Biological Journal of the Linnean Society, 2021, **133**, 817–834. With 6 figures.

Comparative phylogeography of two Northern Rocky Mountain endemics: the widespread *Anguispira kochi occidentalis* and the narrow-range *Anguispira nimapuna* (Gastropoda: Discidae)

ANDREW M. RANKIN[1,2,*], FRANK E. ANDERSON[3], STEPHANIE A. CLUTTS[3], ANAHÍ ESPÍNDOLA[4], BRYAN C. CARSTENS[5], MICHAEL LUCID[6], and JACK SULLIVAN[1,2]

[1]*Department of Biological Sciences, University of Idaho, Moscow, ID 83844, USA*
[2]*Institute for Bioinformatics and Evolutionary Studies (IBEST), Department of Biological Sciences, University of Idaho, Moscow, ID 83844, USA*
[3]*School of Biological Sciences, Southern Illinois University, Carbondale, IL 62901, USA*
[4]*Department of Entomology, University of Maryland, College Park, 4291 Fieldhouse Drive, 3138 Plant Sciences Building, College Park, MD 20742–4454, USA*
[5]*Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, Columbus, OH 43210, USA*
[6]*PO Box 733, Sandpoint, ID 83864, USA*

Received 29 October 2020; revised 8 February 2021; accepted for publication 12 February 2021

The Northern Rocky Mountain ecosystem supports rich biological diversity with many endemic and rare species. Extant endemics display two biogeographic patterns: widespread species with fragmented populations, and narrow-range endemics. These distributions are shown by the congeneric snails *Anguispira kochi occidentalis* and *Anguispira nimapuna*. These two taxa are disjunct from the remaining species of the genus, which achieves its greatest diversity in eastern North America. Given the disjunct nature of *A. k. occidentalis* and *A. nimapuna*, we here present a mtDNA phylogeny of the genus that includes both eastern and western species to assess the phylogenetic position of *A. k. occidentalis* and *A. nimapuna*. We then reconstruct the demographic history of *A. k. occidentalis* and *A. nimapuna* by analysing current patterns of genetic variation and interpreting the results considering the historical biogeography of the region. Both *A. k. occidentalis* and *A. nimapuna* represent unique taxa that are genetically and geographically distinct from their congeners. The current distribution and genetic structure of *A. k. occidentalis* has been shaped by both historical isolation in refugia and more recent northward shifts, whereas *A. nimapuna* is represented by two populations with shallow divergence in an area of long-term habitat stability.

ADDITIONAL KEYWORDS: *Anguispira* – demography – endemic – Northern Rocky Mountains – phylogeny.

INTRODUCTION

The Northern Rocky Mountain (NRM) mesic forest ecosystem is a relatively small biogeographic region that ranges from south-eastern British Columbia south through the Idaho Panhandle, into north-central Idaho and the north-eastern corners of Washington and Oregon, and west of the continental divide in Montana. Scattered throughout are discontinuous bands of mesic forest patches [sometimes referred to as inland rainforest (DellaSala, 2011)] that support a rich biodiversity of vascular and non-vascular plants, small-bodied vertebrates and litter-dwelling organisms. Many of these appear to represent relicts, as many NRM species have disjunct relatives in the forests of the Pacific Coast (e.g. *Ascaphus*, *Prophysaon* and *Hemphillia*) and eastern North America (e.g. the Polygyridae and Plethodontidae). The current

*Corresponding author. E-mail: rank3288@gmail.com

NRM biota includes two contrasting biogeographic patterns: (1) widespread species with populations distributed allopatrically in mid-elevation bands separated by stretches of dryer, lower-elevation habitat; (2) regional endemics confined to discrete patches of suitable habitat (Brunsfeld *et al.*, 2001).

For widespread NRM taxa, it has been hypothesized that recurring glacial cycles in concert with altitudinal heterogeneity had a significant effect on the demographic history of regional populations (Brunsfeld *et al.*, 2001). During glacial periods, multiple populations were fragmented in allopatric refugia across a topographically complex area [i.e. multiple inland refugia hypothesis; see Brunsfeld *et al.* (2001)]. The current distribution of widespread taxa would therefore have been achieved by subsequent post-glacial secondary contact among lineages, with the northernmost populations reflecting recent postglacial expansion. Phylogeographic studies of several NRM mesic species have indeed indicated populations survived in isolation long enough for genetic differences to be preserved among separate refugia [i.e. there was more than a single Pleistocene refugium; tailed frogs (Nielson *et al.*, 2001); Constance's bittercress (Brunsfeld & Sullivan, 2005)]. Also present in this ecosystem are narrow-range endemics currently confined to discrete patches of suitable habitat. Numerous rare insects, gastropods, aquatic invertebrates and wetland angiosperms are only known to occur in certain regions that were safe from glaciation during the Pleistocene such as within the Clearwater and Lolo National Forests (Fend & Gustafson, 2001; Stark & Gustafson, 2004; Stagliano *et al.*, 2007; Newell *et al.*, 2008; Stagliano, 2016). Specifically, the Clearwater drainage—a low-elevation tributary canyon of the Snake River in north-central Idaho, rich with moist, old-growth forest floors and higher-order tributaries—harbours many narrow-range endemics that appear to represent relicts that survived regional climatic changes in this protected canyon (Lichthardt & Moseley, 1994) but did not subsequently recolonize other parts of the ecosystem following glacial retreat.

