguity in ancestral range estimation for the split between *S. mongoliensis* and its sister clade of Nearctic species. There is a clear eastward colonization into the Nearctic by the ancestor of *S.* species 2 and its sister clade. Following this, the ancestor to *S. mongoliensis* possibly returned westward into the Palearctic and then its sister clade colonized eastward back into the Nearctic. Alternatively, *S. mongoliensis* could have split and independently colonized westward into the Palearctic, leaving its sister clade of current Nearctic diversity in North America with *S.* species 2, which would mean one less eastward colonization compared to the alternative.

Biogeographic reconstruction for *L.* (*Eugenuris*) initially detected four eastward colonizations and one westward colonization (Figure A5). However, none of the Palearctic specimens provided an ancestral backbone to the analysis, and because we have confidence in a Palearctic origin for the lineage, there is reason to scrutinize the range reconstruction for the basal node of the tree. The placement of *L.* (*E.*) species 1 at the base of the tree influenced the analysis to place its ancestral range in the Nearctic, along with the rest of the *L.* (*Eugenuris*) diversity. However, as we move to the sister clade of *L. (E.)* species 1, the reconstructions along this branch place these ancestors all in the Palearctic. So insead of invoking a back and forth, eastward then westward, then eastward movement with the subsequent Nearctic species, a more parsimonious explanation is that these ancestors remained in the Palearctic, with separate eastward movements of *L. (E.)* species 1 and the other Nearctic species across Beringia. In order to test this more directly, I added an outgroup to the *L.* (*Eugenuris*) tree by using the same branchlenth as the *L. (E.)* species 1 split, added to the base of the tree, and assigned this outgroup to an Asian range. The resulting biogeographic reconstruction is provided in Figure A6, and although there was some ambiguity in the new range reconstructions, the same DEC+J model was selected and the highest probability of the ancestral ranges for the basal nodes did change to the Palearctic. This alternative scenario eliminates the westward colonization back into the Palearctic, but still indicates four separate eastward colonizations, with the most ancestral colonization of *L. (E.)* species 1 shifted to the node of its direct ancestor. The other three inferred colonizations remain unchanged, one leading to *L*. (*E*.) species 2, one with the ancestor of *L*. (*E*.) species 3 and *L*. (*E*.) *talkeetnaeuris* in *O. princeps*, and a third very recent one with *L*. (*E*.) *talkeetnaeuris* found in *O. collaris*.

The relative timing of inferred colonizations for both parasite lineages was assessed first by comparing the median node heights (in substitutions per site) estimated through BEAST analyses. The split between *S. yamashitai* and *S*. species 1, and between the unknown lineage from *O. pallasi* and *S*. species 2 have broadly overlapping 95% highest posterior densities (0.0435-0.1135 and 0.0483-0.1058 respectively), indicating that these two eastward colonizations could have occurred during the same glacial cycle. Branch lengths for the *L*. (*Eugenuris*) phylogeny seem to indicate that all of these colonizations occurred during different time periods; however, 95% confidence intervals on node heights did overlap between all subsequent putative colonizations. The statistical test of simultaneous divergence using MsBayes provided evidence that I can reject simultaneous divergence for all taxon pairs because 95% highest posterior density intervals for Ω (Var[t]/E[t]) did not include 0. The analysis indicated the strongest support for four discrete divergence times out of the six that were tested (Figure A7). Multiple histories could explain this result, but it provides evidence that at least two of the reconstructed colonizations occurred concurrently.

DISCUSSION

Colonization history/Biogeography

These results provide evidence to reject the hypotheses that these parasites colonized the Nearctic during a single glacial cycle, given complex relationships between multiple Nearctic and Palearctic clades that did not diverge simultaneously. This reveals that a Holarctic perspective is necessary in order to unravel the biogeographic history of pikas and their parasites across Beringia. Biogeographic and divergence time analyses most closely support the hypothesis that North American pikas have a complex history of colonization across Beringia with multiple colonizations over more than one glaciation (H1c).

However, the alternative explanation of a single large pika colonization cannot be fully ruled out. If the rate of mutation across the phylogeny has been inconsistent, then we cannot know if branch lengths across the tree are comparable. Although I could not reject clocklike evolution for either phylogeny using likelihood ratio test, much is still unknown about patterns of parasite evolution (Criscione et al., 2005). Here I provide additional evidence for and against each alternative.

A single colonization across Beringia is a more parsimonious explanation than a complex series of migrations and secondary contact between pikas across Beringia. Under a single colonization scenario, high rates of co-infection would be required in order for pikas to maintain high levels of parasite diversity as they moved across Beringia and became established in the Nearctic. Indeed, infection rates from multiple studies on

Nearctic pikas indicate high parasite loads with multiple species of the same major lineage, as well as multiple lineages in a single host (Hobbs, 1980; Galbreath & Hoberg, 2015). Co-infection rates of *L.* (*Eugenuris*) and *Schizorchis* in single individuals of *O. princeps* are about 12% (K. Galbreath, unpublished data). However, looking at co-infections within populations (localities) instead of a single individual, the rate goes up to 84%. Also, beyond a simple co-infection of one *L.* (*Eugenuris*) and one *Schizorchis*, multiple species of each major lineage are often present within a single population.

The single large colonization scenario would not necessitate an extensive history of pikas in the Nearctic (before the Pleistocene), though it would not exclude this hypothesis. Conversely, the multiple colonization hypothesis would require an extensive history for pikas in the Nearctic, allowing them to differentiate and disperse latitudinally on multiple occasions. Estimates for the timing of divergence between *O. princeps* and *O. collaris* are between 3 and 8 Mya (Lanier & Olson, 2009), which is much older than previous Pleistocene estimations. We can apply an approximate divergence estimate of 6 mya for Nearctic pikas to their *Schizorchis* parasites by converting the node age (in substitutions per site) calculated in BEAST for *Schizorchis caballeroi* found only in *O. collaris* and *S.* species 3 + *S. ochotonae* in *O. princeps*. A very rough estimate of colonization for *S.* species 1 and the ancestor of S. species 2 can then be calculated by using the node ages for their splits from Eurasian species. The diversification of *S.* species 1 from *S. yamashitai* is in this way placed at 15.5 Mya, and *S.* species 2 from its sister group at 10 Mya. These estimations place both colonizations within the mid-Miocene, when Beringian faunal migrations were thought to be symmetrical between the continents (Williams, 1998).

Fossil evidence also indicates a long and more geographically widespread history of pikas in North America than current distributions imply. Multiple species of pikas that are now extinct inhabited North America, including *O. spanglei* dated to >5.3 Mya from Oregon (Shotwell, 1956) and *O. whartoni* from Alaska assessed to be from 1.9-0.9 Mya (Repenning, 1987). Two morphological forms of *Ochotona* lived in northeastern North America during the late Pleistocene, including a larger species like *O. whartoni* and a smaller *O. collaris* or *O. princeps*-like form. Excavations from nine cave sites in the northern Appalachian Mountains in eastern North America contained *O. princeps*-like pika remains (Guilday, 1979; Churcher & Dods, 1979; Mead & Grady, 1996) that dated from ~850 kya (Repenning & Grady, 1988), and the youngest to ~30 Kya when it is presumed to have been extirpated from that area (Grady & Garton, 1981; 1982). This evidence does not exclude either hypothesis, but it does highlight that this group of hosts has a complex history in North America.

Tests for simultaneous divergence support the multiple colonization hypothesis, with biogeographic analysis providing an outline of how this movement could have proceeded if parasites did diversify within the Nearctic and then establish secondary contact with Palearctic parasites. Additionally, if we look at the genetic structure within these parasite lineages we do not see a pattern we would expect from haplotypes that have been maintained over large times and distances. In order for the genetic structure of these parasites to be unaffected by recent gene flow between populations, as well as lineage sorting, effective population sizes (N_e) would need to be extremely large. Methods for reliable estimation of N_e have not been developed for parasites (Criscione *et al*., 2005), but one genetic signal for large N_e is haplotypes that are shared at high frequencies between geographically distant and widespread locations (Watterson & Guess, 1977). *Labiostomum (E.) talkeetnaeuris* in both Nearctic *O. collaris* and Palearctic *O. hyperborea* share a single haplotype*,* but this haplotype was only found in two individuals, so was therefore not high frequency. In fact, no haplotypes could be considered high frequency. Shared haplotypes are more likely to represent recent gene flow because even though they span host boundaries, they are typically between localities within close geographic proximity. Additionally, novel haplotypes were found in localities that were unavailable for pikas to colonize during the LGM due to the presence of ice sheets (localities above the heavy dashed line in Figures 2 & 3), which is a possible indication that these parasites are able to accumulate new diversity rapidly through mutation (Galbreath & Hoberg, 2015).

Ochotona princeps/collaris ancestry

Parasite biogeography has been used as an indicator of host biogeography (Nieberding & Olivieri, 2007; Criscione *et al.*, 2005; Manter, 1966) and the use of multiple parasite lineages allows us to improve our confidence that a given pattern reflects the history of the entire assemblage rather than the parasite independently. Previously, a consistent pattern of *O. collaris* parasite diversity nested within *O. princeps* parasites was detected among five independent parasite lineages, including *Schizorchis* and *L. (Eugenuris)*, providing strong evidence that *O. collaris* is derived from *O. princeps* and moved north into Alaska and northwestern Canada from a southern ancestor (Galbreath & Hoberg, 2012). However, with the currently updated phylogenies of *Schizorchis* and *L. (Eugenuris)*, different biogeographical patterns are revealed. Under the scenario proposed by biogeographical analyses, *L. (Eugenuris) talkeetnaeuris* found in *O. collaris* may represent a recent colonizer from the Palearctic with no direct link to *O. princeps* parasites, while the origin of *O. collaris* parasite species *S. caballeroi* is more ambiguous. As previously described, *S. caballeroi* could be descendent from a pika recolonization into the Nearctic with subsequent speciation into *S. caballeroi* and sister *O. princeps* species *S.* species 3 + *S. ochotonae* without any revealed directionality. Alternatively, *S. caballeroi* could be descended from a common ancestor shared by *S.* species 2 in *O. princeps*, and as their pika hosts moved north, a subset moved across Beringia, their ancestors differentiating into *S. mongoliensis* in Siberia. Either way, with the additional diversity added to the phylogeny, a clear case of *O. princeps* parasite paraphyly is no longer consistently supported. We cannot clearly determine the diversification history of *O. collaris* and *O. princeps* using their parasites. Recent colonization of *O. collaris* parasites could suggest host switching between Eurasian and *O. collaris* pikas after they were established in the Nearctic. However, the alternate explanation of a single large pika colonization into the Nearctic would mean that the phylogenetic associations of these parasites were established before arrival in North America and genetic associations would not provide any evidence for Nearctic pika ancestry.

The geographic ranges of parasite species found in *O. princeps* do not correlate with the pika populations associated with distinct mountain ranges of the Intermountain West (Figure 2 & 3) (Galbreath & Hoberg, 2015) and the results of these biogeographic analyses could imply a different colonization history of parasites into this region than what has been suggested for their host. Lineage divergence times place *O. princeps* in the Intermountain West since the middle Pleistocene (781 – 126 Kya), with diversification proceeding from the northwest, south and then west. Divergence of the Cascade Range lineage significantly pre-dates the origin of the three eastern lineages (Galbreath *et al*., 2010; Hafner & Sullivan, 1995). Biogeographic analysis indicates that *S*. species 1, *S*. species 2 and *L*. (*E.*) species 1 have deep associations with the more southern *O. princeps* populations. This could reflect extirpation of pikas found in northern ranges or movement between major pika populations during cooler periods when their ranges expanded and connected in some regions. Movement of parasites between mountain ranges after host populations became established was likely, due to evidence of gene flow between pika populations found in nuclear DNA (Galbreath *et al*. 2009; 2010; Galbreath & Hoberg, 2015).

Taxonomy and phylogenetics

The systematics and taxonomy of helminth parasites is incomplete, and there are many challenges associated with this essential base-level work, including misidentifications, and improper documentation, preservation, storage and handling of voucher parasites (Brooks *et al.*, 2014; Hoberg *et al.*, 2009). As an example of this, the current phylogenetic analysis suggests the existence of previously unknown species-level diversity within *Schizorchis* and *L.* (*Eugenuris*). More work needs to be done in identifying many of the Eurasian parasites to species as well as revisiting the species level designations currently in place for Nearctic pinworms. With the additional loci and individuals added to the *L*. (*Eugenuris*) phylogeny, it appears that the currently designated *L*. (*Eugenuris*) *talkeetnaeuris* is paraphyletic if *L*. (*E.*) species 3 is in fact a distinct species (Figure 3). Current taxonomy of the subgenus *Labiostomum (Eugenuris*) and the putative monophyletic subfamily Labiostomatinae in lagomorph hosts is according to Petter and Quentin (1976), but rearrangement of group designations may be necessary due to the difficulty in distinguishing among closely related genera and species (eg. Akhtar, 1941, 1953; Leiby, 1961; Hobbs, 1976; Grundmann and Lombardi, 1976). We rely on catalogued voucher specimens in natural history museums as references for these morphological determinations. However, a recent examination of type specimens of Labiostomatinae pinworms revealed that multiple specimens were identified incorrectly relative to current taxonomy and there is an absence of adequate type series to document the range of morphological variation as a basis for taxonomic descriptions (Hoberg *et al.,* 2009).

CONCLUSIONS

We are just beginning to understand the taxonomic diversity and host associations of helminths, a group that constitutes multiple branches on the tree of life. It is therefore not surprising that initial hypotheses on how these organisms interact with their environment to evolve and move across the landscape were oversimplifications of a complex process. At its simplest, parasites were thought to be so attached to their hosts that their evolution would exactly track their diversification, similar to the genome of a host organelle (Nieberding & Olivieri, 2007). As evidence has grown that parasites rarely track their hosts with true fidelity, processes such as host switching and independent diversification are used to explain complex host-parasite associations. The concept of ecological fitting has recently been used to explain the patterns of major radiations of host-parasite assemblages as associated with ecological disruptions and breakdowns (Hoberg & Brooks, 2008). It is also not unreasonable to assume some level of co-diversification between hosts and their parasites during periods of ecological stability, but my study has demonstrated that there is much more we need to understand about these understudied organisms before we can clearly draw such connections. Compared to Europe and North America, the biogeographical history of Asia is poorly understood, even though it must have played a critical role in structuring high-latitude biodiversity (Tarasov *et al.*, 2000; Binney *et al.*, 2009). The addition of more parasite diversity from Asia has shown that many related lineages of these pika parasites are trans-Beringian and may reflect complex colonization patterns that are at least partially independent from each other and their host, leading us to question the underlying assumptions that allow us to use these parasites as a proxy for understanding host history. This analysis indicates the importance of more in-depth sampling throughout Asia in order to parse out complete species histories.

REFERENCES

Abbott, R.J. & Brochmann, C. (2003) History and evolution of the arctic flora: in the footsteps of Eric Hulten, *Molecular Ecology*, **12**, 299-313.

Akhtar, S.A. (1941) A new genus and species of nematodes parasitic in a pika from Afghanistan. *Records of the Indian Museum*, **43**, 217-219.

Akhtar, S.A. (1956) On the nematodes (family: Oxyuridae Cobb., subfamily Labiostominae n. subf.) parasitic in the pika of Alaska Part I. *Pakistan Journal of Scientific Research*, **8**, 133-139.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., & Johnson, E.A. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, **3**, 3376.

Ballard, J.W.O., & Whitlock, M.C. (2004) The incomplete natural history of mitochondria. *Molecular Ecology*. **13**, 729-744.

Bermingham, E. & Moritz, C. (1998) Comparative phylogeography; Concepts and applications. *Molecular Ecology*, **7**, 367-369.

Binney, H.A., Willis, K.J., Edwards, M.E., Bhagwat, S.A., Anderson, P.M., Andreev, A.A., Blaauw, M., Damblon, F., Haesaerts, P. & Kienast, F. (2009) The distribution of late Quaternary woody taxa in northern Eurasia: evidence from a new macrofossil database. *Quaternary Science Reviews*, **28**, 2445-2464.

Brant, S.V. & Gardner, S.L. (2000) Phylogeny of species of the genus *Litomosoides* (Nematoda; Onchocercidae): evidence of rampant host switching. *Journal of Parasitology*, **86**, 545-554.

Brooks, D.R. (1979) Testing the context and extent of host-parasite coevolution. *Systematic Zoology*, **28**, 299-307.

Brooks, D.R., & Ferrao, A.L. (2005) The historical biogeography of co-evolution: emerging infectious diseases are evolutionary accidents waiting to happen. *Journal of Biogeography,* **32**, 1291-1299.

Brooks, D.R., & Hoberg, E.P. (2006) Systematics and emerging infectious diseases: From management to solution. *Journal of Parasitology,* **92**, 426-429.

Brooks, D.R., & Hoberg, E.P. (2007) How will global climate change affect parasite-host assemblages? *Trends in Parasitology,* **23**, 571-574.

Brooks, D.R., Hoberg, E.P., Boeger, W.A., Gardner, S.L., Galbreath, K.E., Herczeg, D., Mejia-Madrid, H.H., Rácz, S.E., & Dursahinhan, A.T. (2014) Finding them before they find us: informatics, parasites, and environments in accelerating climate change. *Comparitive Parasitology*, **81**, 155-164.

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.

Churcher, C.S. & Dods, R.R. (1979) *Ochotona* and other vertebrates of possible Illinoian age from Kelso Cave, Halton County, Ontario. *Canadian Journal of Earth Sciences*, **16**, 1613-1620.

Cook, J.A., Hoberg, E.P., Koehler, A., Henttonen, H., Wicks-tr€om, L., Haukisalmi, V., Galbreath, K., Chernyavski, F., Dokuchaev, N., Lahzuhtkin, A., MacDonald, S.O., Hope, A., Waltari, E., Runck, A., Veitch, A., Popko, R., Jenkins, E., Kutz, S. & 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.

Criscione, C.D., Poulin, R. & Blouin, M.S. (2005) Molecular ecology of parasites: elucidating ecological and microevolutionary processes. *Molecular Ecology*, **14**, 2247-2257.

Dawson, N.G., Hope, A.G., Talbot, S.L. & Cook, J.A. (2013) A multilocus evaluation of ermine (*Mustela ermine*) across the Holarctic, testing hypotheses of Pleistocene diversification in response to climate change. *Journal of Biogeography*, **41**, 464-475.

DeChaine, E.G. (2008) A bridge or a barrier? Beringia's influence on the distribution and diversity of tundraplants. *Plant Ecology and Diversity*, **1**, 197–207.

Denver, D.R., Morris, K., Lynch, M. Vassilieva, L.L., & Thomas, W.K. (2000) High direct estimate of the mutation rate in the mitochondrial genome of *Caenorhabditis elegans. Science*, **289**, 2342-2344.

Dobson, A., Lafferty, K.D., Kuris, A.M., Hechinger, R.F., & Jetz, W. (2008) Homage to Linnaeus: How many parasites? How many hosts? *Proceedings of the National Academy of Sciences of the United States of America,* **105**, 11482-11489.

Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. (2012) Bayesian phylogenetics with BEAUTi and the BEAST 1.7. *Molecular Biology and Evolution,* **29**, 1969-1973.

Dyke, A.S. & Prest, V.K. (1987) Late Wisconsinan and Holocene history of the laurentide ice sheet. *Gèographie Physique et Quaternaire*, **42**, 237-263.

Edwards, M.E., Brubaker, L.B., Lozhkin, A.V., & Anderson, P.M. (2005) Structurally novel biomes: a response to past warmings in Beringia. *Ecology*, **86**, 1696-1703.

Elias, S.A., Short, S.K., Nelson, C.H., & Birks, H.H. (1996) Life and times of the Bering land bridge. *Nature*, **382**, 60-62.

Emerson, K.J., Merz, C.R., Catchen, J.M., Hohenlohe, P.A., Cresko, W.A., Bradshaw, W.E., & Holzapfel, C.M. (2010) Resolving postglacial phylogeography using high-throughput sequencing. *PNAS,* **107**, 16196–16200.

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

Formozov, N.A., Grigor'eva, T.V. and Surin, V.L. (2006) Molecular systematics of pikas of the subgenus *Pika (Ochotona*, Lagomorpha). *Zoologicheskii zhurnal*, **85**, 1465-1473. (in Russian with English summary)

Galbreath, K.E. (2009) Episodic host-switching drives parasite diversification in a pika-parasite assemblage. *Doctoral Dissertation, Cornell University*, 93-135.

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., Hafner, D.J. & Zamudio, K.R. (2009) When cold is better: climate-driven elevation shifts yield complex patterns of diversification and demography in an alpine specialist (American pika, *Ochotona princeps*). *Evolution*, **63**, 2848-2863.

Galbreath, K.E., Hafner, D.J., Zamudio, K.R., & Agnew, K. (2010) Isolation and introgression in the Intermountain West: contrasting gene genealogies reveal the complex biogeographic history of the American pika (Ochotona princeps). *Journal of Biogeography*, **37**, 344–362.

Galbreath, K.E., Cook, J.A., Eddingsaas, A.A., & DeChaine, E.G. (2011) Diversity and demography in Beringia: multilocus tests of paeodistribution 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 of Biological Science*, **279**, 371–378.

Galbreath, E.G. & Hoberg, E.P. (2015) Host responses to cycles of climate change shape parasite diversity across North America's Intermountain West. *Folia Zoologica*, **64**, 218-232.