Tiger snails of the genus *Anguispira* (family Discidae) are common, terrestrial snails that occupy moist, forested areas across portions of eastern and western North America. They achieve their greatest diversity in the east, where 12 of the 13 recognized species occur. In the west, only two forms occur: the subspecific *A. kochi occidentalis* and *A. nimapuna*. Specifically, *A. k. occidentalis* (Von Martens, 1882) is widespread in the moist, mixed coniferous forest of the NRM, whereas the type *Anguispira kochi kochi* (Pfeiffer, 1845) is a species of eastern North America. However, the original taxonomic description and subspecific designation of *A. k. occidentalis* was based primarily on subjective descriptions of shell characters that may be prone to homoplasy (e.g. Emberton, 1991a, 1995). Conversely, *A. nimapuna* (Baker, 1932), a narrow-range endemic known only from a few watersheds in Idaho County, Idaho, has a G1 global conservation status rank [Critically Imperiled (NatureServe, 2017)] and an S3 state rank [Species of Greatest Conservation Need Tier 3 (idfg.idaho.gov, retrieved Jan 15, 2021)] due to its rarity. The distribution of *A. nimapuna* is restricted to a small portion of the Clearwater drainage at the confluence of the Lochsa and Selway rivers, a V-shaped unglaciated valley where elevations rise abruptly from the river bottom to form steep slopes (Lichthardt & Moseley, 1994). *A. nimapuna* may represent a relictual population that has persisted in an ancient refugium but has not expanded, perhaps due to the surrounding topological barriers.

Given the disjunction of *A. k. occidentalis* and *A. nimapuna* from their congeners—separated by ~2000 km from eastern populations—and the importance of our understanding of hierarchical evolutionary relationships, determining the phylogenetic position of these two species with respect to other members of *Anguispira* is critical. However, the only phylogenetic study on *Anguispira* published to date (Haskell & Pan, 2013) was limited to species/populations in the Cumberland Plateau region (southern Tennessee and northern Alabama) and did not include western forms. Here we present a mitochondrial DNA phylogeny of the genus with both eastern and western species to determine the phylogenetic position of *A. k. occidentalis* and *A. nimapuna*. After inferring this phylogeny, we focus on reconstructing the demographic history of *A. k. occidentalis* and *A. nimapuna* by analysing current patterns of genetic variation and diversity of intraspecific lineages and interpreting the results considering the historical biogeography of the NRM region. Because contemporary patterns of genetic variation and distribution of populations are often the result of historical biogeographic processes, comparing the phylogeographic histories of species may highlight the different or similar effects that the biogeography of a region may have for species with contrasting distribution patterns. We first apply a traditional phylogeographic approach—reconstructing the history of the two species by assessing intraspecific patterns of diversity—then use coalescent simulations to compare alternative models of population history, and lastly evaluate our interpretations considering Species Distribution Models (SDM) produced for both species under

current and Last Glacial Maximum (LGM) climate conditions using survey-presence records.

MATERIAL AND METHODS

Sampling and tissue acquisition

We obtained sequence data from 120 *A. k. occidentalis* and 60 *A. nimapuna* specimens, the majority of which were obtained by field collection of live snails. To these, we added tissue samples from the California Academy of Sciences, Carnegie Museum of Natural History, Delaware Museum of Natural History, Field Museum of Natural History, Montana Natural Heritage Program, Royal British Columbia Museum and University of Florida Museum. Additional sequence data from 176 individuals representing nine of the remaining 11 currently recognized *Anguispira* species [missing are *Anguispira knoxensis* and *Anguispira rugoderma* (Integrated Taxonomic Information System; http://www.itis.gov; retrieved June 18, 2020)] were obtained from GenBank, including both unpublished and published (Haskell & Pan, 2013; accession numbers: JN544577 - JN544696) data, as well as data from an unpublished thesis (Clutts, 2008). To serve as confamilial phylogenetic comparison, sequence data from 96 individuals representing seven of the 12 currently recognized *Discus* species was obtained from museums, field collections and GenBank (accession numbers: AF063140, AF063141, AF064406 - AF064409, AF064415 - AF064438, FJ917285, FJ969586 - FJ969703, JX911298, KM611953, KM612165, MF544381, MF544595, MF544774, MF544853, MF545101, MG421413, MG421515, MG421676, MG422014, MG422282, MG422422, MG422625, MG422626, MG422738, MG422943, MG423043, MG423120, MG423227, MG423525). For new data generated for this study and Clutts (2008), see Supporting Information (Data S1) for details of specimen IDs, localities, museum catalogue numbers and GenBank accession numbers. Following Nordsieck (1986) and Emberton (1991b), we included the Oreohelicidae (*Oreohelix* and *Radiocentrum*) as a possible sister group as well as the Megomphicidae (*Polygyrella* and *Megomphix*) as a presumably more distantly related outgroup to root the discid phylogeny.

Laboratory work and data generation

Genomic DNA was extracted from tissue using either an E.Z.N.A. Blood and Tissue Kit (Omega Bio-tek, Doraville, CA, USA; this study) or a DNeasy Extraction Kit (Qiagen, Valencia, CA, USA; Clutts, 2008) as per the manufacturer's protocols. Regions of three mitochondrial genes—cytochrome *c* oxidase subunit I (*COI*), cytochrome b (*cytb*) and 16S ribosomal RNA (16S)—were amplified via polymerase chain reaction (PCR) with universal metazoan and molluscan primer pairs: *COI* (Folmer *et al.*, 1994), 16S [Lucid *et al.* (2018) for this study and Geller *et al.* (1997) for Clutts (2008)] and *cytb* (Merritt *et al.*, 1998). These genes were chosen to match data available from Clutts (2008) and in GenBank. Amplifications for newly generated data (this study) were performed in 25 μL reactions containing 2 μL DNA extract, 18 μL H_2O, 2.5 μL buffer, 0.75 μL 50 mM $MgCl_2$, 0.5 μL 10 mM dNTPs, 0.5 μL 10 mM forward primer, 0.5 μL 10 mM reverse primer and 1.25 U New England Biolabs Taq polymerase, while Clutts (2008) used HotStart Master Mix (Qiagen) following the manufacturer's protocols (half-volume reactions). Thermal cycling profiles included an initial denaturation step at 95° C for 2 min, followed by 30–35 cycles of denaturation (95° C for 45 s), primer annealing (45 s), extension (72° C for 60 s), followed by a terminal extension cycle of 72° C for 7 min. Primer annealing temperatures were: *COI* (52 °C); 16S (50 °C); *cytb* (48 °C). Amplicons were then purified using the QiaQuick PCR Cleanup Kit (Qiagen; this study) or were gel-purified from 1% agarose gels using the QiaQuick Gel Extraction Kit (Qiagen; Clutts, 2008). Bi-directional DNA Sanger sequencing was outsourced to Eurofins MWG Operon, Louisvillle, KY, USA (http://www.eurofins.fr; this study) or cycle-sequenced using the BigDye Terminator Cycle Sequencing Ready Reaction Kit v.3.0 (Applied Biosystems, CA) and cleaned on Sephadex spin columns (Princeton Separations, Adelphia, NJ) and run out on an ABI Prism 377 (Applied Biosystems) automated DNA sequencer at Southern Illinois University (Clutts, 2008). Chromatograms in both directions were compared and consensus sequences were assembled using either Chromas v.2.6.2 [Technelysium, http://www.technelysium.com.au/chromas.html (this study)] or Sequencher v.3.0 [Gene Codes, Ann Arbor, http://www.genecodes.com/ (Clutts, 2008)].

Because *COI*, *cytb* and 16S genes each reside on the same mitochondrial DNA molecule and are inherited as one linkage group, they share a genealogical history and may share a single gene tree. Therefore, the three mitochondrial gene segments were concatenated into a single data set. We also explored partitioning sites into seven partitions, *COI* codon positions+*cytb* codon positions+16S; however, this partitioning scheme did not influence the results (e.g. poorly supported nodes were still poorly supported and highly supported nodes were still highly supported), so we focussed our discussion to results of a single partition. From these data, we produced three separate data matrices, including intraspecific sets for both *A. k. occidentalis* and *A. nimapuna*, as well as an interspecific set that contained all unique

haplotypes (across the concatenated *COI-cytb*-16S matrix) of the remaining *Anguispira* species, *Discus* representatives and other outgroups. For each data matrix, multiple sequence alignments were constructed using MAFFT online with the default FFT-NS-2 option (http://www.ebi.ac.uk/Tools/msa/mafft/). For the interspecific data matrix only, several alignment-ambiguous loop regions of the 16S gene were too divergent to be aligned across lineages and we therefore used the Gblocks Server v.0.91b algorithm with the least restrictive settings available [(Castresana, 2000); http://molevol.cmima.csic.es/castresana/Gblocks_server.html] to exclude these regions from phylogenetic inference.

Interspecific phylogeny

We analysed the interspecific data matrix under maximum parsimony (MP), maximum likelihood (ML) and Bayesian frameworks. An MP analysis was performed in PAUP* (Swofford, 2003) using the heuristic search algorithm with ten random-addition replicates, tree bisection and reconnection (TBR) swapping and all nucleotide substitutions treated as equal weight. Nodal support was estimated with 500 parsimony bootstrap replicates (Felsenstein, 1985). For the model-based ML analysis, we first selected a model of nucleotide sequence evolution via the automodel command in PAUP* under the Bayesian Information Criterion (BIC) and decision theory [DT (Minin *et al.*, 2003)]. ML analyses were performed in Garli (Zwickl, 2006) using the GTR+I+Γ model (determined to be best fitting by PAUP*; with parameters estimated in Garli), conducting 100 replicate runs with random starting trees. Nodal support was assessed using 200 bootstrap replicates with two tree searches per bootstrap. We used the resulting phylogeny to test the assumption that the data set has evolved in a clock-like fashion by testing for a global molecular clock in PAUP* using the likelihood-ratio test of Felsenstein (1988). Then, a Bayesian analysis was performed with BEAST v.2.4.4 (Bouckaert *et al.*, 2014) using the same GTR+I+Γ model determined earlier (with parameters estimated in BEAST) as well as a relaxed lognormal molecular clock (based on results of clock analysis; Likelihood ratio test (LRT) = 822.3, d.f. = 328, $P << 0.01$) and a birth-death speciation tree prior [the birth-death prior has been shown to outperform the Yule prior when analysing mixed inter- and intraspecies data sets (Ritchie *et al.*, 2017)]. Posterior probabilities were estimated and used to assess support for each branch in the inferred phylogeny. Because we lacked fossils for calibration and recognizing that molecular clock estimates are often dubious (Hillis *et al.*, 1996), we fixed the mean substitution rate to a value of 1 so that branch lengths would be reported in units of substitutions per site. Results were viewed in Tracer v.1.7 (Rambaut & Drummond, 2007) to ensure all parameters had converged and Effective Sample Size (ESS) values for all parameters exceeded 200, and maximum clade credibility trees were produced with the BEAST application TreeAnnotator.