Grady, F., & Garton, E.R. (1981) The collared lemming *Dicrostonyx hudsonius* (Pallas) from a Pleistocene cave deposit in West Virginia. *Proceedings of the Eight International Congress of Speleology*, **1/2**, 279-281.

Grady, F. & Garton, E.R. (1982) Pleistocene fauna from New Trout Cave. *Capitol Area Cavers Bulletin*, **1**, 62-69.

Greenspoon, P.B. & M'Gonigle, L.K. (2013) The evolution of mutation rate in an antagonistic coevolutionary model with maternal transmission of parasites. *Proceedings of the Royal Society Biology*, **280**, 20130647.

Guilday, J.E. (1979) Eastern North American Pleistocene *Ochotona* (Lagomorpha: Mammalia). *Annals of Carnegie Museum*, **48**, 435-444.

Grundmann, A.W., & Lombardi, P.S. (1976) Parasitism of the pika, *Ochotona princeps* Richardson (Mammalia, Lagomorpha), in Utah and Nevada with description of *Eugenuris utahensis* sp. n. (Nematoda, Oxyuridae). *Proceedings of the Helminthological Society of Washington*, **43**, 39-46.

Guan, J. & Lin, Y. (1988) Life cycle of *Schizorchis altaica* Gvozdev (Cestoda: Anoplocephalidae). *Journal of Xiamen University Natural Science*, **27**, 709–713.

Guthrie, R.D. (1973) Mummified pika (*Ochotona*) carcass and dung pellets from pleistocene deposits in interior Alaska. J*ournal of Mammalogy*, **54**, 970-971.

Hafner, D.J. & Sullivan, R.M. (1995) Historical and ecological biogeography of Nearctic pikas (Lagomorpha, Ochotonidae). *Journal of Mammalogy*, **76**, 302-321.

Hafner, M. S. & Page, R.D. (1995) Molecular phylogenies and host-parasite cospeciation - gophers and lice as a model system. *Philosophical Transactions of the Royal Society of London Biology*, **349**, 77-83.

Haukisalmi, V., Hardman, L.M., Fedorov, V.B., Hoberg, E.P. & Henttonen, H. (2015) 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.

Heled, J. & Drummond, A.J. (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, **27**, 570-580.

Hobbs, R.P. (1976) A revision of the genus *Cephaluris* Akhtar, 1947 (Nematoda: Oxyuridae) with redescriptions of the North American species. *Proceedings of the Helminthological Society of Washington*, **43**, 151-160.

Hobbs, R.P. (1980) Interspecific interactions among gastrointestinal helminths in pikas of North America. *American Midland Naturalist*, **103**, 15-25.

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. (2013) Episodic processes, invasion, and faunal mosaics in evolutionary and ecological time. In: Rohde K. (ed.), *The balance of nature and human impact*, pp. 199-213. Cambridge 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 of London. Series B*, **370**, 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. In: Rollinson D. and Hay S. I. (eds.). *Advances in parasitology*, pp. 1-97. Academic Press.

Hoberg, E.P., Pilitt, P.A., Galbreath, K.E. (2009) Why museums matter: A tale of pinworms (Oxyuroidae: Heteroxynematidae) among pikas (*Ochotona princeps* and *O. collaris*) in the American west. *Journal of Parasitology*, **95,** 490-501.

Hoffmann, R.S. & Smith, A.T. (2005) Lagomorpha. *Mammal species of the world* (ed. by D. E. Wison and D. M. Reeder), pp. 185-211. Johns Hopkins University Press, Baltimore, MD.

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.

Hopkins, D.M. (1959) Cenozoic history of the Bering land bridge. *Science*, **129**, 1519-1528.

Huang, W., Takebayashi, N., Qi, Y. & Hickerson, M.J. (2011) MTML-msBayes: Approximate Bayesian comparative phylogeographic inference from multiple taxa and multiple loci with rate heterogeneity. *BMC Bioinformatics*, **12**: 1.

Huyse, T. & Volckaert, F.A.M. (2005) Comparing host and parasite phylogenies: *Gyrodactylus* flatworms jumping from goby to goby. *Systematic Biology*, **54**, 710-718.

Johnson, K.P., Adams, R.J., & Clayton, D.H. (2002) The phylogeny of the louse genus *Bruellia* does not reflect host phylogeny. *Biological Journal of the Linnean Society*, **77**, 233-247.

Johnson, K.P., Adams, R.J., Page, R.D.M., & Clayton, D.H. (2003) When do parasites fail to speciate in response to host speciation? *Systematic Biology*, **52**, 37-47.

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**, 1647-1649.

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.

Lanier, H.C. & Olson, L.E. (2009) Inferring divergence times within pikas (*Ochotona* spp.) using mtDNA and relaxed molecular dating techniques. *Molecular Phylogenetics and Evolution*, **53**, 1–12.

Leaché, A.D. & Rannala, B. (2011) The accuracy of species tree estimation under simulation: a comparison of methods. *Systematic Biology*, **60**, 126-137.

Lehtonen, J. & Lanfear, R. (2014) Generation time, life history and the substitution rate of neutral mutations. *Biology Letters*, **10**, 20140801.

Leiby, P.D. (1961) *Labiostomum coloradensis* n. sp. (Oxyuridae: Nematoda) from the pika *Ochotona princeps*. *Journal of Parasitology*, **47**, 733-734.

Li, W-H., Ellsworth, D.L., Krushkal, J., Chang, B.H.J., Hewett-Emmett, D. (1996) Rates of nucleotide substitution in primates and rodents and the generation-time effect hypothesis. *Molecular Phylogenetic Evolution*, **5**, 182-187.

Li, M., Shi, S.F., Brown, C.L., & Yang, T.B. (2011) Phylogeographical pattern of *Mazocraeoides gonialosae* (Monogenea, Mazocraeidae) on the dotted gizzard shad, *Konosirus punctatus*, along the cost of China. *The international Journal for Parasitology* **41**, 1263-1272.

Librado, P. & Rozas, J. (2009) DnaSP v5: A software for comprehensive analysis of DNA
polymorphism data. *Bioinformatics,* **25**, 1451-1452.

Lissovsky, A.A., Ivanova, N.V. & Borisenko, A.V. (2007) Molecular phylogenetics and taxonomy of the subgenus Pika (*Ochotona*, Lagomorpha). *Journal of Mammalogy*, **88**, 1195–1204.

MacArthur, R.A. & Wang, L.C.H. (1974) Behavioral thermoregulation in the pika *Ochotona princeps*: a field study using radiotelemetry. *Canadian Journal of Zoology*, **52**, 353-358.

Manter, H.W. (1966) Parasites of fishes as biological indicators of recent and ancient conditions. In *Host-parasite relationships* (ed. J.E. McCauley), pp. 59-71. Corvallis, OR: Oregon State University Press.

Matzke, N.J. (2013) Probabilistic Historical Biogeography: New Models for Founder-Event Speciation, Imperfect Detection, and Fossils Allow Improved Accuracy and Model-Testing. *Frontiers of Biogeography*, **5**, 242-248.

Marlowe, K. & Hufford, L. (2008) Evolution of *Synthyris* sect. *Dissecta* (Plantiginaceae) on sky islands in the Northern Rocky Mountains. *American Journal of Botany*, **95**, 381-392.

Martin, A.P. & Palumbi, S.R. (1993) Body size, metabolic rate, generation time and the molecular clock. *Proceedings of the National Academy of Sciences USA*, **90**, 4087-4091.

McCormack, J.E., Huang, H., & Knowles, L.L. (2009) Maximum likelihood estimates of species trees: how accuracy of phylogenetic inference depends upon the divergence history and sampling design. *Systematic Biology,* **58**, 501-508.

Mead, J.I. & Grady, F. (1996) *Ochotona* (Lagomorpha) from Late Quaternary cave deposits in eastern North America. *Quaternary Research*, **45**, 93-101.

Nieberding, C.M., & Olivieri, I. (2007) Parasites: proxies for host genealogy and ecology. *Trends in Ecology & Evolution*, **22**, 156-165.

Niu, Y.D., Wei, F.W., Li, M., Liu X.M. & Feng Z.J. (2004) Phylogeny of pikas (Lagomorpha, *Ochotona*) inferred from mitochondrial cytochrome *b* sequences. *Folia Zoologica*, 53, 141-155.

Petter, A.J. & Quentin, J.C. (1976) Keys to the genera of the Oxyuroidae, Part 4. *In* CIH keys to the nematode parasites of vertebrates, Anderson, R.C., Chabaud, A.G., & Willmot, S. (eds.). *Commonwealth Agricultural Bureaux,* Rarnham Royal, England, p. 1-30.

Pollak, E. (1982) The rate of mutant substitution in populations with overlapping generations. *Genetic Research*, **40**, 89-94.

Posada, D. (2008) JModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 141-155.

Poulin, R., & Morand, S. (2000) The diversity of parasites. *Quarterly Review of Biology*, **75**, 277-293.

Rambaut, A., & A. Drummond. (2009) FigTree v1. 3.1 <http://tree.bio.ed.ac.uk/software/figtree>.

Rambaut, A., Suchard, M.A., Xie, D. & Drummond, A.J. (2014) Tracer v1.6 <http://beast.bio.ed.ac.uk/Tracer>.

Rausch, V.R. & Ritter, D.G. (1973) Somatic chromosomes of a male collared pika, *Ochotona collaris* (Nelson) (Lagomorpha: Ochotonidae). *Mammal Chromosome Newsletter*, **14**, 109-111.

Rausch, R.L. & Smirnova, L.V. (1984) The genus *Schizorchis* (Cestoda, Anoplocephalidae) in Eurasian pikas, *Ochotona* spp. (Lagomorpha, Ochotonidae), with descriptions of new species. *Transactions of the American Microscopical Society*, **103**, 144-156.

Repenning, C.A. (1987) Biochronology of the microtine rodents of the United States. *In* Cenozoic Mammals of North America. Geochronology and Biochronology (M.O. Woodburne, Ed.), pp. 236-268. University of California Press, Berkeley.

Repenning, C.A. & Grady, F. (1988) The microtine rodents of the Cheetah Room fauna, Hamilton Cave, West Virginia, and the spontaneous origin of *Synaptomys. U. S. Geological Survey Bulletin*, **1853**, 1-32.

Roberts, B.L. (1991) Modeling the Cordilleran ice-sheet. *Gèographie Physique et Quaternaire*, **45**, 287-299.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Hohna, S., Larget, B., Liu, L., & Suchard, M.A. (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**, 1-4.

Sato, H., Une, Y., Kawakami, S., Saito, E., Kamiya, H., Akao, N. & Furuoka, H. (2005) Fatal *Baylisascaris* larva migrans in a colony of Japanese macaques kept by a safari-style zoo in Japan. *Journal of Parasitology*, **91**, 716-719.

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

Shotwell, J.A. (1956) Hemphillian mammalian assemlage from northeastern Oregon. *Geological Society of America Bulletin*, **67**, 717-738.

Smith, A.T. (1974) The distribution and dispersal of pikas: influences of behavior and climate. *Ecology,* **55**, 1368–1376.

Spellmen, G.M., Riddle, B. & Klicka, J. (2007) Phylogeography of the mountain chickadee (*Poecile gambeli*): diversification, introgression, and expansion in response to Quaternary climate change. *Molecular Ecology*, **16**, 1055-1068.

Stapp, P. & Salkeld, D.J. (2009) Inferring host-parasite relationships using stable isotopes: implications for disease transmission and host specificity. *Ecology*, **90**, 3268–3273.

Swofford, D.L. (2000) *PAUP**: Phylogenetic analysis using parsimony (*and other methods, Version 4.0. Sunderland, MA: *Sinauer Associates*.

Tarsov, P.E., Volkova, V.S., Webb, T., Guiot, J., Andreev, A.A., Bezusko, L.G., Bezusko, T.V., Bykova, G.V., Dorofeyuk, N.I., Kvavadze, E.V. Osipova, I.M. & Panova, N.K. (2000) Last glacial maximum biomes reconstructed from pollen and plant macrofossil data from northern Eurasia. *Journal of Biogeography*, **27**, 609-620.

Thomas, J.A., Welch, J.J., Lanfear, R., Bromham, L. (2010) A generation time effect on the rate of molecular evolution in invertebrates. *Molecular Biology and Evolution*, **27**, 1173-1180.

Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22,** 4673-4680.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. & Rozen, S.G. (2012) Primer3--new capabilities and interfaces. *Nucleic Acids Research*, **40**, e115.

von Nickisch-Rosenegk, M., Brown, W.M., & Boore, J.L. (2001) Complete sequence of the mitochondrial genome of the tapeworm *hymenolepis diminuta*: gene arrangements indicate that platyhelminths are eutrochozoans. *Molecular Biology and Evolution*, **18**, 721-730.

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

Watterson, G.A. & Guss, H.A. (1977) Is the most frequent allele the oldest? *Theoretical Population Biology*, **11**, 141-160.

Wickström, L.M., Haukisalmi, V., Varis, S., Hantula, J., Fedorov, V.B. & Henttonen, H. (2003) Phylogeography of the circumpolar *Paranoplocephala arctica* species complex (Cestoda: Anoplocephalidae) parasitizing collard lemmings (*Dicrostonyx* spp.). *Molecular Ecology*, **12**, 3359-3371.

Williams, M., Dunkerley, D., De Deckker, P. Kershaw, P. & Chappell, J. (1998) Quaternary Environments Second Addition. *Oxford University Press Inc*. pp. 107-122.

Zarlenga, D.S., Rosenthal, B.M., La Rosa, G., Pozio, E. & Hoberg, E.P. (2006) Post-Miocene expansion, colonization, and host switching drove speciation among extant nematodes of the archaic genus *Trichinella*. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 7354-7359.

Zwickl, D. J. (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin.

APPENDIX A

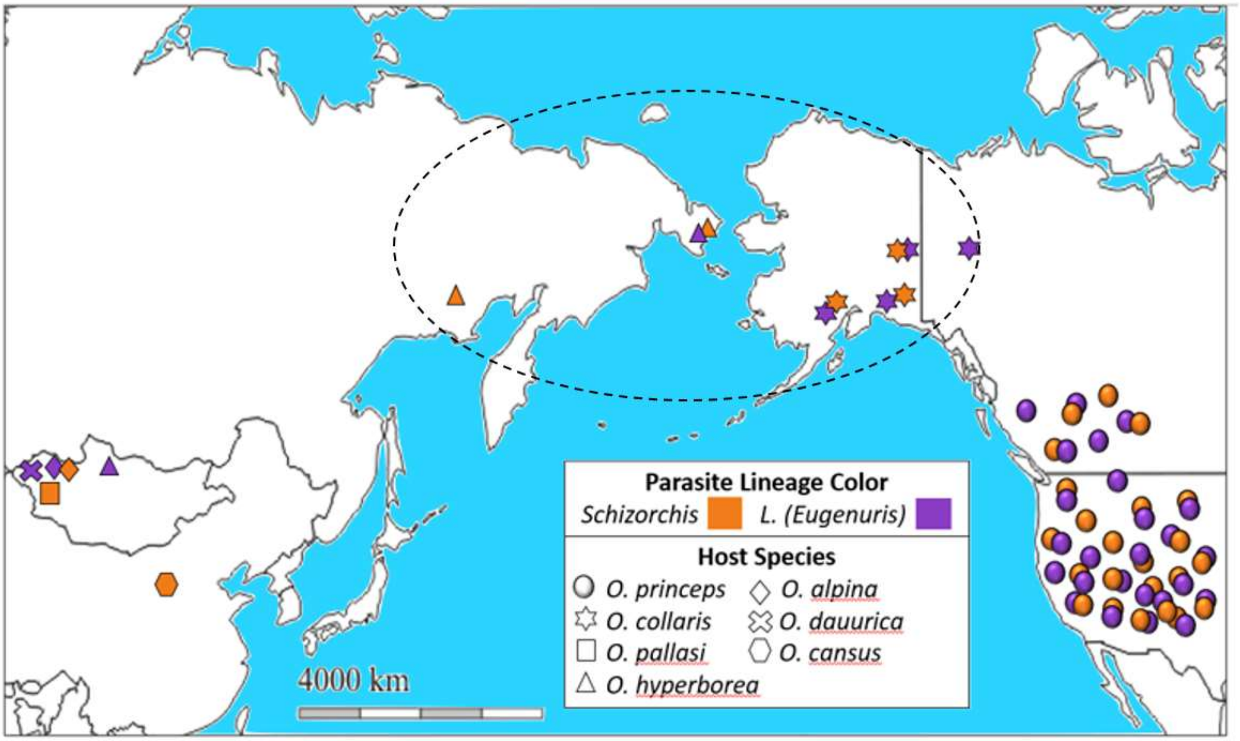


Figure A1. Range map of *Schizorchis* and *L*. (*Eugenuris*) collection localities used in this study. Beringia is the area generally within the dashed oval.

Table A1. Primers used for PCR.

PCR Primers		
Schizorchis	**Primer Name**	**Sequence**
Cytb	HYM01	5' ATTGTGGTTYTGTTGAATAC
	HYMLEM02	5' CCCACAATAGCAAAYCCCAARCATACATG
12S16S	HYM16SF	5' TTATAAATGGCCGTATATTGAC
	HYM12SR	5' ATCGTCCTTTATAACACACCTTCCC
Eugenuris		
ITS2	EUITS2-F	5' CGAGTATCGATGAAGAACGCAGC
	EUITS2-R	5'ATATGCTTAAGTTCAGCGGG
Cytb	EUCYB-F	5' TAATTTTGAAGAGATTGAG
	EUCYB-R	5' AATCTTRTCATAATCACCATGACA
CO1	BPCOX1-F1	5' TTTGGTCATCCTGAGGTTTATATT
	BPCOX1-R1	5' ATGAAAATGTCTAACTACATAATAAGTATC

Table A2. Loci sample counts and model selection information for each Nearctic parasite species. N = number of sequenced individuals used and Model (AIC) the nucleotide substitution model chosen by jModelTest under the Akaike information criterion used in phylogenetic analyses. Results from PartitionFinder are given below each coding locus, indicating how they were partitioned in analyses.

L. (Eugenuris)					
Locus	**species**	**# base pairs**	**N**	**N_H**	**Model (AIC)**
ITS2	All	609	115	23	TIM1+G
	talkeetnaeuris		66	13	HKY
	sp. 1		14	6	F81
	sp. 2		6	2	F81
	sp. 3		12	4	TPM3uf
Cox1	All	369	104	38	TIM1+G
	talkeetnaeuris		68	21	HKY+I
(partitioned by all 3 codons)	sp. 1		9	5	TrN
	sp. 2		6	5	TrN+I
	sp. 3		11	7	TrN
Cytb	All	566	111	51	TPM3uf+G
(partit. 2	talkeetnaeuris		77	16	TPM3uf+I
subsets, codons	Sp. 2		3	3	TPM1uf
1&2, 3 sep.)	Sp. 3		24	13	TPM1uf
Schizorchis					
Locus	**species**	**length**	**N**	**N_H**	**Model (AIC)**
12s16s	all	836	154	56	TPM3uf+G
	ochotonae		94	36	TPM3uf+G
	sp. 1		9	4	TPM3uf+G
	sp. 2		19	9	TIM3+G
	sp. 3		13	4	HKY
	caballeroi		8	6	TrN
Cytb	all	546	82	38	TIM3+I+G
	ochotonae		45	27	TIM3+G
(partit. 2	sp. 1		7	6	TPM3uf+I
subsets, codons	sp. 2		5	4	TPM2uf+G
1&2, 3 sep.)	sp. 3		3	2	TPM2uf
	caballeroi		6	5	TPM3uf

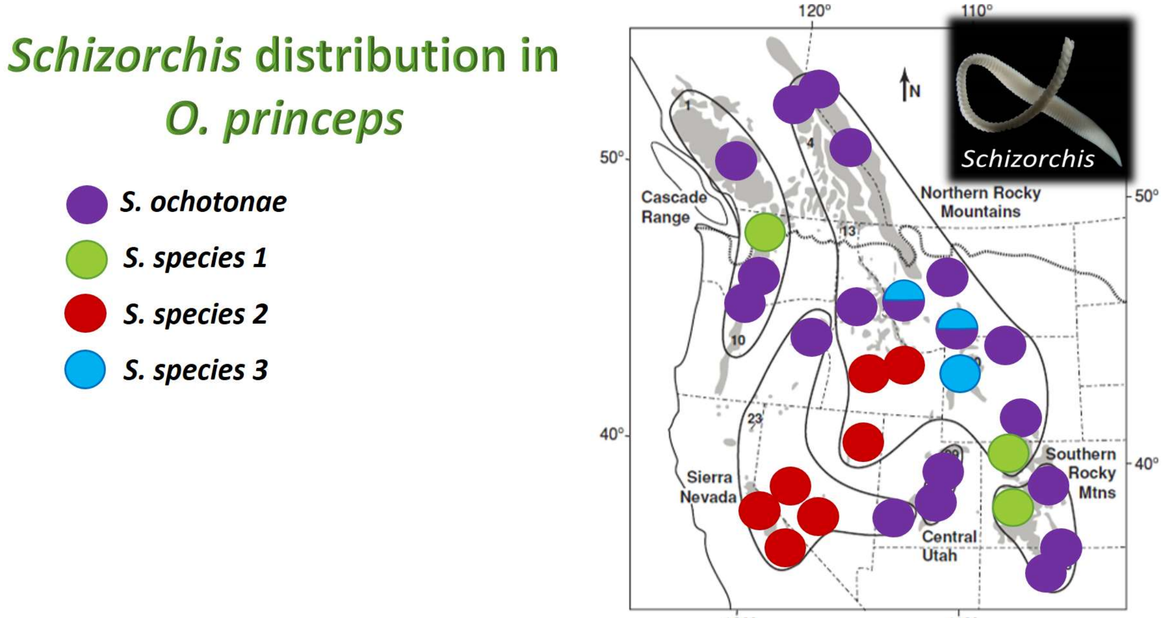


Figure A2. Distribution of *Schizorchis* species within the range of its *O. princeps* host. The red southwestern species (*S.* species 2) corresponds to the southwestern range in BioGeoBEARS, and the other three correspond to the American mountain range.

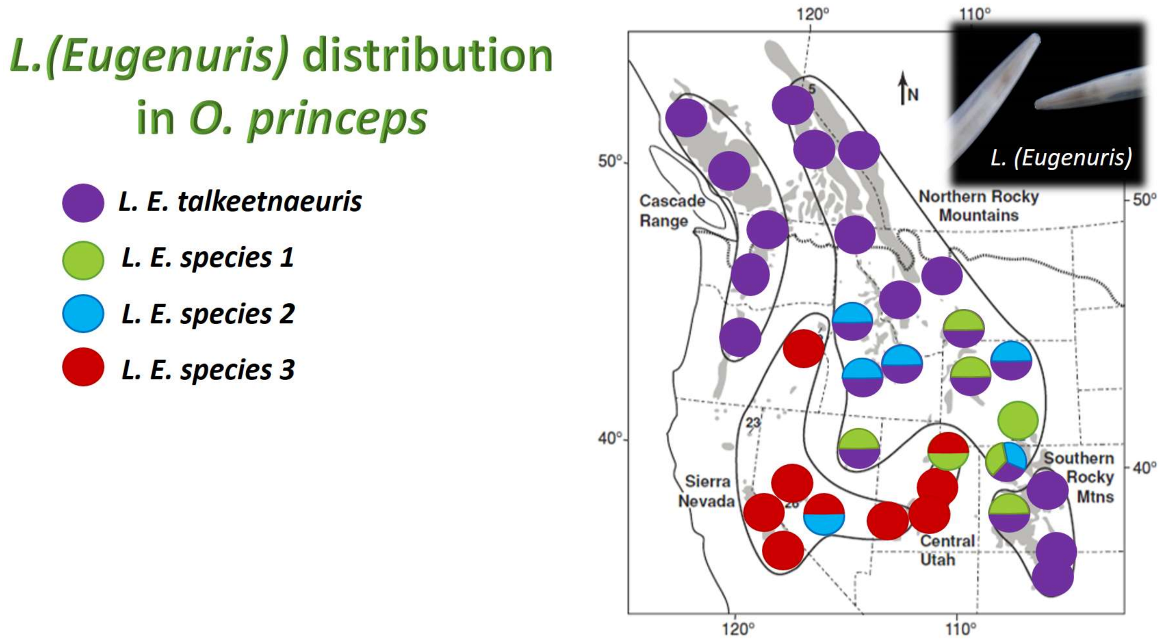


Figure A3. Distribution of *L*. (*Eugenuris*) species within the range of its *O. princeps* host. The red southwestern species (*L*. (*E*.) species 3) corresponds to the southwestern range in BioGeoBEARS, and the other three correspond to the American mountain range.

Table A3: Taxon pairs used for msBayes analysis. Each pair consists of two monophyletic clades, each containing one or more species, one located only in the Palearctic and the other only in the Nearctic.

	Nearctic Clade	Palearctic Clade
Taxon pair 1	*L. (Eugenuris) talkeetnaeuris* and species 3 from *O. princeps* hosts	Unknown species of *L.* (*Eugenuris*) from *O. hyperborea* and *O. alpina* hosts
Taxon pair 2	*L. (Eugenuris) talkeetnaeuris* from *O. collaris* hosts	*L. (Eugenuris) talkeetnaeuris* from *O. hyperborea* hosts
Taxon pair 3	*L. (Eugenuris)* species 2 from *O. princeps* hosts	Unknown species of *L.* (*Eugenuris*) from *O. dauurica* hosts
Taxon pair 4	*Schizorchis caballeroi* from *O. collaris* hosts	*Schizorchis mongoliensis* from *O. hyperborea* hosts
Taxon pair 5	*Schizorchis* species 2 from *O. princeps* hosts	Unknown species of *Schizorchis* from *O. pallasi* hosts
Taxon pair 6	*Schizorchis* species 1 from *O. princeps* hosts	*Schizorchis yamashitai* from *O. cansus* hosts

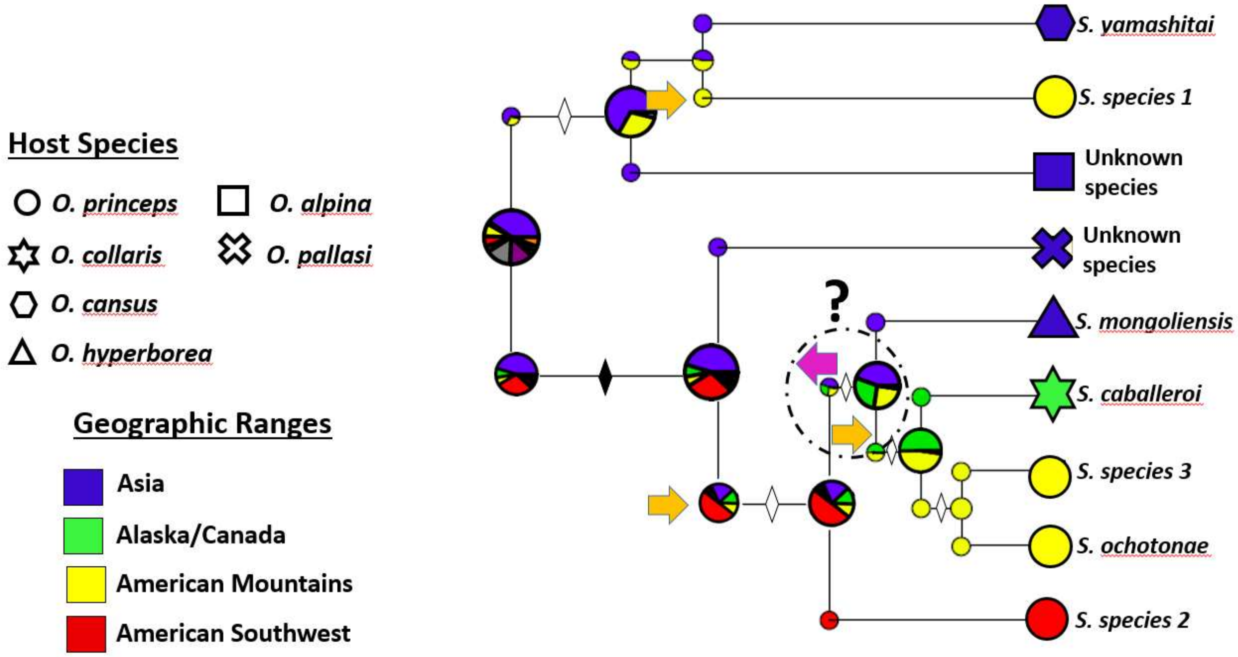


Figure A4. *Schizorchis* BioGeoBEARS biogeographic reconstruction using the BEAST species tree backbone. Open diamonds along branches indicate Bayesian posterior probabilities >.65 and filled diamonds >.95 for subsequent nodal support. Shapes on branch tips indicate host species and color indicates known current geographic ranges. Pie charts on nodes are the probability of ancestral range location/s and directional arrows indicate inferred cross continental colonization across the Bering land bridge either east into North America (orange arrows) or west into Eurasia (pink arrow). The dashed circle surrounding colonization arrows and node indicate uncertainty in range estimation and therefore history of colonization. Individuals from Mongolian pikas identified to genus are indicated as unknown species and S. species 1, 2 and 3 are named as in Galbreath and Hoberg (2012).

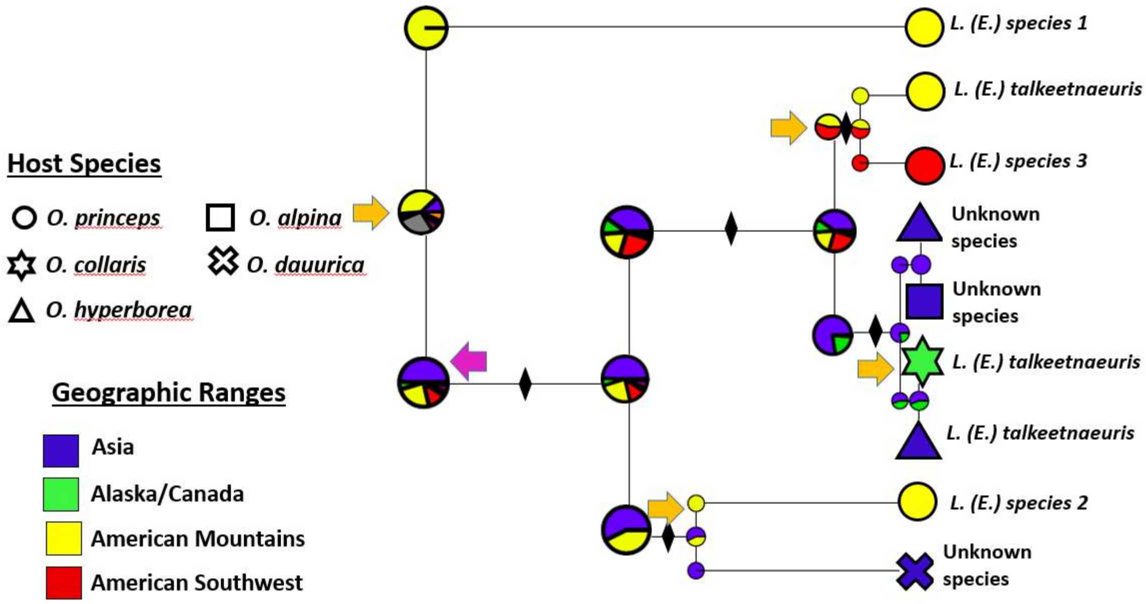


Figure A5. *L*. (*Eugenuris*) BioGeoBEARS biogeographic reconstruction using the BEAST species tree backbone. Open diamonds along branches indicate Bayesian posterior probabilities >.80 and filled diamonds >.95 for subsequent nodal support. Shapes on branch tips indicate host species and color indicates known current geographic ranges. Pie charts on nodes are the probability of ancestral range location/s and directional arrows indicate inferred cross continental colonization across the Bering land bridge either east into North America (orange arrows) or west into Eurasia (pink arrow). Individuals from Mongolian pikas identified to genus are indicated as unknown species, and L. E. species 1, 2 and 3 are named as in Galbreath and Hoberg (2012).

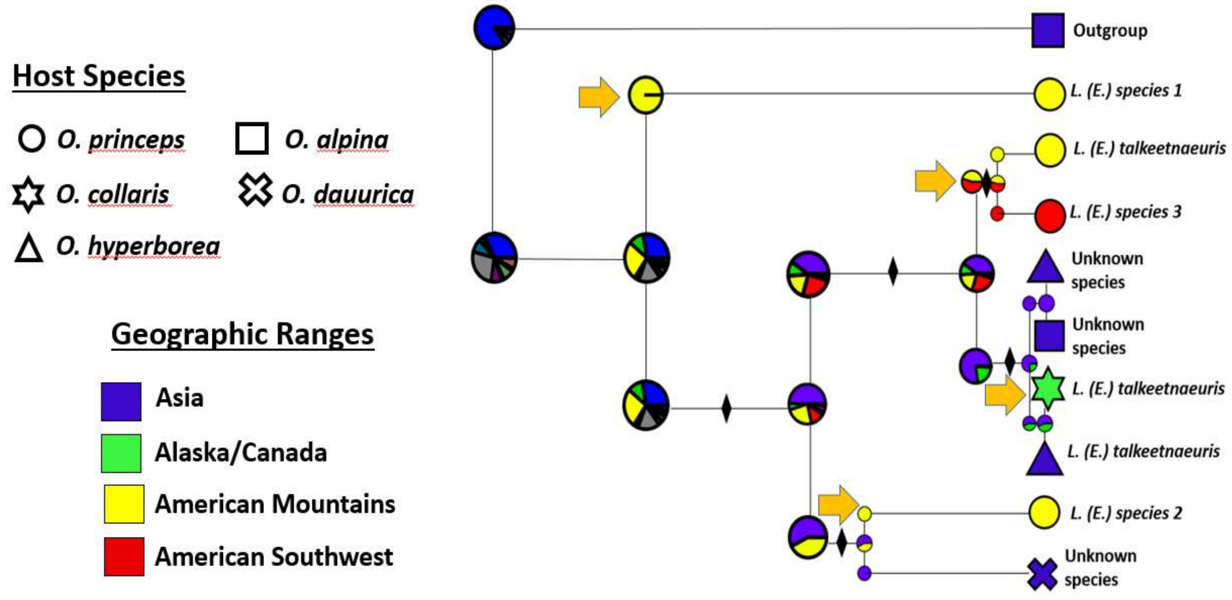


Figure A6. Same *L*. (*Eugenuris*) BioGeoBEARS biogeographic reconstruction as in Figure A5 but with an added simulated outgroup from Asia, which reflects the know origin for the lineage. Updated Beringian colonization events are inferred, with directional arrows now all indicating eastward colonization into North America (orange arrows).

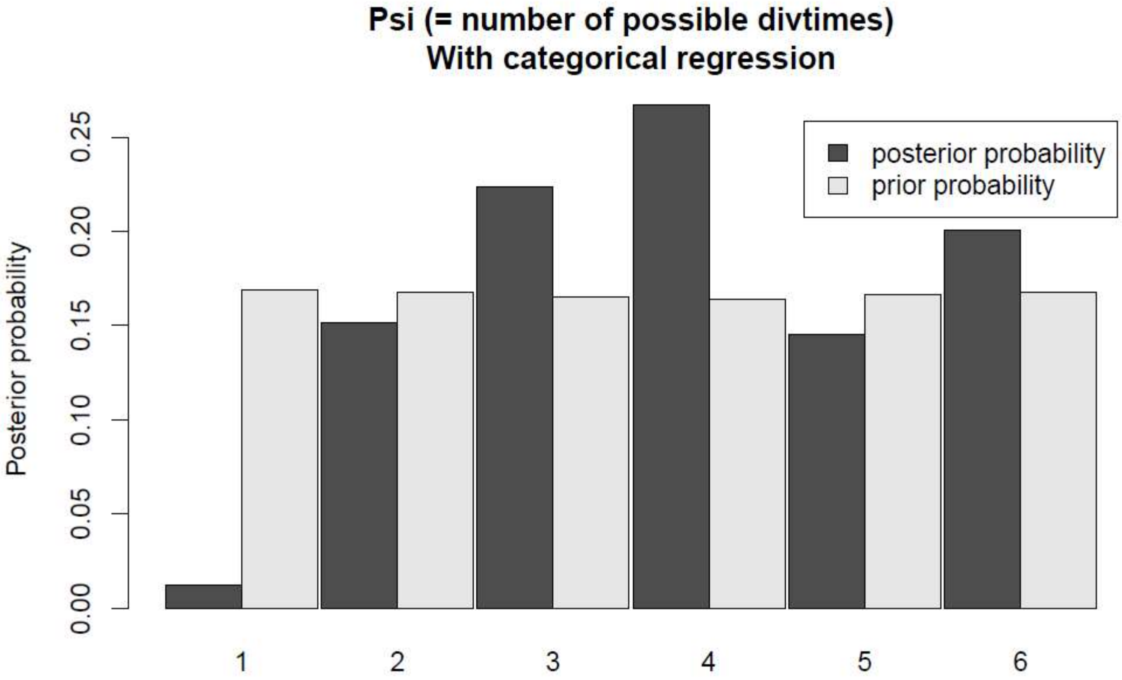


Figure A7. Posterior probability comparisons between the number of possible divergence times of taxon pairs using Msbayes.

APPENDIX B

Table B1: *Schizorchis* specimen information list. Museum accession numbers for parasites and associated hosts as well as GenBank numbers for sequenced loci.

Host Field ID	**Host Archive (CMV/MSB)**	**Schiz ID**	**Species**	**16sF/12sR**	**CytB**	**GenBank #**	**Parasite Accession**	**Locality**
KG 103	20116		O. princeps					USA,NM,TAOS CO. CARSON NATL FOR, 2.5 KM NW TWINING
		103c1	S. ochotonae	X	X	HQ189787.1		
		103c2	S. ochotonae	X	X	KP876383.1		
		103c3	S. ochotonae	X		KP876384.1		
		103c4	S. ochotonae	X	X	KP876385.1		
		103c5	S. ochotonae	X		KP876386.1		
		103c6	S. ochotonae	X		KP876387.1		
		103c7	S. ochotonae	X		KP876388.1		
		103c8	S. ochotonae	X		KP876389.1		
KG 111	20124		O. princeps					USA,NM,TAOS CO. CARSON NATL FOR, 2.5 KM NW TWINING
		111c1	S. ochotonae	X	X	HQ189788.1		
		111c2	S. ochotonae	X		KP876390.1		
		111c3	S. ochotonae	X		KP876391.1		
		111c4	S. ochotonae	X	X	KP876392.1		
		111c5	S. ochotonae	X		KP876393.1		
		111c6	S. ochotonae	X		KP876394.1		

		111c7	S. ochotonae	X	X	KP876395.1		
		111c8	S. ochotonae	X		KP876396.1		
KG 114	20125		O. princeps					USA,NM,TAOS CO. CARSON NATL FOR, 2.5 KM NW TWINING
		114c1	S. ochotonae	X		HQ189789.1		
		114c2	S. ochotonae	X	X	HQ189790.1		
		114c3	S. ochotonae	X		KP876397.1		
		114c4	S. ochotonae	X		KP876398.1		
		114c5	S. ochotonae	X		KP876399.1		
		114c6	S. ochotonae	X		KP876400.1		
		114c7	S. ochotonae	X		KP876401.1		
		114c8	S. ochotonae	X		KP876402.1		
		114c9	S. ochotonae	X		KP876403.1		
		114c10	S. ochotonae	X		KP876404.1		
		114c11	S. ochotonae	X		KP876405.1		
		114c12	S. ochotonae	X		KP876406.1		
		114c13	S. ochotonae	X		KP876407.1		
		114c14	S. ochotonae	X		KP876408.1		
		114c15	S. ochotonae	X	X	KP876409.1		
		114c16	S. ochotonae	X		KP876410.1		
KG 128	20137		O. princeps					USA,CO,CLEAR CK CO. ARAPAHOE NATL FOR, W. SIDE BERTHOUD PASS
		128c1	S. ochotonae	X		KP876411.1		

		128c2	S. ochotonae	X		KP876412.1		
		128c3	S. ochotonae	X		KP876413.1		
		128c4	S. ochotonae	X	X	KP876414.1		
		128c5	S. ochotonae	X		HQ189791.1		
		128c6	S. ochotonae	X		KP876415.1		
		128c7	S. ochotonae	X		KP876416.1		
		128c8	S. ochotonae	X		KP876417.1		
		128c9	S. ochotonae	X		KP876418.1		
KG 139			O. princeps					USA, CO, GRAND MESA NATL. FOR., 5 KM S, 33 KM E GRAND JUNCTION
		139c1	S. sp 1	X	X	HQ189817.1		
		139c2	S. sp 1	X	X	HQ189818.1		
		139c3	S. sp 1	X	X	KP876447.1		
		139c4	S. sp 1	X	X	KP876448.1		
		139c5	S. sp 1	X		KP876449.1		
KG 152	20154		O. princeps					USA, CO, WHITE RIVER NATIONAL FOREST, 1KM NW END TRAPPER'S LAKE
		152c1	S. sp 1	X	X	HQ189819.1		
KG 156	20156		O. princeps					USA, CO, WHITE RIVER NATIONAL FOREST, 1KM NW END TRAPPER'S LAKE
		156c1	S. sp 1	X		KP876450.1		
KG 157	20157		O. princeps					USA, CO, WHITE RIVER NATIONAL FOREST, 1KM NW END TRAPPER'S LAKE
		157c1	S. sp 1	X	X	HQ189820.1		
		157c2	S. sp 1					
KG 168	20165		O. princeps					USA,WY, CARBON CO. MEDICINE BOW NATL FOREST
		168c1	S. ochotonae	X	X	HQ189792.1		

		168c2	S. ochotonae	X		KP876419.1		
		168c3	S. ochotonae	X	X	KP876420.1		
		168c4	S. ochotonae	X		KP876421.1		
KG 191	20177		O. princeps					USA,WY, CARBON CO. MEDICINE BOW NATL FOREST
		191c1	S. ochotonae	X	X	HQ189793.1		
		191c2	S. ochotonae	X		KP876422.1		
		191c3	S. ochotonae	X	X	KP876423.1		
		191c4	S. ochotonae	X		KP876424.1		
		191c5	S. ochotonae	X		KP876425.1		
KG 195	20180		O. princeps					USA,WY, TETON CO. SHOSHONE NAT FOR, 2 KM E TOGWOTEE PASS
		195c1	S. sp 3	X		HQ189834.1		
		195c2	S. sp 3	X	first half	KP876458.1		
		195c3	S. sp 3	X	first half	KP876459.1		
		195c4	S. sp 3	X	X	KP876460.1		
KG 205	20189		O. princeps					USA,WY, TETON CO. SHOSHONE NAT FOR, 2 KM E TOGWOTEE PASS
		205c1	S. sp 3	X	X	HQ189835.1		
		205c2	S. sp 3	X	X	HQ189836.1		
		205c3	S. sp 3	X	first half	KP876461.1		
		205c4	S. sp 3	X		KP876462.1		
		205c5	S. sp 3	X		KP876463.1		
		205c6	S. sp 3	X	X	KP876464.1		
KG 214	20198		O. princeps					USA,WY, BIGHORN CO. BIGHORN NAT FOR, WEST SIDE DUNCUM Mt.
		214c1	S.ochotonae	X	X	HQ189794.1		