Intraspecific phylogeny

We used a phylogeny-based framework to identify possible phylogeographic structure within *A. k. occidentalis* and *A. nimapuna*. For each data set, we first selected an appropriate model of nucleotide substitution using the automodel command in PAUP* under the BIC and DT (the best fitting models were HKY+I+Γ for *A. k. occidentalis* and HKY+I for *A. nimapuna*). Then, a ML tree was determined (from 100 replicate runs) for only unique haplotypes (across the concatenated *COI-cytb*-16S matrix; $N = 64$ for *A. k. occidentalis* and $N = 40$ for *A. nimapuna*) using Garli (Zwickl, 2006), and nodal support was assessed based on 200 bootstrap replicates (five tree searches per bootstrap). Next, we estimated trees under Bayesian inference with BEAST 2.4.4. Using all alleles ($N = 120$ for *A. k. occidentalis* and $N = 60$ for *A. nimapuna*), we specified both a coalescent constant population size tree prior and a Bayesian skyline tree prior with the same substitution models mentioned above. The skyline analysis attempts to assess putative fluctuations in effective population size over time, which is done by estimating the posterior distribution for effective population size at intervals along the phylogeny (see Ho & Shapiro, 2011). Each BEAST analysis was run for 100 million generations and samples were drawn every 10 000 generations, and we discarded 2500 samples from each run as burn-in. Results from each run were viewed in Tracer v.1.7 (Rambaut & Drummond, 2007) to ensure all parameters had converged and Effective Sample Size (ESS) values for all parameters exceeded 200, and maximum clade credibility trees were produced with the BEAST application Tree Annotator. Bayesian skyline plots (BSP) were also reconstructed using Tracer.

Intraspecific polymorphism

To estimate the extent of mtDNA variation, we calculated two intraspecific diversity statistics using MEGA6 (Tamura *et al.*, 2013): the Watterson estimator [θ, estimated from the number of segregating sites (Watterson, 1975)] and nucleotide diversity [π, mean number of pairwise differences per site (Nei & Li, 1979)]. We then examined the behaviour of these estimates as a function of sample

size by calculating them for each n, where n is the number of sequences, and plotting the results of ten replicates. We randomly selected n samples from a data set, calculated θ and π, and repeated this ten times for each n. Because both θ and π are unbiased estimates of neutral diversity, the neutral model predicts that $\theta \approx \pi$; however, this assumption fails by violation of the infinite-sites model and under various influences of selection and demographic history (Tajima, 1989a, b). Therefore, θ and π can be compared to infer the types of variants present in each sample (e. g., when there are a lot of singletons, the latter underestimates polymorphisms compared to the former, whereas when there are many alleles at intermediate frequencies, pairwise differences are inflated when compared to segregating sites). To do so, we performed Tajima's D intraspecific polymorphism statistical test for neutrality to test the prediction that $\theta \approx \pi$ using Arlequin v.3.5 (Excoffier *et al.*, 2007), running 1000 coalescent simulations under the infinite-sites model to test the significance of D.

Nucleotide diversity (π) is a useful measure of the extent of DNA polymorphism. However, the variance (σ^2) of an estimate of π has historically not been well studied (Nei & Jin, 1989) despite σ^2 being a potential source of information of demographic history (Wakeley, 1996a, 1996b). For example, consider two species having the same mean π but one species is panmictic whereas the other species is divided into several subpopulations between which there may or may not be genetic exchange. This may cause the two species to have different variances of π. Therefore, we quantified the variance of π for both *A. k. occidentalis* and *A. nimapuna* and compared those values to simulated data sets that fitted a single, neutrally evolving population of constant size, i.e. we compared the variance of π to the variance of a simulated population having the same mean of π. We quantified variance by calculating π for 200 randomly generated subsamples of size $n = 20$, plotted the distribution of values, and qualitatively compared the actual and simulated distributions against each other.

To estimate the most likely number of genetically differentiated clusters present in each species, a Bayesian Analyses of Population Structure (BAPS) was performed in BAPS 6.0 (Corander *et al.*, 2008). BAPS performs a *clustering with linked loci* (codon linkage model) analysis that takes into consideration the non-independence of linked loci and attempts to infer the most likely number of putative genetic clusters (K) by maximizing the Hardy-Weinberg equilibrium amongst clusters [i.e. minimizing the Walhund effect (1928)] through a stochastic learning algorithm. We estimated the probability of a different number of genetic clusters ($K = 1$ to 15) with ten runs for each K, and the clustering of groups with the lowest log likelihood was selected.

Support for evolutionary scenarios estimated using ABC

Although descriptive analyses of genetic variation and phylogenies are useful to identify patterns and compare hypotheses, Approximate Bayesian Computation (ABC) allows for the quantitative comparisons of alternative scenarios via simulation and estimation of the posterior distributions of important parameters. Such simulation approaches, which rely on implied assumptions of the many parameters, are therefore valuable when used in conjunction with other methods that do not rely on highly parameterized models. Thus, to better understand the evolutionary history of *A. k. occidentalis* and *A. nimapuna*, we performed coalescent simulations in an ABC framework using DIYABC v.1.0.4.46 (Cornuet *et al.*, 2008), simulating many thousands of genealogies and retaining those simulations that produced genetic variation patterns close to the empirical data which were then used to discriminate among a set of alternative historical scenarios.

For *A. k. occidentalis*, we partitioned the data into five clusters that were recovered by the BAPS analysis and compared six alternative scenarios that considered divergence and population size variation (Fig. 1A). The *A. nimapuna* data were partitioned into two clusters that were recovered by BAPS and we compared five alternative scenarios (Fig. 1B) that considered population size variation. The similarity between the simulations and the empirical data was measured using both within- and between-population summary statistics, including number of segregating sites, mean pairwise differences, Tajima's D, and private segregating sites for a single population, and number of segregating sites, mean pairwise differences within, and mean pairwise differences between pairs of populations. For each species, 100 000 simulated data sets were generated for each scenario to build a reference under a mutation model with mean rate ranging from 1.00×10^{-09} to 1.00×10^{-07} and uniform prior distribution. A pre-evaluation step based on a principal component analysis (PCA) was performed to ensure scenarios and priors produced simulated data sets similar enough to the empirical data. The relative posterior probabilities of the competing scenarios were estimated via logistic regression on the 10% of simulated data sets closest to the empirical (Cornuet *et al.*, 2010). The model with the highest posterior probability was considered the best model.