KG 215	20199		O. princeps					USA,WY, BIGHORN CO. BIGHORN NAT FOR, WEST SIDE DUNCUM Mt.
		215c1	S. ochotonae	X	X	HQ189795.1		
		215c2	S. ochotonae	X		KP876426.1		
		215c3	S. ochotonae	X		KP876427.1		
		215c4	S. ochotonae	X	X	KP876428.1		
		215c5	S. ochotonae	X		KP876429.1		
KG 227	20210		O. princeps					USA, MT, LEWIS & CLARK NATL FOREST, 9 KM E of NEIHART
		227c1	S. ochotonae	X	X	HQ189796.1		
KG 228	20056		O. princeps					USA, MT, LEWIS & CLARK NATL FOREST, 9 KM E of NEIHART
		228c1	S. ochotonae	X	X	KP876430.1		
KG 249	20226		O. princeps					USA, MT, CUSTER NATL FOREST, 1 KM E OF EMERALD LAKE
		249c1	S. sp 3	X		HQ189837.1		
KG 260	20236		O. princeps					USA, MT, CUSTER NATL FOREST, 1 KM E OF EMERALD LAKE
		260c1	S. sp 3	X	X	HQ189838.1		
		260c2	S. ochotonae	X		KP876431.1		
KG 264	20238		O. princeps					USA, MT, BITTERROOT NATL FOREST, 9 KM W OF DARBY
		264c1	S. sp 3	X	X	HQ189839.1		
		264c2	S. ochotonae	X		KP876432.1		
KG 265	20239		O. princeps					USA, MT, BITTERROOT NATL FOREST, 9 KM W OF DARBY

		265c1	S. ochotonae	X	X	HQ189797.1		
		265c2	S. ochotonae	X		KP876433.1		
KG 268	20241		O. princeps					USA,MT, RAVALLI CO. BITTERROOT NATL. FOR. 9 KM W. OF DARBY
		268c1	S. sp 2	X	X	HQ189822.1		
KG 270	20243		O. princeps					USA,MT, RAVALLI CO. BITTERROOT NATL. FOR. 9 KM W. OF DARBY
		270c1	S. sp 2	X	X	HQ189823.1		
		270c2	S. sp 2	X		KP876451.1		
KG 275	20248		O. princeps					USA,ID, BOISE CO, BOISE NATL. FOR. 14 KM W OF FEATHERVILLE
		275c1	S. sp 2	X	X	HQ189824.1		
		275c2	S. sp 2					
KG 281	20253		O. princeps					USA,ID, BOISE CO, BOISE NATL. FOR. 14 KM W OF FEATHERVILLE
		281c1	S. sp 2	X	X	HQ189825.1		
		281c2	S. sp 2	X		KP876452.1		
KG 283	20255		O. princeps					USA,ID, ADAMS CO, PAYETTE NATL. FOR. 1 KM W BLACK LAKE
		283c1	S. ochotonae	X	X	KP876434.1		
KG 285	20257		O. princeps					USA,ID, ADAMS CO, PAYETTE NATL. FOR. 1 KM W BLACK LAKE
		285c2		X	X			
KG 288	20260		O. princeps					USA,ID, ADAMS CO, PAYETTE NATL. FOR. 1 KM W BLACK LAKE
		288c1	S. ochotonae	X	X	KP876435.1		
KG 290	20262		O. princeps					USA,ID, ADAMS CO, PAYETTE NATL. FOR. 1 KM W BLACK LAKE
		290c1	S. ochotonae	X	X	HQ189798.1		

		290c2	S. ochotonae	X	X	HQ189799.1		
KG 296	20268		O. princeps					USA,ID, ADAMS CO, PAYETTE NATL. FOR. 1 KM W BLACK LAKE
		296c1	S. ochotonae	X	X	KP876436.1		
		296c2	S. ochotonae	X	X	KP876437.1		
KG 307	20278		O. princeps					USA,ID,BOUNDARY CO,PANHANDLE NATL. FOR. ROMAN NOSE LAKES
		307c1	S. sp 1		X			
		307c2	S. sp 1	X	X			
		307c3	?	X				
KG 313	20284		O. princeps					USA,WA,OKANOGON CO. OKANOGAN NATL FOR, 2 KM SE WASH. PASS
		313c1	S. sp 1	first half	X			
KG 316	20287		O. princeps					USA,WA,OKANOGON CO. OKANOGAN NATL FOR, 2 KM SE WASH. PASS
		316c1	S. sp 1	X	X	HQ189821.1		
KG 317	20288		O. princeps					USA,WA,OKANOGON CO. OKANOGAN NATL FOR, 2 KM SE WASH. PASS
		317c1	S. ochotonae		X			
KG 331	20300		O. princeps					USA, WA, GIFFORD-PINCHOT NATL FOR, 0.5 KM E SUNRISE PEAK
		331c1	S. ochotonae	X	X	HQ189800.1		
KG 333	20302		O. princeps					USA, WA, GIFFORD-PINCHOT NATL FOR, 0.5 KM E SUNRISE PEAK
		333c1	S. ochotonae	X	X	KP876438.1		
KG 335	20304		O. princeps					USA, WA, GIFFORD-PINCHOT NATL FOR, 0.5 KM E SUNRISE PEAK
		335c1	S. ochotonae	X	X	HQ189801.1		

KG 338	20307		O. princeps					USA, WA, GIFFORD-PINCHOT NATL FOREST, 19 KM W TROUT LAKE
		338c1	S. ochotonae	X	X	HQ189802.1		
KG 341	20310		O. princeps					USA, WA, GIFFORD-PINCHOT NATL FOREST, 19 KM W TROUT LAKE
		341c1	S. ochotonae	X	X	HQ189803.1		
KG 346	20315		O. princeps					USA, NV, HUMBOLDT-TOIYABE NATL FOREST, 5 KM N OF ARC DOME
		346c1	S. sp 2	X		HQ189826.1		
		346c2	S. sp 2	X		HQ189827.1		
KG 348	20317		O. princeps					USA, NV, HUMBOLDT-TOIYABE NATL FOREST, 3 KM S OF MOUNT JEFFERSON
		348c1	S. sp 2	X		HQ189828.1		
		348c2	S. sp 2	X		KP876453.1		
		348c3	S. sp 2	X		KP876454.1		
KG 349	20318		O. princeps					USA, NV, HUMBOLDT-TOIYABE NATL FOREST, 3 KM S OF MOUNT JEFFERSON
		349c1	S. sp 2	X		HQ189829.1		
		349c2	S. sp 2	X		KP876455.1		
KG 351	20320		O. princeps					USA, NV, HUMBOLDT-TOIYABE NATL FOREST, 3 KM S OF MOUNT JEFFERSON
		351c1	S. sp 2	X		KP876456.1		
KG 354	20323		O. princeps					USA,NM,SANTA FE CO, SANTA FE NATL. FOR., W. SLOPE LAKE PK
		354c1	S. ochotonae	X		HQ189804.1		
KG 356	20325		O. princeps					USA,NM,SANTA FE CO, SANTA FE NATL. FOR., W. SLOPE LAKE PK

		356c1	S. ochotonae	X		KP876439.1		
KG 360	20329		O. princeps					USA,NM,SANTA FE CO, SANTA FE NATL. FOR., W. SLOPE LAKE PK
		360c1	S. ochotonae	X	X	KP876440.1		
KG 375			O. collaris					CANADA, NWT
		375c1	S. caballeroi	X	X	HQ189783.1		
KG 377			O. collaris					CANADA, NWT
		377c1	S. caballeroi	X		HQ189784.1		
		377c7	S. caballeroi					
KG 383	20594		O. princeps					USA, CA, EBBETS PASS
		383c1	S. sp 2	X		HQ189830.1		
KG 386	20685		O. princeps					USA, CA, ONION VALLEY
		386c1	S. sp 2	X		HQ189831.1		
KG 390	20688		O. princeps					USA, NV, RUBY MOUNTAINS
		390c1	S. sp 2	X	X	HQ189832.1		
KG 396	20693		O. princeps					USA, NV, RUBY MOUNTAINS
		396c1	S. sp 2	X	X	HQ189833.1		
		396c2	S. sp 2	X		KP876457.1		
KG 411	20704		O. princeps					CANADA, AB, CLINE RIVER
		411c1	S. ochotonae	X		HQ189805.1		
		411c2	S. ochotonae	X	X	KP876441.1		
KG 421	20590		O. princeps					CANADA, BC, MCBRIDE
		421c1	S. ochotonae	X	X	KP876442.1		
		421c2	S. ochotonae	X	X	HQ189806.1		
KG 422	20712		O. princeps					CANADA, BC, MCBRIDE
		422c1	S. ochotonae	X	X	HQ189807.1		
		422c2	S. ochotonae	X	X	KP876443.1		
KG 430	20589		O. princeps					CANADA, AB, GRANDE CACHE
		430c1	S. ochotonae	X	X	HQ189808.1		
KG 433	20722		O. princeps					CANADA, AB, GRANDE CACHE

		433c1	S. ochotonae	X	X	KP876444.1		
KG 470	20750		O. princeps					CANADA, BC, PEMBERTON
		470c1	S. ochotonae	X	X	HQ189809.1		
		470c2			X			
KG 516	20574		O. princeps					USA, OR, WALLOWA MOUNTAINS
		516c1	S. ochotonae	X		HQ189810.1		
KG 517	20649		O. princeps					USA, OR, WALLOWA MOUNTAINS
		517c1	S. ochotonae	X	X	KP876445.1		
KG 531	21155		O. princeps					USA, UT, HELIOTROPE MOUNTAINS
		531c1	S. ochotonae	X	X	HQ189811.1		
KG 534	20659		O. princeps					USA, UT, HELIOTROPE MOUNTAINS
		534c1	S. ochotonae	X	X	HQ189812.1		
KG 537	20569		O. princeps					USA, UT, FLAT TOP
		537c1	S. ochotonae	X	X	HQ189813.1		
KG 540	20663		O. princeps					USA, UT, FLAT TOP
		540c1	S. ochotonae	X	X	HQ189814.1		
KG 555	21157		O. princeps					USA, UT, MUD LAKE
		555c1	S. ochotonae	X	X	HQ189815.1		
		555c2	S. ochotonae	X		KP876446.1		
KG 557	20677		O. princeps					USA, UT, MUD LAKE
		557c1	S. ochotonae	X		HQ189816.1		
KG 558								CHINA
		558c1	S. yamashitai	X		HQ189840.1		
AF38918	UAM:Mamm:80029							RUSSIA, MAGADAN OBLAST
		38918c1	S. mongoliensis	X	X	HQ189785.1		

		38918c2	S. mongoliensis	X		HQ189786.1		
AF49315	UAM:Mamm:58407		O. collaris					USA, ALASKA, CHARLEY RIVER QUAD, YUKON-CHARLEY RIVERS NATL PRESERVE
		49315c2	S. caballeroi	X	X	HQ189777.1		
AF49340	UAM:Mamm:58445		O. collaris					USA, ALASKA, CHARLEY RIVER QUAD, YUKON-CHARLEY RIVERS NATL PRESERVE
		49340c2	S. caballeroi	X	X	HQ189778.1		
AF49615	UAM:Mamm:58257		O. collaris					USA, ALASKA, EAGLE QUAD, YUKON-CHARLEY RIVERS NATL PRESERVE
		49615c1	S. caballeroi	X	X	HQ189779.1		
AF55181	UAM:Mamm:57920		O. collaris					USA, ALASKA, NABESNA QUAD, WRANGELL-ST. ELIAS NATL PARK AND PRESERVE
		55181c3	S. caballeroi	X	X	HQ189780.1		
IF9737	UAM: Mamm:93191		O. collaris					USA, ALASKA, LAKE CLARK QUAD, LAKE CLARK NATIONAL PARK AND PRESERVE
		9737c3	S. caballeroi	X	X	HQ189781.1		
		9737c4	S. caballeroi	X	X	HQ189782.1		
NK270062	MSB:Mamm:285672		O. pallasi					MONGOLIA, UVS, OLGII SUM
		A3JT2		X			MSB:Para: 24655	
NK270428	MSB:Mamm:285675		O. alpina					MONGOLIA, BAYAN-OLGII, SAGSAI SUM
		A3JW9		X	X		MSB:Para: 24681	
NK270393	MSB:Mamm:285674		O. alpina					MONGOLIA, BAYAN-OLGII, SAGSAI SUM
		A3JWD		X	X		MSB:Para: 24678	

Table B2: *L. (Eugenuris)* specimen information list. Museum accession numbers for parasites and associated hosts as well as GenBank numbers for sequenced loci.

Host Field ID	**Host Archive (Cornell)**	**Eug. ID**	**Species**	**COX1**	**CytB**	**ITS2: EUITS2-F/-R**	**GenBank #**	**USNPC Accession**	**Locality**
KG 102	20115		O. princeps						USA,NM,TAOS CO. CARSON NATL FOR, 2.5 KM NW TWINING
		102p3e	L. (E.) talkeetnaeuris	X	X	X		103508.00	
		102p4e	L. (E.) talkeetnaeuris	X	X	X	HQ189956.1	103508.00	
KG 107	20120		O. princeps						USA,NM,TAOS CO. CARSON NATL FOR, 2.5 KM NW TWINING
		107p3e	L. (E.) talkeetnaeuris	X	X	X	HQ189957.1	103509.00	
KG 108	20121		O. princeps						USA,NM,TAOS CO. CARSON NATL FOR, 2.5 KM NW TWINING
		108p3e	L. (E.) talkeetnaeuris	X	X	X	KP876217.1	103510.00	
KG 110	20123		O. princeps						USA,NM,TAOS CO. CARSON NATL FOR, 2.5 KM NW TWINING
		110p2e	L. (E.) talkeetnaeuris	X	X	X	KP876218.1	103511.00	
KG 111	20124		O. princeps						USA,NM,TAOS CO. CARSON NATL FOR, 2.5 KM NW TWINING
		111p3e	L. (E.) talkeetnaeuris	X	X	X	KP876219.1	103562.00	
KG 117	20128		O. princeps						USA,CO,CLEAR CK CO. ARAPAHOE NATL FOR, W. SIDE BERTHOUD PASS
		117p3e	L. (E.) talkeetnaeuris	X	X	X	HQ189958.1	103512.00	
		117p4e	L. (E.) talkeetnaeuris	X	X	X	KP876220.1		

KG 125	20134		O. princeps						USA,CO,CLEAR CK CO. ARAPAHOE NATL FOR, W. SIDE BERTHOUD PASS
		125p1e	L. (E.) talkeetnaeuris	X	X	X	KP876221.1	103513.00	
KG 126	20135		O. princeps						USA,CO,CLEAR CK CO. ARAPAHOE NATL FOR, W. SIDE BERTHOUD PASS
		126p1e	L. (E.) talkeetnaeuris	X		X	KP876222.1	103514.00	
KG 129	20138		O. princeps						USA,CO,CLEAR CK CO. ARAPAHOE NATL FOR, W. SIDE BERTHOUD PASS
		129p3e	L. (E.) talkeetnaeuris	X	X	X	KP876223.1	103515.00	
KG 134	20140		O. princeps						USA,CO,MESA CO,GRAND MESA NATL FOR
		134p3e	L. (E.) talkeetnaeuris	X	X	X	KP876224.1	103516.00	
		134p4e	L. (E.) talkeetnaeuris	X	X	X			
KG 137	20143		O. princeps						USA,CO,MESA CO. GRAND MESA NATL FOR
		137p1e	L. (E.) sp. 1	X			HQ189933.1	103588.00	
KG 140	20146		O. princeps						USA,CO,MESA CO,GRAND MESA NATL Forest
		140p3e	L. (E.) sp. 1	X				103705.00	
KG 145	20149		O. princeps						USA,CO,MESA CO. GRAND MESA NATL FOR
		145p3e	L. (E.) sp. 1	X		X	HQ189934.1	103589.00	
KG 148	20152		O. princeps						USA,CO,MESA CO. GRAND MESA NATL FOR
		148p3e	L. (E.) talkeetnaeuris	X		X	KP876225.1	103517.00	

KG 151	20045		O. princeps						USA,CO, GARFIELD CO. WHITE RIVER NATL FOREST
		151p3e	L. (E.) talkeetnaeuris	X	X	X	KP876226.1	103518.00	
KG 152	20154		O. princeps						USA,CO, GARFIELD CO. WHITE RIVER NATL FOREST
		152p3e	L. (E.) sp. 1	X		X	KP876267.1	103590.00	
KG 154	20046		O. princeps						USA,CO,MESA CO,GRAND MESA NATL Forest
		154p3e	L. (E.) talkeetnaeuris	X	X	X	KP876227.1	103519.00	
KG 157	20157		O. princeps						USA,CO, GARFIELD CO. WHITE RIVER NATL FOREST
		157p3e	L. (E.) sp. 1	X		X	KP876268.1	103591.00	
KG 159	20159		O. princeps						USA,CO, GARFIELD CO. WHITE RIVER NATL FOREST
		159p3e	L. (E.) sp. 2	X	X	X	HQ189938.1	103851.00	
KG 173	20170		O. princeps						USA,CO, GARFIELD CO. WHITE RIVER NATL FOREST
		173p1e	L. (E.) sp. 1	X		X	HQ189935.1	103592.00	
KG 190	20176		O. princeps						USA,WY, CARBON CO. MEDICINE BOW NATL FOREST
		190p3e	L. (E.) sp. 1	X		X	KP876269.1	103593.00	
		190p4e	L. (E.) sp. 1	X		X	KP876270.1		
KG 198	20183		O. princeps						USA,WY, TETON CO. SHOSHONE NAT FOREST
		198p3e	L. (E.) talkeetnaeuris	X	X	X	HQ189959.1		
KG 203	20188		O. princeps						USA,WY, TETON CO. SHOSHONE NAT FOREST
		203p3e	L. (E.) talkeetnaeuris				KP876228.1	103520.00	

KG 205	20189		O. princeps						USA,WY, TETON CO. SHOSHONE NAT FOREST
		205p3e	L. (E.) talkeetnaeuris	X	X	X	KP876229.1	103521.00	
KG 206	20190		O. princeps						USA,WY, TETON CO. SHOSHONE NAT FOREST
		206p3e	L. (E.) sp. 1	X		X	KP876271.1	103594.00	
KG 209	20193		O. princeps						USA,WY, TETON CO. SHOSHONE NAT FOREST
		209p1e	L. (E.) sp. 1	X			KP876272.1	103595.00	
KG 210	20194		O. princeps						USA,WY, BIGHORN CO. BIGHORN NAT FOR, 4 km N DUNCUM Mt.
		210p3e	L. (E.) talkeetnaeuris	X	X	X	HQ189960.1	103522.00	
KG 213	20197		O. princeps						USA,WY, BIGHORN CO. BIGHORN NAT FOR, WEST SIDE DUNCUM Mt.
		213p3e	L. (E.) talkeetnaeuris	X	X		KP876230.1	103523.00	
KG 217	20201		O. princeps						USA,WY, BIGHORN CO. BIGHORN NAT FOR, 1 KM NORTH DUNCUM Mt.
		217p3e	L. (E.) sp. 2	X	X	X	HQ189961.1	103524.00	
KG 219	20203		O. princeps						USA,WY, BIGHORN CO. BIGHORN NAT FOR, 1 KM NORTH DUNCUM Mt.
		219p3e	L. (E.) talkeetnaeuris	X	X		KP876231.1	103525.00	
KG 224	20208		O. princeps						USA,WY, BIGHORN CO. BIGHORN NAT FOR, 5 KM NORTH DUNCUM Mt.
		224p3e	L. (E.) sp. 2	X	X	X	KP876278.1	103852.00	

KG 231	20211		O. princeps						USA,MT,JUDITH BASIN CO,LEWIS & CLARK NATL FOR.
		231p2e	L. (E.) talkeetnaeuris	X	X		KP876232.1	103526.00	
KG 245	20059		O. princeps						USA,MT, CARBON CO. CUSTER NATL FOR. 5 KM EAST GLACIER LAKE
		245p3e	L. (E.) talkeetnaeuris	X	X		HQ189962.1	103527.00	
KG 246	20224		O. princeps						USA,MT, CARBON CO. CUSTER NATL FOR. 5 KM EAST GLACIER LAKE
		246p3e	L. (E.) talkeetnaeuris	X			KP876233.1	103528.00	
KG 249	20226		O. princeps						USA,MT, CARBON CO. CUSTER NATL FOR. 1 KM EAST EMERALD LAKE
		249p3e	L. (E.) sp. 1	X			KP876273.1	103596.00	
KG 254	20231		O. princeps						USA,MT, CARBON CO. CUSTER NATL FOR. 1 KM EAST EMERALD LAKE
		254p3e	L. (E.) talkeetnaeuris	X	X	X	KP876234.1	103529.00	
KG 258	20234		O. princeps						USA,MT,CARBON CO,CUSTER NATL FOR. 1 KM EAST EMERALD LAKE
		258p3e	L. (E.) sp. 2	X	X	X		103725.00	
KG 263	20237		O. princeps						USA,MT, RAVALLI CO. BITTERROOT NATL. FOR. 9 KM W. OF DARBY
		263p3e	L. (E.) talkeetnaeuris	X	X	X	HQ189963.1	103530.00	
KG 264	20238		O. princeps						USA,MT, RAVALLI CO. BITTERROOT NATL. FOR. 9 KM W. OF DARBY

		264p1e	L. (E.) talkeetnaeuris	X			KP876235.1	103531.00	
		264p2e	L. (E.) talkeetnaeuris	X	X	X	HQ189964.1		
KG 265	20239		O. princeps						USA,MT, RAVALLI CO. BITTERROOT NATL. FOR. 9 KM W. OF DARBY
		265p3e	L. (E.) talkeetnaeuris	X	X	X	KP876236.1	103532.00	
KG 268	20241		O. princeps						USA,ID,CUSTER C, CHALLIS NATL. FOREST.
		268p3e	L. (E.) sp. 2	X			HQ189939.1	103853.00	
KG 273	20246		O. princeps						USA,ID,CUSTER C, CHALLIS NATL. FOREST.
		273p1e	L. (E.) talkeetnaeuris	X			KP876237.1	103726.00	
		273p2e	L. (E.) talkeetnaeuris	X	X	X	KP876238.1		
KG 276	20249		O. princeps						USA,ID, BOISE CO, BOISE NATL. FOR. 14 KM W OF FEATHERVILLE
		276p1e	L. (E.) talkeetnaeuris	X			HQ189965.1		
		276p2e	L. (E.) talkeetnaeuris	X	X	X	HQ189966.1	103533.00	
KG 277	20250		O. princeps						USA,ID, BOISE CO, BOISE NATL. FOR. 14 KM W OF FEATHERVILLE
		277p3e	L. (E.) talkeetnaeuris	X			KP876239.1	103534.00	
		277p4e	L. (E.) talkeetnaeuris	X	X	X			
KG 278	20251		O. princeps						USA,ID, BOISE CO, BOISE NATL. FOR. 14 KM W OF FEATHERVILLE
		278p1e	L. (E.) talkeetnaeuris	X	X		KP876240.1	103535.00	

KG 280	20252		O. princeps						USA,ID, BOISE CO, BOISE NATL. FOR. 14 KM W OF FEATHERVILLE
		280p3e	L. (E.) sp. 2	X		X	KP876279.1	103854.00	
KG 287	20259		O. princeps						USA,ID, ADAMS CO, PAYETTE NATL. FOR. 1 KM W BLACK LAKE
		287p1e	L. (E.) talkeetnaeuris	X	X	X	HQ189967.1	103536.00	
KG 290	20262		O. princeps						USA,ID, ADAMS CO, PAYETTE NATL. FOR.
		290p3e	L. (E.) sp. 2	X	X	X	HQ189940.1	103855.00	
		290p4e	L. (E.) talkeetnaeuris	X	X	X	KP876241.1		
KG 298	20270		O. princeps						USA,ID,BOUNDARY CO,PANHANDLE NATL. FOR. ROMAN NOSE LAKES
		298p6e	L. (E.) talkeetnaeuris	X	X		HQ189968.1	103727.00	
KG 300	20272		O. princeps						USA,ID BOUNDARY CO,PANHANDLE NATL. FOR. ROMAN NOSE LAKES
		300p1e	L. (E.) talkeetnaeuris	X	X	X	HQ189969.1	103549.00	
KG 303	20275		O. princeps						USA,ID, BOUNDARY CO, PANHANDLE NATL. FOR. ROMAN NOSE LAKES
		303p3e	L. (E.) talkeetnaeuris	X	X	X	KP876242.1	103550.00	
KG 305	20277		O. princeps						USA,ID, BOUNDARY CO, PANHANDLE NATL. FOR. ROMAN NOSE LAKES
		305p1e	L. (E.) sp. 2	X	X	X	KP876243.1	103551.00	
KG 307	20278		O. princeps						USA,ID, BOUNDARY CO, PANHANDLE NATL. FOR. ROMAN NOSE LAKES

		307p2e	L. (E.) talkeetnaeuris	X	X	X	KP876244.1	103552.00	
KG 309	20280		O. princeps						USA,ID, BOUNDARY CO, PANHANDLE NATL. FOR. ROMAN NOSE LAKES
		309p1e	L. (E.) talkeetnaeuris	X	X	X	KP876245.1	103553.00	
KG 312	20283		O. princeps						USA,WA,OKANOGON CO. OKANOGAN NATL FOR, 2 KM SE WASH. PASS
		312p3e	L. (E.) talkeetnaeuris	X	X	X	HQ189970.1	103554.00	
KG 313	20284		O. princeps						USA,WA,OKANOGON CO. OKANOGAN NATL FOR, 2 KM SE WASH. PASS
		313p1e	L. (E.) talkeetnaeuris	X	X	X	KP876246.1	103555.00	
KG 321	20291		O. princeps						USA,WA,OKANOGON CO. OKANOGAN NATL FOR, 2 KM SE WASH. PASS
		321p1e	L. (E.) talkeetnaeuris	X	X	X	HQ189971.1	103731.00	
KG 322	20292		O. princeps						USA,WA,OKANOGON CO. OKANOGAN NATL FOR, 2 KM SE WASH. PASS
		322p1e	L. (E.) talkeetnaeuris	X		X	KP876249.1	103556.00	
KG 327	20297		O. princeps						USA,WA,OKANOGON CO. OKANOGAN NATL FOR, 2 KM SE WASH. PASS
		327p3e	L. (E.) talkeetnaeuris	X	X		KP876250.1	103557.00	
KG 337	20306		O. princeps						USA,WA,SKAMANIA CO,GIFFORD-PINCHOT NATL. FOR.
		337p3e	L. (E.) talkeetnaeuris	X	X	X	HQ189973.1	103733.00	

KG 342	20311		O. princeps						USA,WA,SKAMANIA CO, GIFFORD-PINCHOT NATL. FOR.
		342p1e	L. (E.) talkeetnaeuris	X	X	X	KP876251.1	103558.00	
KG 343	20312		O. princeps						USA,NV,NYE CO,HUMBOLDT-TOIYABE NATL FOREST
		343p1e	L. (E.) sp. 3	X	X	X	HQ189943.1	103734.00	
KG 345	20314		O. princeps						USA,NV NYE CO,HUMBOLDT-TOIYABE NATL FOREST
		345p2e	L. (E.) sp. 3	X	X	X	HQ189944.1	103735.00	
		345p3e	L. (E.) sp. 3	X	X	X	KP876280.1		
KG 347	20316		O. princeps						USA,NV,NYE CO,HUMBOLDT-TOIYABE NATL FOREST
		347p4e	L. (E.) sp. 2	X		X	HQ189941.1	103856.00	
		347p5e	L. (E.) sp. 2	X		X	HQ189942.1		
KG 350	20319		O. princeps						USA,NV,NYE CO,HUMBOLDT-TOIYABE NATL FOREST
		350p1e	L. (E.) sp. 3	X	X	X	KP876281.1	103736.00	
KG 352	20321		O. princeps						USA,NV,NYE CO,HUMBOLDT-TOIYABE NATL FOREST
		352p3e	L. (E.) sp. 3	X	X		KP876282.1	103737.00	
KG 353	20322		O. princeps						USA,NV NYE CO,HUMBOLDT-TOIYABE NATL FOREST
		353p3e	L. (E.) sp. 3	X			KP876283.1	103738.00	
KG 357	20326		O. princeps						USA,NM,SANTA FE CO, SANTA FE NATL. FOR., W. SLOPE LAKE PK
		357p3e	L. (E.) talkeetnaeuris	X			HQ189973.1	103559.00	

KG 360	20329		O. princeps						USA,NM,SANTA FE CO, SANTA FE NATL. FOR., W. SLOPE LAKE PK
		360p2e	L. (E.) talkeetnaeuris	X	X	X	KP876252.1	103560.00	
KG 362	20331		O. princeps						USA,NM,SANTA FE CO, SANTA FE NATL. FOR., W. SLOPE LAKE PK
		362p3e	L. (E.) talkeetnaeuris	X	X	X	KP876253.1	103561.00	
KG 375			O. collaris						CANADA, NWT
		375p3e	L. (E.) talkeetnaeuris	X			HQ189974.1		
KG 377			O. collaris						CANADA, NWT
		377p3e	L. (E.) talkeetnaeuris	X			HQ189975.1		
KG 378			O. collaris						CANADA, NWT
		378p3e	L. (E.) talkeetnaeuris	X	X	X	HQ189976.1		
KG 379			O. collaris						CANADA, NWT
		379p1e	L. (E.) talkeetnaeuris	X	X		HQ189977.1		
KG 385	20684		O. princeps						USA, CA, ALPINE CO, STANISLAUS NAT. FOREST
		385p3e	L. (E.) sp. 3	X	X	X	KP876284.1	103857.00	
KG 388	20686		O. princeps						USA, CA, INYO CO, INYO NAT. FOREST
		388p3e	L. (E.) sp. 3	X	X	X	HQ189945.1	103858.00	
		388p4e	L. (E.) sp. 3	X	X	X	HQ189946.1		
KG 389	20687		O. princeps						USA, CA, INYO CO, INYO NAT. FOREST
		389p3e	L. (E.) sp. 3	X	X	X	KP876285.1	103859.00	
KG 390	20688		O. princeps						USA,NV,ELKO CO,HUMBOLDT NAT. FOR.
		390p3e	L. (E.) sp. 1	X		X	HQ189936.1	103597.00	
KG 391	20689		O. princeps						USA,NV,ELKO CO,HUMBOLDT NAT. FOR.
		391p1e	L. (E.) sp. 1	X		X	KP876274.1	103598.00	

		391p2e	L. (E.) sp. 1	X		X			
KG 392	20592		O. princeps						USA,NV,ELKO CO,HUMBOLDT NAT. FOR.
		392p1e	L. (E.) sp. 1	X		X	KP876275.1	103599.00	
		392p2e	L. (E.) sp. 1	X					
KG 395	20692		O. princeps						USA,NV, ELKO CO, HUMBOLDT NAT. FOR. 15 km S, 8 km E LAMOILLE
		395p3e	L. (E.) talkeetnaeuris	X	X		KP876254.1	103570.00	
KG 399	20694		O. princeps						USA,NV,ELKO CO,HUMBOLDT NAT. FOR.
		399p3e	L. (E.) talkeetnaeuris	X		X	KP876276.1	103864.00	
KG 404	20697		O. princeps						CANADA, AB, BIGHORN WILDLAND PROVINCIAL RECREATION AREA
		404p3e	L. (E.) talkeetnaeuris	X	X	X	HQ189978.1	103571.00	
KG 406	20699		O. princeps						CANADA, AB, BIGHORN WILDLAND PROVINCIAL RECREATION AREA
		406p3e	L. (E.) talkeetnaeuris	X	X	X	KP876255.1	103572.00	
KG 418	20710		O. princeps						CANADA,AB,BIGHORN WILDLAND PROVINCIAL RECREATION AREA
		418p2e	L. (E.) talkeetnaeuris	X	X	X		103867.00	
KG 419	20711		O. princeps						CANADA, AB, BIGHORN WILDLAND PROVINCIAL RECREATION AREA
		419p3e	L. (E.) talkeetnaeuris	X	X	X	KP876256.1	103573.00	
KG 420	21159		O. princeps						CANADA, AB, BIGHORN WILDLAND

									PROVINCIAL RECREATION AREA
		420p1e	L. (E.) talkeetnaeuris	X	X	X	KP876257.1	103574.00	
KG 428	20717		O. princeps						CANADA, BC, MCBRIDE
		428p1e	L. (E.) talkeetnaeuris	X	X	X	HQ189979.1		
KG 447	20735		O. princeps						CANADA, BC, COAST LAND DISTRICT RANGE 3
		447p3e	L. (E.) talkeetnaeuris	X	X	X	HQ189980.1	103575.00	
KG 452	20740		O. princeps						CANADA,BC,COAST LAND DIST. RANGE 3
		452p3e	L. (E.) talkeetnaeuris	X		X		103870.00	
KG 453	20741		O. princeps						CANADA, BC, COAST LAND DISTRICT RANGE 3
		453p2e	L. (E.) talkeetnaeuris	X	X	X	KP876258.1	103576.00	
KG 457	20745		O. princeps						CANADA, BC, LILLOOET LAND DISTRICT, RAILROAD MT.
		457p3e	L. (E.) talkeetnaeuris	X	X	X	HQ189981.1	103577.00	
KG 461	20585		O. princeps						CANADA, BC, LILLOOET LAND DISTRICT, RAILROAD MT.
		461p3e	L. (E.) talkeetnaeuris	X	X	X	HQ189982.1	103578.00	
		461p4e	L. (E.) talkeetnaeuris		X	X			
KG 462	20584		O. princeps						CANADA, BC, LILLOOET LAND DISTRICT, RAILROAD MT.
		462p3e	L. (E.) talkeetnaeuris	X	X	X	KP876259.1	103579.00	

KG 470	20750		O. princeps						CANADA, BC, LILLOOET LAND DISTRICT, RAILROAD MT.
		470p1e	L. (E.) talkeetnaeuris	X	X	X	KP876260.1		
KG 475	20754		O. princeps						CANADA,BC,LILLOOET LAND DIST., RAILROAD MT.
		475p3e	L. (E.) sp. 1	X	X	X		103878.00	
KG 481	20757		O. princeps						CANADA, BC, RAFT PROVINCIAL FOREST, RAFT MT.
		481p3e	L. (E.) talkeetnaeuris	X			HQ189983.1	103581.00	
KG 484	20759		O. princeps						CANADA,BC,RAFT PROVINCIAL FOREST, RAFT MT.
		484p3e	L. (E.) talkeetnaeuris	X	X	X	KP876261.1	103879.00	
KG 485	20771		O. princeps						CANADA, BC, RAFT PROVINCIAL FOREST, RAFT MT.
		485p3e	L. (E.) talkeetnaeuris	X	X	X	KP876262.1	103582.00	
		485p4e	L. (E.) talkeetnaeuris	X	X	X	KP876263.1		
KG 488	21152		O. princeps						USA,OR,LANE CO.WILLAMETTE NAT. FOREST
		488p3e	L. (E.) talkeetnaeuris	X	X	X	HQ189984.1	103583.00	
KG 489	21161		O. princeps						USA,OR,LANE CO.WILLAMETTE NAT. FOREST
		489p2e	L. (E.) talkeetnaeuris	X	X	X	KP876264.1	103584.00	
KG 490	20628		O. princeps						USA,OR,LANE CO.WILLAMETTE NAT. FOREST
		490p2e	L. (E.) talkeetnaeuris	X	X	X	KP876265.1	103585.00	

KG 493	20578		O. princeps						USA,OR,JEFFERSON CO.WILLAMETTE NAT. FOREST
		493p2e	L. (E.) talkeetnaeuris	X	X		KP876266.1	103586.00	
KG 494	20631		O. princeps						USA,OR,JEFFERSON CO.WILLAMETTE NAT. FOREST
		494p2e	L. (E.) talkeetnaeuris	X		X	HQ189985.1	103587.00	
KG 508	20642		O. princeps						USA,OR,BAKER CO,WHITMAN NAT. FOR.,
		508p1e	L. (E.) sp. 3	X	X		HQ189947.1	103880.00	
KG 512	20576		O. princeps						USA,OR,BAKER CO,WHITMAN NAT. FOR.,
		512p3e	L. (E.) sp. 3	X	X	X	KP876286.1	103881.00	
KG 518	20573		O. princeps						USA,UT,SUMMIT CO,WASATCH NAT. FOR.,24 km E, 6 km N OAKLEY
		518p1e	L. (E.) sp. 1			X		103884.00	
KG 522	20653		O. princeps						USA,UT,SUMMIT CO,WASATCH NAT. FOR. 24 km E, 6 km N of OAKLEY
		522p3e	L. (E.) sp. 1	X		X	HQ189937.1	103600.00	
		522p4e	L. (E.) sp. 1			X			
KG 524	20655		O. princeps						USA,UT,SUMMIT CO,WASATCH NAT. FOR.,22 km E, 5 km N OAKLEY
		524p3e	L. (E.) sp. 3	X	X			103885.00	
KG 525	20656		O. princeps						USA,UT,SUMMIT CO,WASATCH NAT. FOR. 24 km E, 6 km N of OAKLEY
		525p3e	L. (E.) sp. 1	X		X	KP876277.1	103601.00	

KG 527	20657		O. princeps						USA,UT,SUMMIT CO,WASATCH NAT. FOR.,22 km E, 5 km N OAKLEY
		527p3e	L. (E.) sp. 3	X	X	X	HQ189948.1	103887.00	
KG 528	20658		O. princeps						USA,UT,SUMMIT CO,WASATCH NAT. FOR. 24 km E, 6 km N of OAKLEY
		528p1e	L. (E.) sp. 1			X		103602.00	
KG 529	20678		O. princeps						USA,UTAH,SANPETE CO,MANTI-LASAL NAT.FOR. 0.5 km SW ISLAND L.
		529p3e	L. (E.) sp. 3	X	X		HQ189949.1	103889.00	
KG 530	20571		O. princeps						USA,UTAH,SANPETE CO,MANTI-LASAL NAT.FOR. 0.5 km SW ISLAND L.
		530p3e	L. (E.) sp. 3	X	X	X	HQ189950.1	103890.00	
KG 532	20743		O. princeps						USA,UTAH,SANPETE CO,MANTI-LASAL NAT.FOR.
		532p3e	L. (E.) sp. 3	X	X	X	KP876287.1	103891.00	
KG 533	20570		O. princeps						USA,UTAH,SANPETE CO,MANTI-LASAL NAT.FOR.
		533p3e	L. (E.) sp. 3	X	X		KP876288.1	103892.00	
KG 535	20660		O. princeps						USA,UTAH,SANPETE CO,MANTI-LASAL NAT.FOR.
		535p3e	L. (E.) sp. 3	X	X	X	KP876289.1	103893.00	
KG 540	20663		O. princeps						USA,UTAH,WAYNE CO,FISH LAKE NAT.FOR.NW SLOPE FLAT TOP MT.
		540p3e	L. (E.) sp. 3	X	X		HQ189951.1	103894.00	
		540p4e	L. (E.) sp. 3	X	X	X	KP876290.1		
KG 542	20666		O. princeps						USA,UT,WAYNE CO,FISH LAKE NAT.FOR. NW slope Flat Top Mt.

		542p3e	L. (E.) sp. 3	X	X		KP876291.1	103898.00	
KG 553	20674		O. princeps						USA,UT, MILLARD CO, FISHLAKE NAT. FOR. 3 KM E, 2 KM S, MT BALDY
		553p3e	L. (E.) sp. 3	X	X	X	HQ189952.1	103896.00	
		553p4e	L. (E.) sp. 3	X	X	X	KP876292.1		
KG 555	21157		O. princeps						USA,UT, MILLARD CO, FISHLAKE NAT. FOR. 3 KM E, 2 KM S, MT BALDY
		555p1e	L. (E.) sp. 3	X	X	X	KP876293.1	103897.00	
KG 563			O. collaris						CANADA, NWT
		563p2e	L. (E.) talkeetnaeuris	X	X	X	HQ189986.1		
		563p3e	L. (E.) talkeetnaeuris	X	X		HQ189987.1		
IF7687	UAM:Mamm:83952		O. hyperborea						RUSSIA, CHUKOTKA AUTONOMOUS OKRUG, YTTYGRAN ISLAND
		IF7687n2p3e	L. (E.) talkeetnaeuris	X	X		HQ189953.1		
IF9635	UAM:Mamm:93095		O. collaris						USA, ALASKA, LAKE CLARK QUAD, LAKE CLARK NATL PARK AND PRESERVE
		IF9635np3e	L. (E.) talkeetnaeuris	X	X	X	HQ189954.1		
IF9737	UAM: Mamm:93191		O. collaris						USA, ALASKA, LAKE CLARK QUAD, LAKE CLARK NATL PARK AND PRESERVE
		IF9737n2p3e	L. (E.) talkeetnaeuris	X		X	HQ189955.1		
NK272257	MSB:Mamm:285686		O. hyperborea						MONGOLIA, KHOVSGOL, CHANDMANI-ONDOR SUM
		A3K0H	L. (E.) sp.,	X	X	X			

NK272257	MSB:Mamm:285686		O. hyperborea						MONGOLIA, KHOVSGOL, CHANDMANI-ONDOR SUM
		A3K09	L. (Labiostomum) sp.	X		X			
NK270429	MSB:Mamm:288874		O. alpina						MONGOLIA, BAYAN-OLGII, SAGSAI SUM
		A3K00	L. (E.) sp.,	X	X	X			
NK270629	MSB:Mamm:288886		O. dauurica						MONGOLIA, UVS, TURGEN SUM
		A3JZR	L. (E.) sp.,	X	X	X			
NK270393	MSB:Mamm:285674		O. alpina						MONGOLIA, BAYAN-OLGII, SAGSAI SUM
		A3JZI	L. (E.) sp.,	X	X	X			
NK270393	MSB:Mamm:285674		O. alpina						MONGOLIA, BAYAN-OLGII, SAGSAI SUM
		A3JZB	L. (Labiostomum) sp.	X	X	X			

12S16S

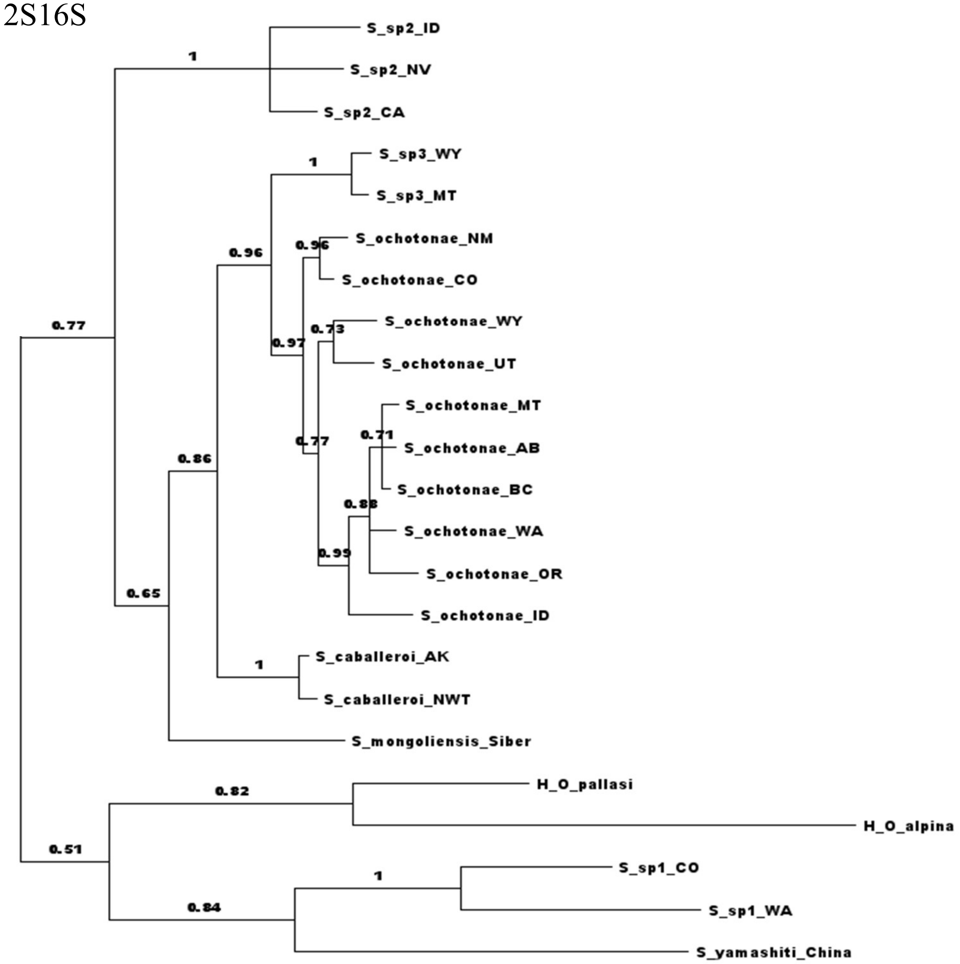


Figure B1: MrBayes gene tree for *Schizorchis* 12S16S using representative haplotypes. Names given as parasite species and region or host name (H_) if parasite species not yet identified. Nodal support provided in posterior probabilities.