Species distribution models

To examine the potential relationship between genetic structure and both current predicted habitat and

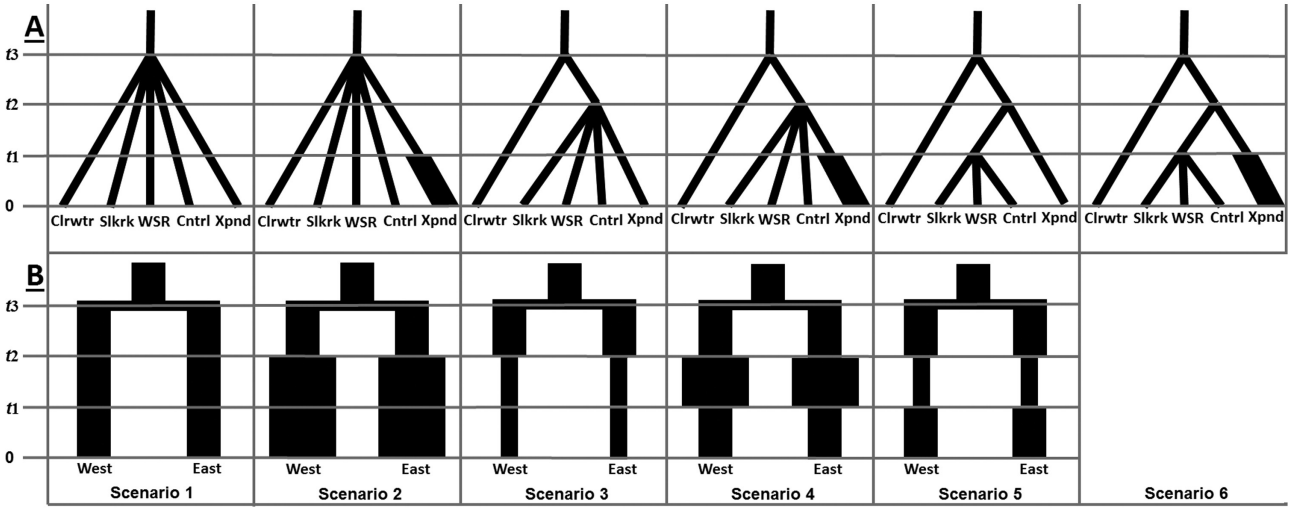

Figure 1. Simulated historical scenarios tested in DIYABC for (A) five *A. k. occidentalis* BAPS clusters and (B) two *A. nimapuna* BAPS clusters. In these scenarios, *t* represents timescale in terms of the number of generations and width of the graph represents relative effective population size during the time period (e.g. 0 – *t*1).

predicted historical habitat, we prepared ensemble Species Distribution Models [SDMs (Peterson, 2011)] for *A. k. occidentalis* and *A. nimapuna* under both current and LGM (~21 500 years ago) climate conditions. This approach allowed us to evaluate how temporal patterns of habitat availability could explain patterns of current genetic structure. We compiled georeferenced locality records for *A. k. occidentalis* and *A. nimapuna* by combining specimen locales from this study and localities collected during field surveys conducted by the Multi-species Baseline Initiative (see https://idfg.idaho.gov/baseline; Lucid, 2018). The modelling was done using 19 standard bioclimatic variables obtained from WORLDCLIM (Hijmans *et al.*, 2005; resolution: 2.5 arc min) and a working area polygon of 72°–20° latitude and 180°–100° longitude. We set the working area polygon to include only western North America to avoid sampling background points from areas that fall well outside the species' known geographic distributions. We constructed SDMs using the ensemble method implemented in the R (R Development Core Team, 2014) package biomod2 (Thuiller *et al.*, 2009). For this we calculated niche models using nine techniques available in biomod2: Generalized Linear Models (GLM), Generalized Boosting Models (GBM), Generalized Additive Models (GAM), Classification Tree Analysis (CTA), Artificial Neural Network (ANN), Surface Range Envelop (SRE), Flexible Discriminant Analysis (FDA), Multiple Adaptive Regression Splines (MARS), Random Forest (RF) and then ran an ensemble approach by using the area under the curve (AUC) values as a weight for the contribution of each individual model to the final ensemble model, forecasting the ensemble model onto current climate conditions, and also projecting a palaeodistribution using the same methods estimated for the LGM under the community climate system model [CCM3 (Collins *et al.*, 2006)].

RESULTS

Interspecific phylogeny

The interspecific data matrix consisted of 1374 bp and, excluding outgroups, included 665 variable sites, 617 of which were parsimony informative and 48 were singletons. The maximum parsimony (Supporting Information, Fig. S1), maximum likelihood (Fig. 2) and Bayesian (Supporting Information, Fig. S2) trees resulted in slightly different topologies; however, three general concordant results were revealed: the Discidae is monophyletic (MP bootstrap 91; ML bootstrap 98; Posterior Probability 0.96); the eastern *Anguispira* species (apart from *A. k. kochi*) form a strongly supported monophyletic group (MP 87; ML 83; PP 0.97) and *A. k. kochi*+*A. k. occidentalis* grouped together, though the support is moderate (MP < 70; ML < 70; PP 0.98).