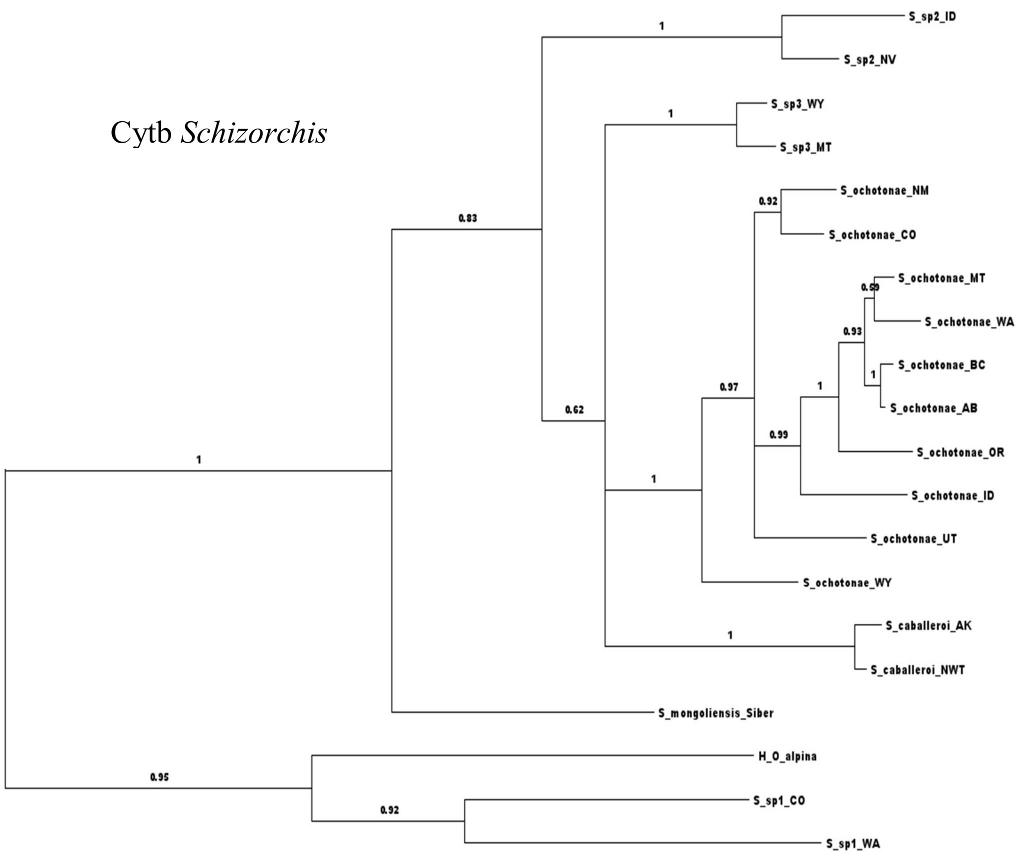
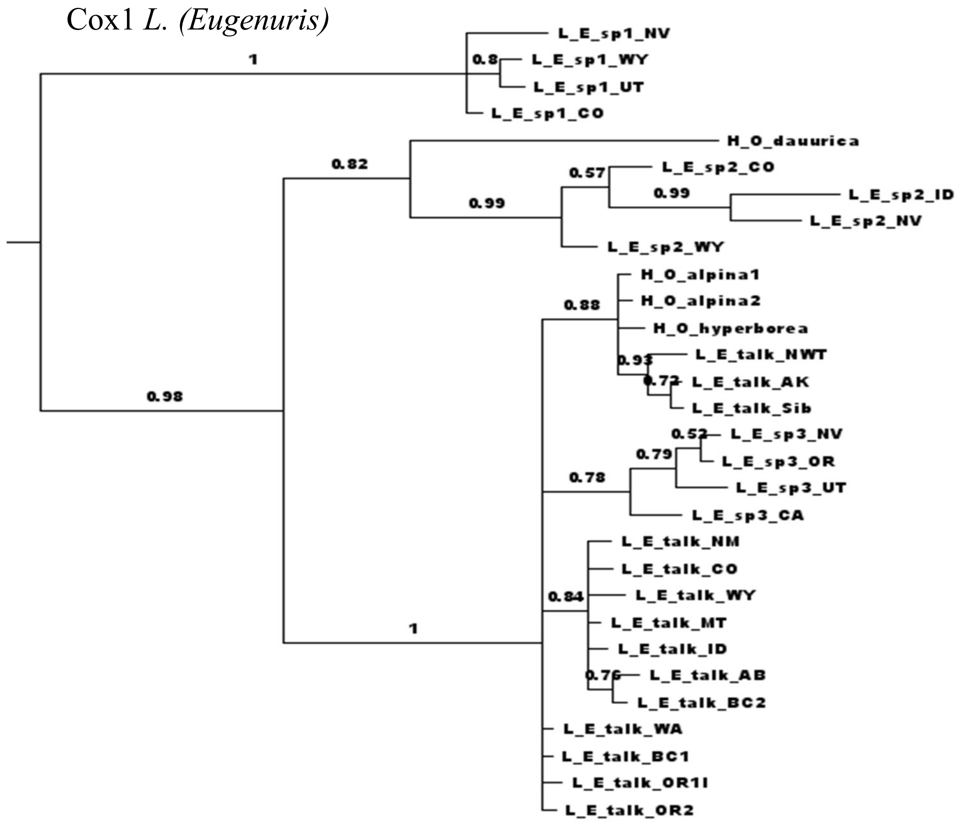


Figure B2: MrBayes gene tree for *Schizorchis* Cytb using representative haplotypes. Names given as parasite species and region or host name (H_) if parasite species not yet identified. Nodal support provided in posterior probabilities.



Cox1 *L. (Eugenuris)*

Figure B3: MrBayes gene tree for *L. (Eugenuris)* Cox1 using representative haplotypes. Names given as parasite species and region or host name (H_) if parasite species not yet identified. Nodal support provided in posterior probabilities.

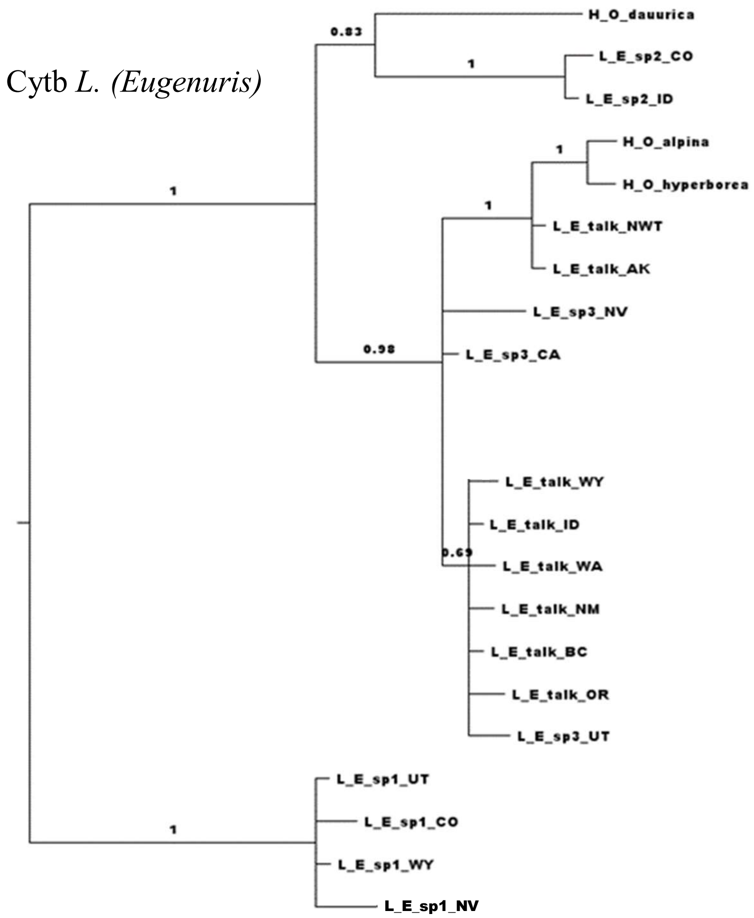


Figure B4: MrBayes gene tree for *L. (Eugenuris)* Cytb using representative haplotypes. Names given as parasite species and region or host name (H_) if parasite species not yet identified. Nodal support provided in posterior probabilities.

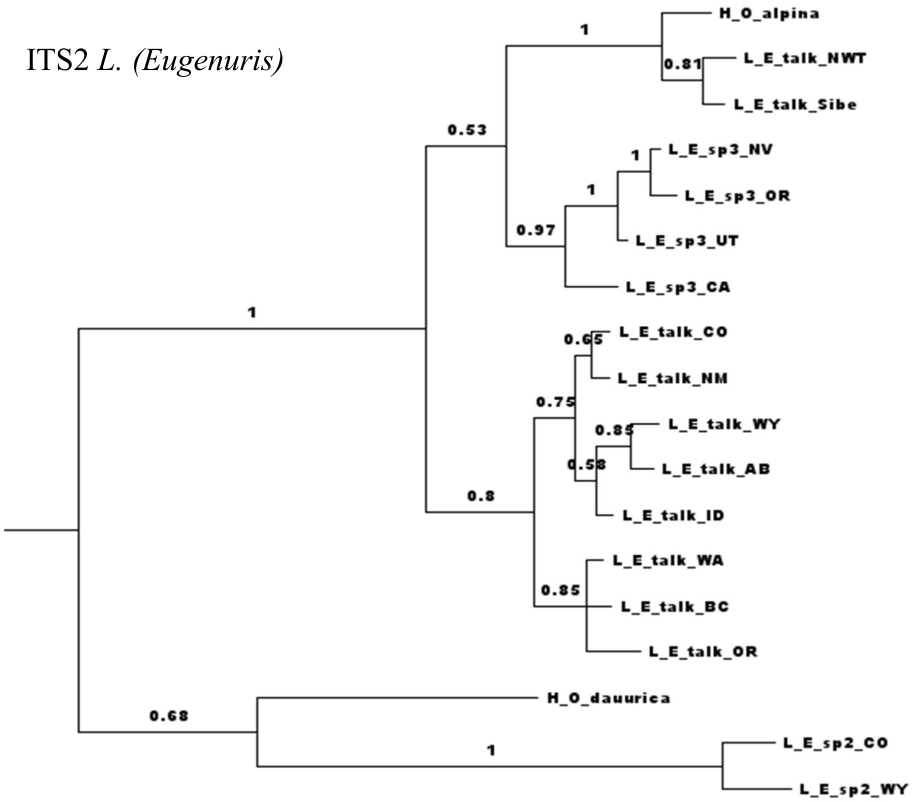


Figure B5: MrBayes gene tree for *L. (Eugenuris)* ITS2 using representative haplotypes. Names given as parasite species and region or host name (H_) if parasite species not yet identified. Nodal support provided in posterior probabilities.

Table B3: Dispersal multiplier matrix for BioGeoBEARS. Represents relative ease of access from one region to another on a scale of 0-1. P (Palearctic), C (Canada), M (Intermountain West), S (American southwest).

Dispersal multiplier matrix for BioGeoBEARS			
P	C	M	S
1	0.5	0.25	0
0.5	1	1	0.25
0.25	1	1	1
0	0.25	1	1