The parsimony analysis produced a single most parsimonious tree. *A. k. kochi* and *A. k. occidentalis* group together and are distantly related to the eastern *Anguispira* species, and *A. nimapuna* clusters with *Discus* species (with < 70 bootstrap support). In the likelihood analysis, all *Angusipira* species cluster together with < 70 bootstrap support. *Anguispira nimapuna* is sister to a clade comprising the *A. kochi*

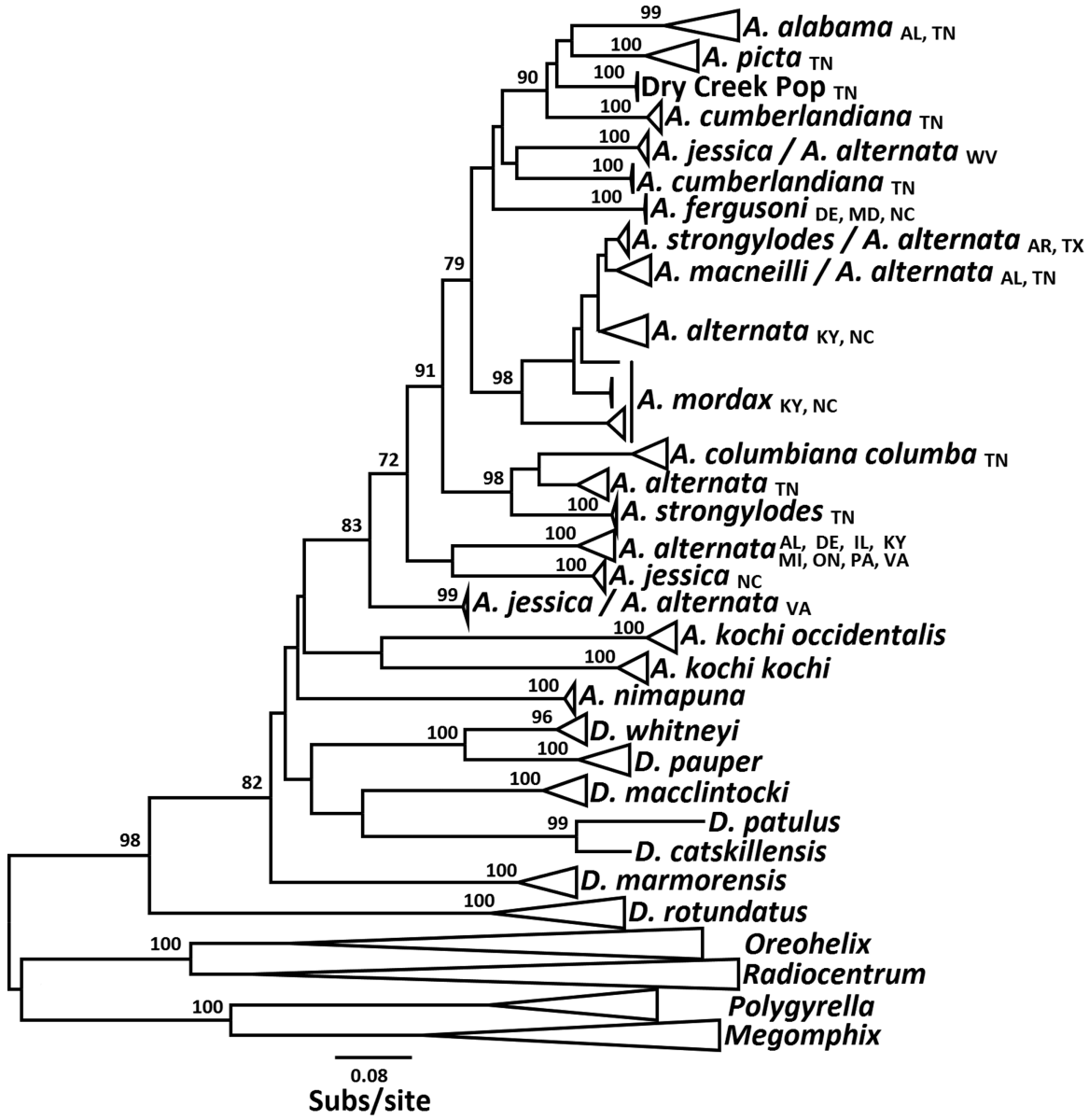

Figure 2. Best maximum likelihood phylogeny for *Anguispira* and *Discus* species based on concatenated *COI*, *cytb* and 16S mtDNA sequences. Bootstrap values on nodes indicate relationships that are well supported (≥ 70).

clade and a well-supported eastern *Anguispira* clade. In the Bayesian analysis, the *A. kochi* samples are monophyletic, and, along with *A. nimapuna*, cluster with *Discus* species with < 0.90 posterior probability. Thus, the main difference among analyses is whether *Anguispira* as a whole is monophyletic.

A. K. OCCIDENTALIS PHYLOGENY AND POLYMORPHISM

The *A. k. occidentalis* data set consisted of 120 sequences of 1486 bp. There were 250 variables sites, 189 of which were parsimony informative and 61 singletons, and 64 unique haplotypes, and for the protein-coding sections (*COI* and *cytb*) substitutions occurred primarily at third positions and synonymous substitutions outnumbered non-synonymous. Phylogenetic analyses were comparable between maximum likelihood (Fig. 3) and Bayesian trees (Fig. 4). There are two well-supported clades: a small clade composed of only Clearwater drainage individuals and a much larger clade composed of all remaining *A. k. occidentalis* individuals. The latter group of individuals are partitioned into subgroups in both