Table B4: R Script for *Schizorchis* BioGeoBEARS Analysis (modified from: http://phylo.wikidot.com/biogeobears)

```
install.packages("optimx", dependencies=TRUE, repos="http://cran.rstudio.com")

# Install BioGeoBEARS from CRAN 0-cloud:
install.packages("BioGeoBEARS", dependencies=TRUE, repos="http://cran.rstudio.com")

library(optimx)  # (either 2012 or 2013 version, as of January 2014)
library(parallel)
library(FD)       # for FD::maxent() (make sure this is up-to-date)
library(snow)     # (if you want to use multicore functionality; prob. better than library(parallel))
library(BioGeoBEARS)
source("http://phylo.wdfiles.com/local--files/biogeobears/cladoRcpp.R") # (needed now that traits model added; source FIRST!)
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_add_fossils_randomly_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_basics_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_calc_transition_matrices_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_classes_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_detection_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_DNA_cladogenesis_sim_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_extract_Qmat_COOmat_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_generics_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_models_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_on_multiple_trees_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_plots_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_readwrite_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_simulate_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_SSEsim_makePlots_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_SSEsim_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_stochastic_mapping_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_stratified_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_univ_model_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/calc_uppass_probs_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/calc_loglike_sp_v01.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/get_stratified_subbranch_top_downpass_likelihoods_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/runBSM_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/stochastic_map_given_inputs.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/summarize_BSM_tables_v1.R")
source("http://phylo.wdfiles.com/local--files/biogeobears/BioGeoBEARS_traits_v1.R") # added traits model
```

```
calc_loglike_sp = compiler::cmpfun(calc_loglike_sp_prebyte)    # crucial to fix bug in uppass calculations
calc_independent_likelihoods_on_each_branch = compiler::cmpfun(calc_independent_likelihoods_on_each_branch_prebyte)

setwd("C:\\R_BGB\\Schiz2.2017")
getwd()
trfn="SchizSpeciesBGB.newick"
geogfn="SchizSp4_geog.data"

moref(geogfn)

# Look at your geographic range data:
tipranges = getranges_from_LagrangePHYLIP(lgdata_fn=geogfn)
tipranges

tr = read.tree(trfn)
tr
plot(tr)
axisPhylo() # plots timescale
max_range_size = 3

############################################################
# Run DEC
############################################################
# Intitialize a default model (DEC model)
BioGeoBEARS_run_object = define_BioGeoBEARS_run()
GeoBEARS_run_object$force_sparse=FALSE    # sparse=TRUE causes pathology & isn't much faster at this scale
BioGeoBEARS_run_object$speedup=TRUE          # shorcuts to speed ML search; use FALSE if worried (e.g. >3 params)
BioGeoBEARS_run_object$use_optimx = TRUE
BioGeoBEARS_run_object$calc_ancprobs=TRUE    # get ancestral states from optim run
BioGeoBEARS_run_object$include_null_range = TRUE

# Give BioGeoBEARS the location of the phylogeny Newick file
BioGeoBEARS_run_object$trfn = trfn

# Give BioGeoBEARS the location of the geography text file
BioGeoBEARS_run_object$geogfn = geogfn
# Input the maximum range size
BioGeoBEARS_run_object$max_range_size = max_range_size

#BioGeoBEARS_run_object$min_branchlength = 0.000001    # Min to treat tip as a direct ancestor (no speciation event)
BioGeoBEARS_run_object$include_null_range = TRUE    # set to FALSE for e.g. DEC* model, DEC*+J, etc.
```

```
#(For DEC* and other "*" models, please cite: Massana, Kathryn A.; Beaulieu, Jeremy M.; Matzke, Nicholas J.; O'Meara, Brian C. (2015). Non-null Effects of the Null Range in Biogeographic Models: Exploring Parameter Estimation in the DEC Model. bioRxiv,  http://biorxiv.org/content/early/2015/09/16/026914 )
# also search script on "include_null_range" for other places to change

# Set up a time-stratified analysis:
# 1. Here, un-comment ONLY the files you want to use.
# 2. Also un-comment "BioGeoBEARS_run_object = section_the_tree(...", below.
# 3. For example files see (a) extdata_dir,
#  or (b) http://phylo.wikidot.com/biogeobears#files
#  and BioGeoBEARS Google Group posts for further hints)
#
# Uncomment files you wish to use in time-stratified analyses:
BioGeoBEARS_run_object$timesfn = "timeperiods.txt"
BioGeoBEARS_run_object$dispersal_multipliers_fn = "dispersal_multipliers.txt"
#BioGeoBEARS_run_object$areas_allowed_fn = "areas_allowed.txt"
#BioGeoBEARS_run_object$areas_adjacency_fn = "areas_adjacency.txt"
#BioGeoBEARS_run_object$distsfn = "distances_matrix.txt"
# See notes on the distances model on PhyloWiki's BioGeoBEARS updates page.

# Speed options and multicore processing if desired
BioGeoBEARS_run_object$num_cores_to_use = 1
# (use more cores to speed it up; this requires
# library(parallel) and/or library(snow). The package "parallel"
# is now default on Macs in R 3.0+, but apparently still
# has to be typed on some Windows machines. Note: apparently
# parallel works on Mac command-line R, but not R.app.
# BioGeoBEARS checks for this and resets to 1
# core with R.app)

# This function loads the dispersal multiplier matrix etc. from the text files into the model object. Required for these to work!
# (It also runs some checks on these inputs for certain errors.)
BioGeoBEARS_run_object = readfiles_BioGeoBEARS_run(BioGeoBEARS_run_object)

# Divide the tree up by timeperiods/strata (uncomment this for stratified analysis)
BioGeoBEARS_run_object = section_the_tree(inputs=BioGeoBEARS_run_object, make_master_table=TRUE, plot_pieces=FALSE)
# The stratified tree is described in this table:
BioGeoBEARS_run_object$master_table

# Good default settings to get ancestral states
BioGeoBEARS_run_object$return_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_TTL_loglike_from_condlikes_table = TRUE
```

```
BioGeoBEARS_run_object$calc_ancprobs = TRUE    # get ancestral states from optim run

# Set up DEC model
# Max values for d, e
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","max"] = 400
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","max"] = 400

# Look at the BioGeoBEARS_run_object; it's just a list of settings etc.
BioGeoBEARS_run_object

# This contains the model object
BioGeoBEARS_run_object$BioGeoBEARS_model_object

# This table contains the parameters of the model
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table

# Run this to check inputs. Read the error messages if you get them!
check_BioGeoBEARS_run(BioGeoBEARS_run_object)

# For a slow analysis, run once, then set runslow=FALSE to just
# load the saved result.
runslow = TRUE
resfn = "Schiz_DEC_2017.Rdata"
if (runslow)
{
  res = bears_optim_run(BioGeoBEARS_run_object)
  res

  save(res, file=resfn)
  resDEC = res
} else {
  # Loads to "res"
  load(resfn)
  resDEC = res
}

#######################################################
# Run DEC+J
#######################################################
BioGeoBEARS_run_object = define_BioGeoBEARS_run()
BioGeoBEARS_run_object$trfn = trfn
```

```
BioGeoBEARS_run_object$geogfn = geogfn
BioGeoBEARS_run_object$max_range_size = max_range_size
#BioGeoBEARS_run_object$min_branchlength = 0.000001    # Min to treat tip as a direct ancestor (no speciation event)
#(For DEC* and other "*" models, please cite: Massana, Kathryn A.; Beaulieu, Jeremy M.; Matzke, Nicholas J.; O'Meara, Brian C. (2015). Non-null Effects of the Null Range in
Biogeographic Models: Exploring Parameter Estimation in the DEC Model. bioRxiv,  http://biorxiv.org/content/early/2015/09/16/026914 )
# also search script on "include_null_range" for other places to change

# Set up a time-stratified analysis:
BioGeoBEARS_run_object$timesfn = "timeperiods.txt"
BioGeoBEARS_run_object$dispersal_multipliers_fn = "dispersal_multipliers.txt"
#BioGeoBEARS_run_object$areas_allowed_fn = "areas_allowed.txt"
#BioGeoBEARS_run_object$areas_adjacency_fn = "areas_adjacency.txt"
#BioGeoBEARS_run_object$distsfn = "distances_matrix.txt"
# See notes on the distances model on PhyloWiki's BioGeoBEARS updates page.

# Speed options and multicore processing if desired
BioGeoBEARS_run_object$speedup=TRUE          # shorcuts to speed ML search; use FALSE if worried (e.g. >3 params)
BioGeoBEARS_run_object$use_optimx = TRUE
BioGeoBEARS_run_object$num_cores_to_use=1
BioGeoBEARS_run_object$force_sparse=FALSE    # sparse=FALSE causes pathology & isn't much faster at this scale
BioGeoBEARS_run_object$calc_ancprobs=TRUE    # get ancestral states from optim run
BioGeoBEARS_run_object$include_null_range = TRUE

# This function loads the dispersal multiplier matrix etc. from the text files into the model object. Required for these to work!
# (It also runs some checks on these inputs for certain errors.)
BioGeoBEARS_run_object = readfiles_BioGeoBEARS_run(BioGeoBEARS_run_object)

# Divide the tree up by timeperiods/strata (uncomment this for stratified analysis)
BioGeoBEARS_run_object = section_the_tree(inputs=BioGeoBEARS_run_object, make_master_table=TRUE, plot_pieces=FALSE)
# The stratified tree is described in this table:
BioGeoBEARS_run_object$master_table

# Good default settings to get ancestral states
BioGeoBEARS_run_object$return_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_TTL_loglike_from_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_ancprobs = TRUE    # get ancestral states from optim run

# Set up DEC+J model
# Get the ML parameter values from the 2-parameter nested model
# (this will ensure that the 3-parameter model always does at least as good)
dstart = resDEC$outputs@params_table["d","est"]
```

```
estart = resDEC$outputs@params_table["e","est"]
jstart = 0.0001

# Input starting values for d, e
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","init"] = dstart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","est"] = dstart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","init"] = estart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","est"] = estart

# Max values for d, e
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","max"] = 400
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","max"] = 400

# Add j as a free parameter
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","type"] = "free"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","init"] = jstart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","est"] = jstart

check_BioGeoBEARS_run(BioGeoBEARS_run_object)

resfn = "Schiz_DEC+J_2017.Rdata"
runslow = TRUE
if (runslow)
{
  #sourceall("/Dropbox/_njm/__packages/BioGeoBEARS_setup/")

  res = bears_optim_run(BioGeoBEARS_run_object)
  res

  save(res, file=resfn)

  resDECj = res
} else {
  # Loads to "res"
  load(resfn)
  resDECj = res
}

#######################################################
# PDF plots
#######################################################
```

```
pdffn = "Schiz_DEC_vs_DEC+J_2017.pdf"
pdf(pdffn, width=6, height=6)

########################################################
# Plot ancestral states - DEC
########################################################
analysis_titletxt ="BioGeoBEARS DEC on Schiz"

# Setup
results_object = resDEC
scriptdir = np(system.file("extdata/a_scripts", package="BioGeoBEARS"))

# States
res2 = plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="text", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

# Pie chart
plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="pie", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

########################################################
# Plot ancestral states - DECJ
########################################################
analysis_titletxt ="BioGeoBEARS DEC+J on Schiz"

# Setup
results_object = resDECj
scriptdir = np(system.file("extdata/a_scripts", package="BioGeoBEARS"))

# States
res1 = plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="text", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

# Pie chart
plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="pie", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

dev.off()  # Turn off PDF
cmdstr = paste("open ", pdffn, sep="")
system(cmdstr) # Plot it
```

```
#######################################################
# Run DEC w/o Dispersal Multipliers
#######################################################

BioGeoBEARS_run_object = define_BioGeoBEARS_run()
BioGeoBEARS_run_object$force_sparse=FALSE    # sparse=FALSE causes pathology & isn't much faster at this scale
BioGeoBEARS_run_object$speedup=TRUE          # shorcuts to speed ML search; use FALSE if worried (e.g. >3 params)
BioGeoBEARS_run_object$use_optimx = TRUE
BioGeoBEARS_run_object$calc_ancprobs=TRUE    # get ancestral states from optim run
BioGeoBEARS_run_object$include_null_range = TRUE

# Set up a time-stratified analysis
# (un-comment to use; see example files in extdata_dir,
#  and BioGeoBEARS google group posts for further hints)
#BioGeoBEARS_run_object$timesfn = "timeperiods.txt"
#BioGeoBEARS_run_object$dispersal_multipliers_fn = "dispersal_multipliers.txt"
#BioGeoBEARS_run_object$areas_allowed_fn = "areas_allowed_aros5g.txt"
#BioGeoBEARS_run_object$areas_adjacency_fn = "areas_adjacency.txt"
#BioGeoBEARS_run_object$distsfn = "distances_matrix.txt"
# See notes on the distances model on PhyloWiki's BioGeoBEARS updates page.

# Input the maximum range size
BioGeoBEARS_run_object$max_range_size = max_range_size

# Multicore processing if desired
BioGeoBEARS_run_object$num_cores_to_use=1
# (use more cores to speed it up; this requires
# library(parallel) and/or library(snow). The package "parallel"
# is now default on Macs in R 3.0+, but apparently still
# has to be typed on some Windows machines. Note: apparently
# parallel works on Mac command-line R, but not R.app.
# BioGeoBEARS checks for this and resets to 1
# core with R.app)

# Sparse matrix exponentiation is an option for huge numbers of ranges/states (600+)
BioGeoBEARS_run_object$force_sparse=FALSE

# Give BioGeoBEARS the location of the geography text file
BioGeoBEARS_run_object$geogfn = geogfn

# Give BioGeoBEARS the location of the phylogeny Newick file
```

```
BioGeoBEARS_run_object$trfn = trfn

# This function loads the dispersal multiplier matrix etc. from the text files into the model object. Required for these to work!
# (It also runs some checks on these inputs for certain errors.)
BioGeoBEARS_run_object = readfiles_BioGeoBEARS_run(BioGeoBEARS_run_object)

# Divide the tree up by timeperiods/strata (uncomment this for stratified analysis)
#BioGeoBEARS_run_object = section_the_tree(inputs=BioGeoBEARS_run_object, make_master_table=TRUE, plot_pieces=FALSE)
# The stratified tree is described in this table:
#BioGeoBEARS_run_object$master_table

# Good default settings to get ancestral states
BioGeoBEARS_run_object$return_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_TTL_loglike_from_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_ancprobs = TRUE

# Set up DEC model
# Max values for d, e
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","max"] = 400
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","max"] = 400


# Look at the BioGeoBEARS_run_object; it's just a list of settings etc.
BioGeoBEARS_run_object

# This contains the model object
BioGeoBEARS_run_object$BioGeoBEARS_model_object

# This table contains the parameters of the model
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table

# Run this to check inputs. Read the error messages if you get them!
check_BioGeoBEARS_run(BioGeoBEARS_run_object)

# For a slow analysis, run once, then set runslow=FALSE to just
# load the saved result.
runslow = TRUE
resfn = "Schiz_DEC_2017_DECnoDisMult.Rdata"
if (runslow)
{
  res = bears_optim_run(BioGeoBEARS_run_object)
```

```
  res

  save(res, file=resfn)
  resDECnoMult = res
} else {
  # Loads to "res"
  load(resfn)
  resDECnoMult = res
}

#######################################################
# Run DEC+J W/o dispersal multipliers
#######################################################
BioGeoBEARS_run_object = define_BioGeoBEARS_run()
BioGeoBEARS_run_object$trfn = trfn
BioGeoBEARS_run_object$geogfn = geogfn
BioGeoBEARS_run_object$max_range_size = max_range_size

# Set up the stratified part
#BioGeoBEARS_run_object$timesfn = "timeperiods.txt"
#BioGeoBEARS_run_object$dispersal_multipliers_fn = "dispersal_multipliers.txt"
#BioGeoBEARS_run_object$areas_allowed_fn = "areas_allowed_aros5g.txt"
#BioGeoBEARS_run_object$areas_adjacency_fn = "areas_adjacency.txt"
#BioGeoBEARS_run_object$distsfn = "distances_matrix.txt"
# See notes on the distances model on PhyloWiki's BioGeoBEARS updates page.

BioGeoBEARS_run_object$speedup=TRUE          # shorcuts to speed ML search; use FALSE if worried (e.g. >3 params)
BioGeoBEARS_run_object$use_optimx = TRUE
BioGeoBEARS_run_object$num_cores_to_use=1
BioGeoBEARS_run_object$force_sparse=FALSE    # sparse=FALSE causes pathology & isn't much faster at this scale
BioGeoBEARS_run_object$calc_ancprobs=TRUE    # get ancestral states from optim run
BioGeoBEARS_run_object$include_null_range = TRUE

# This function loads the dispersal multiplier matrix etc. from the text files into the model object. Required for these to work!
# (It also runs some checks on these inputs for certain errors.)
BioGeoBEARS_run_object = readfiles_BioGeoBEARS_run(BioGeoBEARS_run_object)

# Divide the tree up by timeperiods/strata (uncomment this for stratified analysis)
#BioGeoBEARS_run_object = section_the_tree(inputs=BioGeoBEARS_run_object, make_master_table=TRUE, plot_pieces=FALSE)
# The stratified tree is described in this table:
#BioGeoBEARS_run_object$master_table
```

```
# Good default settings to get ancestral states
BioGeoBEARS_run_object$return_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_TTL_loglike_from_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_ancprobs = TRUE

# Set up DEC+J model
# Get the ML parameter values from the 2-parameter nested model
# (this will ensure that the 3-parameter model always does at least as good)
dstart = resDECnoMult$outputs@params_table["d","est"]
estart = resDECnoMult$outputs@params_table["e","est"]
jstart = 0.0001

# Input starting values for d, e
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","init"] = dstart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","est"] = dstart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","init"] = estart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","est"] = estart

# Max values for d, e
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","max"] = 400
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","max"] = 400

# Add j as a free parameter
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","type"] = "free"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","init"] = jstart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","est"] = jstart

check_BioGeoBEARS_run(BioGeoBEARS_run_object)

resfn = "Schiz_DEC+J_2017_DEC+JnoDisMult.Rdata"
runslow = TRUE
if (runslow)
{
  #sourceall("/Dropbox/_njm/__packages/BioGeoBEARS_setup/")

  res = bears_optim_run(BioGeoBEARS_run_object)
  res

  save(res, file=resfn)
```

```
  resDECjnoMult = res
} else {
  # Loads to "res"
  load(resfn)
  resDECjnoMult = res
}

#######################################################
# PDF plots
#######################################################
pdffn = "Schiz_noDM_DEC_vs_DEC+J_2017.pdf"
pdf(pdffn, width=6, height=6)

#######################################################
# Plot ancestral states - DEC
#######################################################
analysis_titletxt ="BioGeoBEARS DEC No Dispersal Multipliers Schiz"

# Setup
results_object = resDECnoMult
scriptdir = np(system.file("extdata/a_scripts", package="BioGeoBEARS"))

# States
res2 = plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="text", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

# Pie chart
plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="pie", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

#######################################################
# Plot ancestral states - DECJ
#######################################################
analysis_titletxt ="BioGeoBEARS DEC+J No Dispersal Multipliers Schiz"

# Setup
results_object = resDECjnoMult
scriptdir = np(system.file("extdata/a_scripts", package="BioGeoBEARS"))

# States
```

```
res1 = plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="text", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8, plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

# Pie chart
plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="pie", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8, plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

dev.off()  # Turn off PDF
cmdstr = paste("open ", pdffn, sep="")
system(cmdstr) # Plot it

#######################################################
#######################################################
# DIVALIKE AND DIVALIKE+J ANALYSIS
#######################################################
#######################################################
# NOTE: The BioGeoBEARS "DIVALIKE" model is not identical with
# Ronquist (1997)'s parsimony DIVA. It is a likelihood
# interpretation of DIVA, constructed by modelling DIVA's
# processes the way DEC does, but only allowing the
# processes DIVA allows (widespread vicariance: yes; subset
# sympatry: no; see Ronquist & Sanmartin 2011, Figure 4).
#
# DIVALIKE is a likelihood interpretation of parsimony
# DIVA, and it is "like DIVA" -- similar to, but not
# identical to, parsimony DIVA.
#
# I thus now call the model "DIVALIKE", and you should also. ;-)
#######################################################
#######################################################

#######################################################
# Run DIVALIKE
#######################################################
BioGeoBEARS_run_object = define_BioGeoBEARS_run()
BioGeoBEARS_run_object$trfn = trfn
BioGeoBEARS_run_object$geogfn = geogfn
BioGeoBEARS_run_object$max_range_size = max_range_size
#BioGeoBEARS_run_object$min_branchlength = 0.000001    # Min to treat tip as a direct ancestor (no speciation event)
BioGeoBEARS_run_object$include_null_range = TRUE    # set to FALSE for e.g. DEC* model, DEC*+J, etc.
```

```
#(For DEC* and other "*" models, please cite: Massana, Kathryn A.; Beaulieu, Jeremy M.; Matzke, Nicholas J.; O'Meara, Brian C. (2015). Non-null Effects of the Null Range in
Biogeographic Models: Exploring Parameter Estimation in the DEC Model. bioRxiv,  http://biorxiv.org/content/early/2015/09/16/026914 )
# also search script on "include_null_range" for other places to change

# Set up a time-stratified analysis:
BioGeoBEARS_run_object$timesfn = "timeperiods.txt"
BioGeoBEARS_run_object$dispersal_multipliers_fn = "dispersal_multipliers.txt"
#BioGeoBEARS_run_object$areas_allowed_fn = "areas_allowed.txt"
#BioGeoBEARS_run_object$areas_adjacency_fn = "areas_adjacency.txt"
#BioGeoBEARS_run_object$distsfn = "distances_matrix.txt"
# See notes on the distances model on PhyloWiki's BioGeoBEARS updates page.

# Speed options and multicore processing if desired
BioGeoBEARS_run_object$use_optimx = TRUE
BioGeoBEARS_run_object$num_cores_to_use=1

BioGeoBEARS_run_object$force_sparse=FALSE    # sparse=FALSE causes pathology & isn't much faster at this scale
BioGeoBEARS_run_object$speedup=TRUE          # shorcuts to speed ML search; use FALSE if worried (e.g. >3 params)
BioGeoBEARS_run_object$calc_ancprobs=TRUE    # get ancestral states from optim run

# This function loads the dispersal multiplier matrix etc. from the text files into the model object. Required for these to work!
# (It also runs some checks on these inputs for certain errors.)
BioGeoBEARS_run_object = readfiles_BioGeoBEARS_run(BioGeoBEARS_run_object)

# Divide the tree up by timeperiods/strata (uncomment this for stratified analysis)
BioGeoBEARS_run_object = section_the_tree(inputs=BioGeoBEARS_run_object, make_master_table=TRUE, plot_pieces=FALSE)
# The stratified tree is described in this table:
BioGeoBEARS_run_object$master_table

# Good default settings to get ancestral states
BioGeoBEARS_run_object$return_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_TTL_loglike_from_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_ancprobs = TRUE    # get ancestral states from optim run

# Set up DIVALIKE model
# Remove subset-sympatry
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["s","type"] = "fixed"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["s","init"] = 0.0
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["s","est"] = 0.0

BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["ysv","type"] = "2-j"
```

```
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["ys","type"] = "ysv*1/2"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["y","type"] = "ysv*1/2"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["v","type"] = "ysv*1/2"

# Allow classic, widespread vicariance; all events equiprobable
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["mx01v","type"] = "fixed"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["mx01v","init"] = 0.5
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["mx01v","est"] = 0.5

# No jump dispersal/founder-event speciation
# BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","type"] = "free"
# BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","init"] = 0.01
# BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","est"] = 0.01

# Max values for d, e
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","max"] = 400
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","max"] = 400

check_BioGeoBEARS_run(BioGeoBEARS_run_object)

runslow = TRUE
resfn = "Schiz_DIVALIKE_2017.Rdata"
if (runslow)
{
  res = bears_optim_run(BioGeoBEARS_run_object)
  res

  save(res, file=resfn)
  resDIVALIKE = res
} else {
  # Loads to "res"
  load(resfn)
  resDIVALIKE = res
}

#######################################################
# Run DIVALIKE+J
#######################################################
BioGeoBEARS_run_object = define_BioGeoBEARS_run()
BioGeoBEARS_run_object$trfn = trfn
BioGeoBEARS_run_object$geogfn = geogfn
```

```
BioGeoBEARS_run_object$max_range_size = max_range_size
#BioGeoBEARS_run_object$min_branchlength = 0.000001    # Min to treat tip as a direct ancestor (no speciation event)
BioGeoBEARS_run_object$include_null_range = TRUE    # set to FALSE for e.g. DEC* model, DEC*+J, etc.
#(For DEC* and other "*" models, please cite: Massana, Kathryn A.; Beaulieu, Jeremy M.; Matzke, Nicholas J.; O'Meara, Brian C. (2015). Non-null Effects of the Null Range in Biogeographic Models: Exploring Parameter Estimation in the DEC Model. bioRxiv,  http://biorxiv.org/content/early/2015/09/16/026914 )
# also search script on "include_null_range" for other places to change

# Set up a time-stratified analysis:
#BioGeoBEARS_run_object$timesfn = "timeperiods.txt"
#BioGeoBEARS_run_object$dispersal_multipliers_fn = "manual_dispersal_multipliers.txt"
#BioGeoBEARS_run_object$areas_allowed_fn = "areas_allowed.txt"