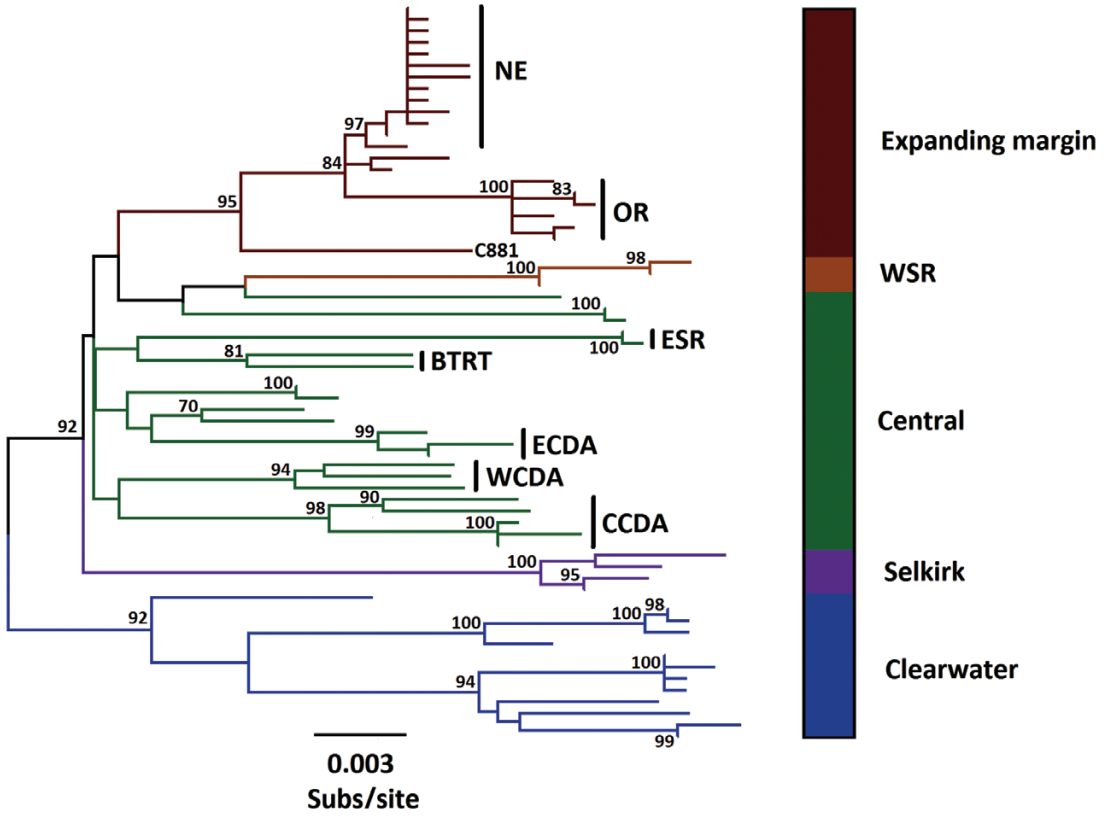

Figure 3. Best maximum likelihood phylogeny for *A. k. occidentalis* based on mtDNA haplotypes. Bootstrap values on nodes indicate relationships that are well supported (≥ 70). BAPS clusters that correspond to the best partition of the data are indicated by different coloured bars.

trees. However, the deepest splits among these clusters are (generally) not well supported. Highly supported nodes include a clade of individuals collected from the southern tip of the Selkirk Mountains, three distantly related clades distributed allopatrically across the Coeur d'Alene (CDA) Mountains of Idaho and Montana (hereafter referred to as western, central and eastern CDA [WCDA], CCDA and ECDA), a pair of localities in the Bitterroot Mountains (BTRT), two distantly related clades distributed on either side of the Salmon River in west-central Idaho (near Riggins) [hereafter referred to as west Salmon River (WSR) and east Salmon River (ESR)], a clade of Oregon individuals, and a genetically depauperate group of individuals collected from northern and eastern sampling locations (NE) (see Figs 3, 4).

Uncorrected nucleotide diversity (π) was 0.019 while the Watterson estimator (θ) was higher at 0.031 (Supporting Information, Fig. S3), although this difference is not significant: Tajima's D = -1.28 ($P > 0.10$). However, we repeated Tajima's D for a single phylogroup (the NE clade), and this clade had Tajima's D = -2.44 ($P < 0.05$), in which case we conclude there is a significant excess of low-frequency polymorphisms in that clade. We compared the distribution curves of π (Supporting Information, Fig. S4) for the *A. k. occidentalis* data set to that of a simulated data set (of a single, neutrally evolving population) produced to have an identical mean. The two distribution curves displayed identical means but qualitatively different variabilities. That is, the variance of the *A. k. occidentalis* is greater, suggesting that *A. k. occidentalis* is likely structured as a group of populations spatially separated across its range.

BAPS clustering analysis found the best partition of data to be K = 5 populations (Fig. 4). Three clusters are composed of samples from small geographic areas and correspond to highly supported clades: Clearwater (N = 15), Selkirk (N = 5) and WSR individuals (N = 9). A fourth cluster (N = 28) includes individuals sampled across the central part of Idaho's panhandle (hereafter referred to as the Central cluster) and a couple of locations in Montana, and includes the WCDA, CCDA, ECDA, BTRT and ESR clades. The fifth cluster (N = 63) comprises individuals collected near the edges of the range, including north Idaho, Montana, British Columbia