#BioGeoBEARS_run_object$areas_adjacency_fn = "areas_adjacency.txt"
#BioGeoBEARS_run_object$distsfn = "distances_matrix.txt"
# See notes on the distances model on PhyloWiki's BioGeoBEARS updates page.

# Speed options and multicore processing if desired
BioGeoBEARS_run_object$use_optimx = TRUE
BioGeoBEARS_run_object$num_cores_to_use=1

BioGeoBEARS_run_object$force_sparse=FALSE    # sparse=FALSE causes pathology & isn't much faster at this scale
BioGeoBEARS_run_object$speedup=TRUE          # shorcuts to speed ML search; use FALSE if worried (e.g. >3 params)
BioGeoBEARS_run_object$calc_ancprobs=TRUE    # get ancestral states from optim run

# This function loads the dispersal multiplier matrix etc. from the text files into the model object. Required for these to work!
# (It also runs some checks on these inputs for certain errors.)
BioGeoBEARS_run_object = readfiles_BioGeoBEARS_run(BioGeoBEARS_run_object)

# Divide the tree up by timeperiods/strata (uncomment this for stratified analysis)
BioGeoBEARS_run_object = section_the_tree(inputs=BioGeoBEARS_run_object, make_master_table=TRUE, plot_pieces=FALSE)
# The stratified tree is described in this table:
BioGeoBEARS_run_object$master_table

# Good default settings to get ancestral states
BioGeoBEARS_run_object$return_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_TTL_loglike_from_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_ancprobs = TRUE    # get ancestral states from optim run

# Set up DIVALIKE+J model
# Get the ML parameter values from the 2-parameter nested model
# (this will ensure that the 3-parameter model always does at least as good)
dstart = resDIVALIKE$outputs@params_table["d","est"]
```

```
estart = resDIVALIKE$outputs@params_table["e","est"]
jstart = 0.0001

# Input starting values for d, e
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","init"] = dstart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","est"] = dstart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","init"] = estart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","est"] = estart

# Max values for d, e
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","max"] = 400
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","max"] = 400

# Remove subset-sympatry
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["s","type"] = "fixed"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["s","init"] = 0.0
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["s","est"] = 0.0

BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["ysv","type"] = "2-j"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["ys","type"] = "ysv*1/2"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["y","type"] = "ysv*1/2"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["v","type"] = "ysv*1/2"

# Allow classic, widespread vicariance; all events equiprobable
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["mx01v","type"] = "fixed"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["mx01v","init"] = 0.5
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["mx01v","est"] = 0.5

# Add jump dispersal/founder-event speciation
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","type"] = "free"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","init"] = jstart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","est"] = jstart

# Under DIVALIKE+J, the max of "j" should be 2, not 3 (as is default in DEC+J)
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","min"] = 0.00001
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","max"] = 1.99999

check_BioGeoBEARS_run(BioGeoBEARS_run_object)

resfn = "Schiz_DIVALIKE+J_2017.Rdata"
runslow = TRUE
```

```
if (runslow)
{
 #sourceall("/Dropbox/_njm/__packages/BioGeoBEARS_setup/")

 res = bears_optim_run(BioGeoBEARS_run_object)
 res

 save(res, file=resfn)

 resDIVALIKEj = res
} else {
 # Loads to "res"
 load(resfn)
 resDIVALIKEj = res
}

pdffn = "Schiz_DIVALIKE_vs_DIVALIKE+J_2017.pdf"
pdf(pdffn, width=6, height=6)

#######################################################
# Plot ancestral states - DIVALIKE
#######################################################
analysis_titletxt ="BioGeoBEARS DIVALIKE on Schizorchis"

# Setup
results_object = resDIVALIKE
scriptdir = np(system.file("extdata/a_scripts", package="BioGeoBEARS"))

# States
res2 = plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="text", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

# Pie chart
plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="pie", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

#######################################################
# Plot ancestral states - DIVALIKE+J
#######################################################
analysis_titletxt ="BioGeoBEARS DIVALIKE+J on Schizorchis"
```

```
# Setup
results_object = resDIVALIKEj
scriptdir = np(system.file("extdata/a_scripts", package="BioGeoBEARS"))

# States
res1 = plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="text", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

# Pie chart
plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="pie", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

dev.off()
cmdstr = paste("open ", pdffn, sep="")
system(cmdstr)

#######################################################
#######################################################
# BAYAREALIKE AND BAYAREALIKE+J ANALYSIS
#######################################################
#######################################################
# NOTE: As with DIVA, the BioGeoBEARS BayArea-like model is
# not identical with the full Bayesian model implemented
# in the "BayArea" program of Landis et al. (2013).
#
# Instead, this is a simplified likelihood interpretation
# of the model.  Basically, in BayArea and BioGeoBEARS-BAYAREALIKE,
# "d" and "e" work like they do in the DEC model of Lagrange
# (and BioGeoBEARS), and then BayArea's cladogenesis assumption
# (which is that nothing in particular happens at cladogenesis) is
# replicated by BioGeoBEARS.
#
# This leaves out 3 important things that are in BayArea:
# 1. Distance dependence (you can add this with a distances
#    matrix + the "x" parameter in BioGeoBEARS, however)
# 2. A correction for disallowing "e" events that drive
#    a species extinct (a null geographic range)
# 3. The neat Bayesian sampling of histories, which allows
#    analyses on large numbers of areas.
#
# The main purpose of having a "BAYAREALIKE" model is
```

```
# to test the importance of the cladogenesis model on
# particular datasets. Does it help or hurt the data
# likelihood if there is no special cladogenesis process?
#
# BAYAREALIKE is a likelihood interpretation of BayArea,
# and it is "like BayArea" -- similar to, but not
# identical to, Bayesian BayArea.
# I thus now call the model "BAYAREALIKE", and you should also. ;-)
#######################################################
#######################################################

#######################################################
# Run BAYAREALIKE
#######################################################
BioGeoBEARS_run_object = define_BioGeoBEARS_run()
BioGeoBEARS_run_object$trfn = trfn
BioGeoBEARS_run_object$geogfn = geogfn
BioGeoBEARS_run_object$max_range_size = max_range_size
#BioGeoBEARS_run_object$min_branchlength = 0.000001    # Min to treat tip as a direct ancestor (no speciation event)
BioGeoBEARS_run_object$include_null_range = TRUE    # set to FALSE for e.g. DEC* model, DEC*+J, etc.
#(For DEC* and other "*" models, please cite: Massana, Kathryn A.; Beaulieu, Jeremy M.; Matzke, Nicholas J.; O'Meara, Brian C. (2015). Non-null Effects of the Null Range in Biogeographic Models: Exploring Parameter Estimation in the DEC Model. bioRxiv,  http://biorxiv.org/content/early/2015/09/16/026914 )
# also search script on "include_null_range" for other places to change

# Set up a time-stratified analysis:
BioGeoBEARS_run_object$timesfn = "timeperiods.txt"
BioGeoBEARS_run_object$dispersal_multipliers_fn = "dispersal_multipliers.txt"
#BioGeoBEARS_run_object$areas_allowed_fn = "areas_allowed.txt"
#BioGeoBEARS_run_object$areas_adjacency_fn = "areas_adjacency.txt"
#BioGeoBEARS_run_object$distsfn = "distances_matrix.txt"
# See notes on the distances model on PhyloWiki's BioGeoBEARS updates page.

# Speed options and multicore processing if desired
BioGeoBEARS_run_object$use_optimx = TRUE
BioGeoBEARS_run_object$num_cores_to_use=1

BioGeoBEARS_run_object$force_sparse=FALSE    # sparse=FALSE causes pathology & isn't much faster at this scale
BioGeoBEARS_run_object$speedup=TRUE          # shorcuts to speed ML search; use FALSE if worried (e.g. >3 params)
BioGeoBEARS_run_object$calc_ancprobs=TRUE    # get ancestral states from optim run

# This function loads the dispersal multiplier matrix etc. from the text files into the model object. Required for these to work!
```

```
# (It also runs some checks on these inputs for certain errors.)
BioGeoBEARS_run_object = readfiles_BioGeoBEARS_run(BioGeoBEARS_run_object)

# Divide the tree up by timeperiods/strata (uncomment this for stratified analysis)
BioGeoBEARS_run_object = section_the_tree(inputs=BioGeoBEARS_run_object, make_master_table=TRUE, plot_pieces=FALSE)
# The stratified tree is described in this table:
BioGeoBEARS_run_object$master_table

# Good default settings to get ancestral states
BioGeoBEARS_run_object$return_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_TTL_loglike_from_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_ancprobs = TRUE    # get ancestral states from optim run

# Set up BAYAREALIKE model
# No subset sympatry
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["s","type"] = "fixed"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["s","init"] = 0.0
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["s","est"] = 0.0

# No vicariance
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["v","type"] = "fixed"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["v","init"] = 0.0
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["v","est"] = 0.0

# No jump dispersal/founder-event speciation
# BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","type"] = "free"
# BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","init"] = 0.01
# BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","est"] = 0.01

# Max values for d, e
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","max"] = 400
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","max"] = 400

# Adjust linkage between parameters
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["ysv","type"] = "1-j"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["ys","type"] = "ysv*1/1"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["y","type"] = "1-j"

# Only sympatric/range-copying (y) events allowed, and with
# exact copying (both descendants always the same size as the ancestor)
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["mx01y","type"] = "fixed"
```

```
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["mx01y","init"] = 0.9999
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["mx01y","est"] = 0.9999

# Check the inputs
check_BioGeoBEARS_run(BioGeoBEARS_run_object)

runslow = TRUE
resfn = "Schiz_BAYAREALIKE_2017.Rdata"
if (runslow)
{
  res = bears_optim_run(BioGeoBEARS_run_object)
  res

  save(res, file=resfn)
  resBAYAREALIKE = res
} else {
  # Loads to "res"
  load(resfn)
  resBAYAREALIKE = res
}

#######################################################
# Run BAYAREALIKE+J
#######################################################
BioGeoBEARS_run_object = define_BioGeoBEARS_run()
BioGeoBEARS_run_object$trfn = trfn
BioGeoBEARS_run_object$geogfn = geogfn
BioGeoBEARS_run_object$max_range_size = max_range_size
#BioGeoBEARS_run_object$min_branchlength = 0.000001    # Min to treat tip as a direct ancestor (no speciation event)
BioGeoBEARS_run_object$include_null_range = TRUE    # set to FALSE for e.g. DEC* model, DEC*+J, etc.
#(For DEC* and other "*" models, please cite: Massana, Kathryn A.; Beaulieu, Jeremy M.; Matzke, Nicholas J.; O'Meara, Brian C. (2015). Non-null Effects of the Null Range in Biogeographic Models: Exploring Parameter Estimation in the DEC Model. bioRxiv,  http://biorxiv.org/content/early/2015/09/16/026914 )
# also search script on "include_null_range" for other places to change

# Set up a time-stratified analysis:
BioGeoBEARS_run_object$timesfn = "timeperiods.txt"
BioGeoBEARS_run_object$dispersal_multipliers_fn = "dispersal_multipliers.txt"
#BioGeoBEARS_run_object$areas_allowed_fn = "areas_allowed.txt"
#BioGeoBEARS_run_object$areas_adjacency_fn = "areas_adjacency.txt"
#BioGeoBEARS_run_object$distsfn = "distances_matrix.txt"
# See notes on the distances model on PhyloWiki's BioGeoBEARS updates page.
```

```
# Speed options and multicore processing if desired
BioGeoBEARS_run_object$use_optimx = TRUE
BioGeoBEARS_run_object$num_cores_to_use=1

BioGeoBEARS_run_object$force_sparse=FALSE    # sparse=FALSE causes pathology & isn't much faster at this scale
BioGeoBEARS_run_object$speedup=TRUE          # shorcuts to speed ML search; use FALSE if worried (e.g. >3 params)
BioGeoBEARS_run_object$calc_ancprobs=TRUE    # get ancestral states from optim run

# This function loads the dispersal multiplier matrix etc. from the text files into the model object. Required for these to work!
# (It also runs some checks on these inputs for certain errors.)
BioGeoBEARS_run_object = readfiles_BioGeoBEARS_run(BioGeoBEARS_run_object)

# Divide the tree up by timeperiods/strata (uncomment this for stratified analysis)
BioGeoBEARS_run_object = section_the_tree(inputs=BioGeoBEARS_run_object, make_master_table=TRUE, plot_pieces=FALSE)
# The stratified tree is described in this table:
BioGeoBEARS_run_object$master_table

# Good default settings to get ancestral states
BioGeoBEARS_run_object$return_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_TTL_loglike_from_condlikes_table = TRUE
BioGeoBEARS_run_object$calc_ancprobs = TRUE    # get ancestral states from optim run

# Set up BAYAREALIKE+J model
# Get the ML parameter values from the 2-parameter nested model
# (this will ensure that the 3-parameter model always does at least as good)
dstart = 4.89
estart = 4.89
#dstart = resBAYAREALIKE$outputs@params_table["d","est"]
#estart = resBAYAREALIKE$outputs@params_table["e","est"]
jstart = 0.0001

# Input starting values for d, e
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","init"] = dstart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","est"] = dstart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","init"] = estart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","est"] = estart

# Max values for d, e
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","max"] = 400
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","max"] = 400
```

```
# No subset sympatry
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["s","type"] = "fixed"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["s","init"] = 0.0
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["s","est"] = 0.0

# No vicariance
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["v","type"] = "fixed"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["v","init"] = 0.0
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["v","est"] = 0.0

# *DO* allow jump dispersal/founder-event speciation (set the starting value close to 0)
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","type"] = "free"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","init"] = jstart
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","est"] = jstart

# Under BAYAREALIKE+J, the max of "j" should be 1, not 3 (as is default in DEC+J) or 2 (as in DIVALIKE+J)
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","max"] = 0.99999

# Adjust linkage between parameters
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["ysv","type"] = "1-j"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["ys","type"] = "ysv*1/1"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["y","type"] = "1-j"

# Only sympatric/range-copying (y) events allowed, and with
# exact copying (both descendants always the same size as the ancestor)
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["mx01y","type"] = "fixed"
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["mx01y","init"] = 0.9999
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["mx01y","est"] = 0.9999

# NOTE (NJM, 2014-04): BAYAREALIKE+J seems to crash on some computers, usually Windows
# machines. I can't replicate this on my Mac machines, but it is almost certainly
# just some precision under-run issue, when optim/optimx tries some parameter value
# just below zero.  The "min" and "max" options on each parameter are supposed to
# prevent this, but apparently optim/optimx sometimes go slightly beyond
# these limits.  Anyway, if you get a crash, try raising "min" and lowering "max"
# slightly for each parameter:
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","min"] = 0.0000001
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["d","max"] = 4.9999999

BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","min"] = 0.0000001
```

```
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["e","max"] = 4.9999999

BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","min"] = 0.00001
BioGeoBEARS_run_object$BioGeoBEARS_model_object@params_table["j","max"] = 0.99999

check_BioGeoBEARS_run(BioGeoBEARS_run_object)

resfn = "Schiz_BAYAREALIKE+J_2017.Rdata"
runslow = TRUE
if (runslow)
{
  res = bears_optim_run(BioGeoBEARS_run_object)
  res

  save(res, file=resfn)

  resBAYAREALIKEj = res
} else {
  # Loads to "res"
  load(resfn)
  resBAYAREALIKEj = res
}

pdffn = "Schiz_BAYAREALIKE_vs_BAYAREALIKE+J_2017.pdf"
pdf(pdffn, width=6, height=6)

#######################################################
# Plot ancestral states - BAYAREALIKE
#######################################################
analysis_titletxt ="BioGeoBEARS BAYAREALIKE on Schizorchis"

# Setup
results_object = resBAYAREALIKE
scriptdir = np(system.file("extdata/a_scripts", package="BioGeoBEARS"))

# States
res2 = plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="text", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

# Pie chart
```

```
plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="pie", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

#########################################################
# Plot ancestral states - BAYAREALIKE+J
#########################################################
analysis_titletxt ="BioGeoBEARS BAYAREALIKE+J on Schizorchis"

# Setup
results_object = resBAYAREALIKEj
scriptdir = np(system.file("extdata/a_scripts", package="BioGeoBEARS"))

# States
res1 = plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="text", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

# Pie chart
plot_BioGeoBEARS_results(results_object, analysis_titletxt, addl_params=list("j"), plotwhat="pie", label.offset=0.45, tipcex=0.7, statecex=0.7, splitcex=0.6, titlecex=0.8,
plotsplits=TRUE, cornercoords_loc=scriptdir, include_null_range=TRUE, tr=tr, tipranges=tipranges)

dev.off()
cmdstr = paste("open ", pdffn, sep="")
system(cmdstr)

############################################################################
############################################################################
############################################################################
############################################################################
#
# CALCULATE SUMMARY STATISTICS TO COMPARE
# DEC, DEC+J, DIVALIKE, DIVALIKE+J, BAYAREALIKE, BAYAREALIKE+J
#
############################################################################
############################################################################
############################################################################
############################################################################

############################################################################
############################################################################
# REQUIRED READING:
#
```

```
# Practical advice / notes / basic principles on statistical model
#    comparison in general, and in BioGeoBEARS:
# http://phylo.wikidot.com/advice-on-statistical-model-comparison-in-biogeobears
#########################################################################
#########################################################################

# Set up empty tables to hold the statistical results
restable = NULL
teststable = NULL

#######################################################
# Statistics -- DEC vs. DEC+J
#######################################################
# We have to extract the log-likelihood differently, depending on the
# version of optim/optimx
LnL_2 = get_LnL_from_BioGeoBEARS_results_object(resDEC)
LnL_1 = get_LnL_from_BioGeoBEARS_results_object(resDECj)

numparams1 = 3
numparams2 = 2
stats = AICstats_2models(LnL_1, LnL_2, numparams1, numparams2)
stats

# DEC, null model for Likelihood Ratio Test (LRT)
res2 = extract_params_from_BioGeoBEARS_results_object(results_object=resDEC, returnwhat="table", addl_params=c("j"), paramsstr_digits=4)
# DEC+J, alternative model for Likelihood Ratio Test (LRT)
res1 = extract_params_from_BioGeoBEARS_results_object(results_object=resDECj, returnwhat="table", addl_params=c("j"), paramsstr_digits=4)

# The null hypothesis for a Likelihood Ratio Test (LRT) is that two models
# confer the same likelihood on the data. See: Brian O'Meara's webpage:
# http://www.brianomeara.info/tutorials/aic
# ...for an intro to LRT, AIC, and AICc

rbind(res2, res1)
tmp_tests = conditional_format_table(stats)

restable = rbind(restable, res2, res1)
teststable = rbind(teststable, tmp_tests)

#######################################################
# Statistics -- DEC vs. DEC+J w/o multipliers
```

```
#######################################################
# We have to extract the log-likelihood differently, depending on the
# version of optim/optimx
LnL_2 = get_LnL_from_BioGeoBEARS_results_object(resDECnoMult)
LnL_1 = get_LnL_from_BioGeoBEARS_results_object(resDECjnoMult)

numparams1 = 3
numparams2 = 2
stats = AICstats_2models(LnL_1, LnL_2, numparams1, numparams2)
stats

# DEC, null model for Likelihood Ratio Test (LRT)
res2 = extract_params_from_BioGeoBEARS_results_object(results_object=resDECnoMult, returnwhat="table", addl_params=c("j"), paramsstr_digits=4)
# DEC+J, alternative model for Likelihood Ratio Test (LRT)
res1 = extract_params_from_BioGeoBEARS_results_object(results_object=resDECjnoMult, returnwhat="table", addl_params=c("j"), paramsstr_digits=4)

# The null hypothesis for a Likelihood Ratio Test (LRT) is that two models
# confer the same likelihood on the data. See: Brian O'Meara's webpage:
# http://www.brianomeara.info/tutorials/aic
# ...for an intro to LRT, AIC, and AICc

rbind(res2, res1)
tmp_tests = conditional_format_table(stats)

restable = rbind(restable, res2, res1)
teststable = rbind(teststable, tmp_tests)

#######################################################
# Statistics -- DIVALIKE vs. DIVALIKE+J
#######################################################
# We have to extract the log-likelihood differently, depending on the
# version of optim/optimx
LnL_2 = get_LnL_from_BioGeoBEARS_results_object(resDIVALIKE)
LnL_1 = get_LnL_from_BioGeoBEARS_results_object(resDIVALIKEj)

numparams1 = 3
numparams2 = 2
stats = AICstats_2models(LnL_1, LnL_2, numparams1, numparams2)
stats

# DIVALIKE, null model for Likelihood Ratio Test (LRT)
```

```
res2 = extract_params_from_BioGeoBEARS_results_object(results_object=resDIVALIKE, returnwhat="table", addl_params=c("j"), paramsstr_digits=4)
# DIVALIKE+J, alternative model for Likelihood Ratio Test (LRT)
res1 = extract_params_from_BioGeoBEARS_results_object(results_object=resDIVALIKEj, returnwhat="table", addl_params=c("j"), paramsstr_digits=4)

rbind(res2, res1)
conditional_format_table(stats)

tmp_tests = conditional_format_table(stats)

restable = rbind(restable, res2, res1)
teststable = rbind(teststable, tmp_tests)

#######################################################
# Statistics -- BAYAREALIKE vs. BAYAREALIKE+J
#######################################################
# We have to extract the log-likelihood differently, depending on the
# version of optim/optimx
LnL_2 = get_LnL_from_BioGeoBEARS_results_object(resBAYAREALIKE)
LnL_1 = get_LnL_from_BioGeoBEARS_results_object(resBAYAREALIKEj)

numparams1 = 3
numparams2 = 2
stats = AICstats_2models(LnL_1, LnL_2, numparams1, numparams2)
stats

# BAYAREALIKE, null model for Likelihood Ratio Test (LRT)
res2 = extract_params_from_BioGeoBEARS_results_object(results_object=resBAYAREALIKE, returnwhat="table", addl_params=c("j"), paramsstr_digits=4)
# BAYAREALIKE+J, alternative model for Likelihood Ratio Test (LRT)
res1 = extract_params_from_BioGeoBEARS_results_object(results_object=resBAYAREALIKEj, returnwhat="table", addl_params=c("j"), paramsstr_digits=4)

rbind(res2, res1)
conditional_format_table(stats)

tmp_tests = conditional_format_table(stats)

restable = rbind(restable, res2, res1)
teststable = rbind(teststable, tmp_tests)

#########################################################################
# RESULTS: DEC, DEC+J, DIVALIKE, DIVALIKE+J, BAYAREALIKE, BAYAREALIKE+J"DECnoMult", "DEC+JnoMult"
#########################################################################
```

```
teststable$alt = c("DEC+J", "DEC+JnoMult", "DIVALIKE+J", "BAYAREALIKE+J")
teststable$null = c("DEC", "DEC noMult", "DIVALIKE", "BAYAREALIKE")
row.names(restable) = c("DEC", "DEC+J","DECnoMult", "DEC+JnoMult", "DIVALIKE", "DIVALIKE+J", "BAYAREALIKE", "BAYAREALIKE+J")

# Look at the results!!
restable
teststable

########################################################
# Save the results tables for later -- check for e.g.
# convergence issues
########################################################

# Loads to "restable"
save(restable, file="restable_v1.Rdata")
load(file="restable_v1.Rdata")

# Loads to "teststable"
save(teststable, file="teststable_v1.Rdata")
load(file="teststable_v1.Rdata")

# Also save to text files
write.table(restable, file="restable.txt", quote=FALSE, sep="\t")
write.table(unlist_df(teststable), file="teststable.txt", quote=FALSE, sep="\t")

########################################################
# Model weights of all six models
########################################################
restable2 = restable

# With AICs:
AICtable = calc_AIC_column(LnL_vals=restable$LnL, nparam_vals=restable$numparams)
restable = cbind(restable, AICtable)
restable_AIC_rellike = AkaikeWeights_on_summary_table(restable=restable, colname_to_use="AIC")
restable_AIC_rellike

# With AICcs -- factors in sample size
samplesize = length(tr$tip.label)
AICtable = calc_AICc_column(LnL_vals=restable$LnL, nparam_vals=restable$numparams, samplesize=samplesize)
restable2 = cbind(restable2, AICtable)
restable_AICc_rellike = AkaikeWeights_on_summary_table(restable=restable2, colname_to_use="AIC")
```

```
restable_AICc_rellike
free_params = row.names(resDECj$output@params_table[resDECj$output@params_table$type=="free",])
names(restable_AICc_rellike) = c("LnL", "numparams", free_params, "AICc", "AICc_wt")

# Also save to text files
write.table(restable_AIC_rellike, file="restable_AIC_rellike.txt", quote=FALSE, sep="\t")
write.table(restable_AICc_rellike, file="restable_AICc_rellike.txt", quote=FALSE, sep="\t")

# Save with nice conditional formatting
write.table(conditional_format_table(restable_AIC_rellike), file="restable_AIC_rellike_formatted.txt", quote=FALSE, sep="\t")
write.table(conditional_format_table(restable_AICc_rellike), file="restable_AICc_rellike_formatted.txt", quote=FALSE, sep="\t")
```