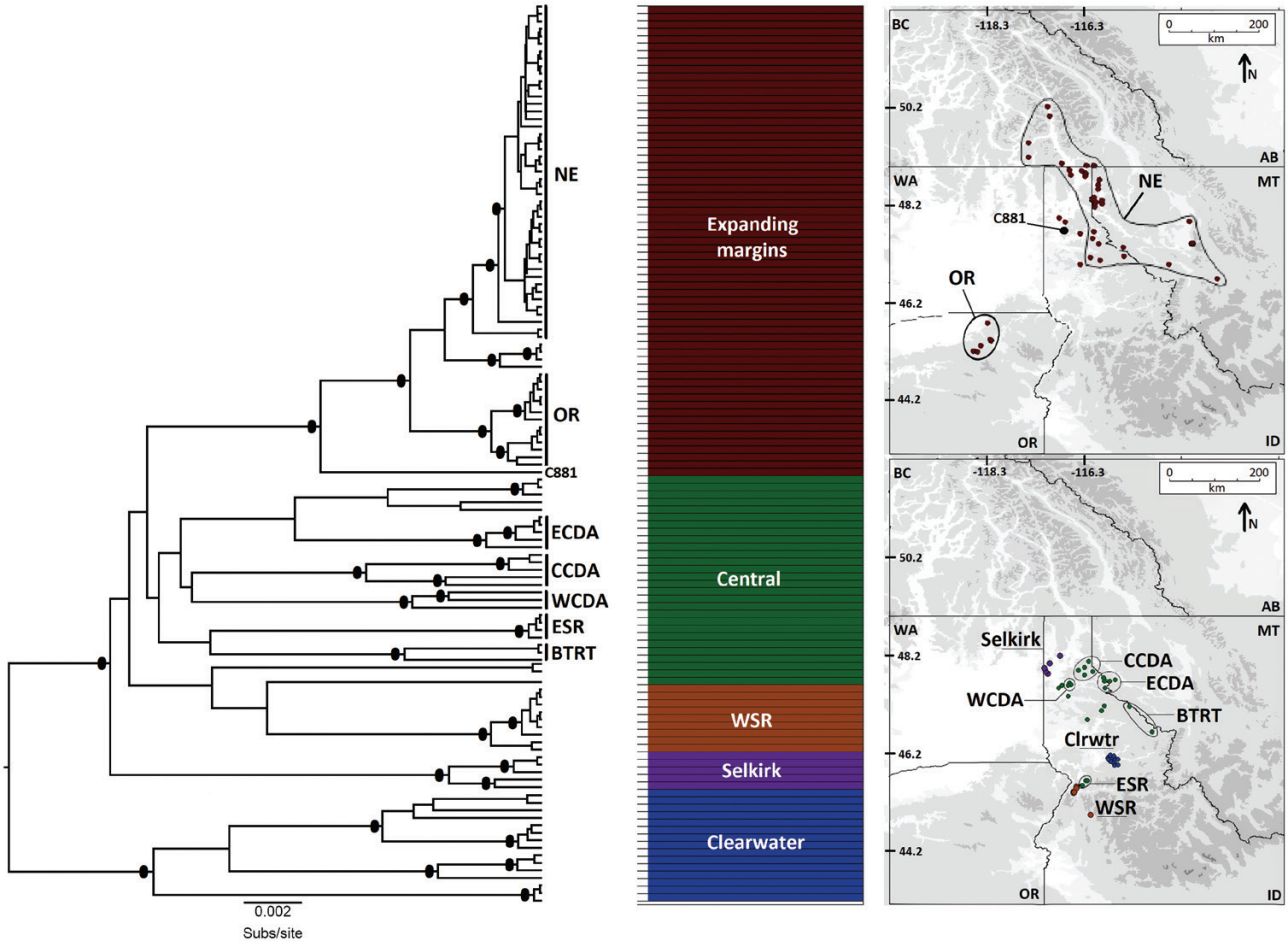

Figure 4. Bayesian phylogeny (constant population size tree prior) based on the mtDNA *A. k. occidentalis* all alleles data set. Circles on branches indicate Bayesian posterior probability ≥ 0.90. BAPS clusters that correspond to the best partition of the data are indicated by different coloured bars. Map showing the locations of sampled individuals. State and province names are abbreviated as follows: AB, Alberta; BC, British Columbia; ID, Idaho; MT, Montana; OR, Oregon; WA, Washington.

and Oregon (hereafter referred to as the expanding margins cluster).

The Bayesian skyline plot revealed a historical signal of stability, followed by a decline and then increase during the most recent interval (Supporting Information, Fig. S5). Due to the putative effects of lumping populations, we repeated the skyline analysis for three of the BAPS clusters (the Selkirk and WSR clusters were excluded due to low sample size). The expanding margins cluster shows a shallow history and a recent, precipitous increase. On the other hand, both the Central cluster and Clearwater cluster show a deeper history (relative to the expanding margins cluster) and a wavering but relatively stable trend. Lastly, the best supported scenario evaluated using DIYABC was scenario 2 (logistic approach PP = 0.39; Fig 1). In this scenario, the oldest event ($t3$) was a simultaneous divergence of the five BAPS clusters with an increase in N_e for the expanding margins cluster in the most recent ($t1 - 0$) interval. The posterior probabilities of remaining scenarios were as follows: scenario 1, PP = 0.13; scenario 3, PP = 0.19; scenario 4, PP = 0.24; scenario 5, PP = 0.01; and scenario 6, PP = 0.04.

A. NIMAPUNA PHYLOGENY AND POLYMORPHISM

The *A. nimapuna* data set consisted of 60 sequences of 1419 bp. There were 78 variable sites, 49 of which were parsimony informative and 29 singletons, and 40 unique haplotypes, and for the protein-coding sections (*COI* and *cytb*) substitutions occurred primarily at third positions and synonymous substitutions outnumbered non-synonymous. Results of phylogenetic analyses were nearly identical between the best maximum likelihood (Fig. 5) and Bayesian trees (Fig. 6). There are two well-supported clades that correspond to an eastern and western group (Fig. 6).

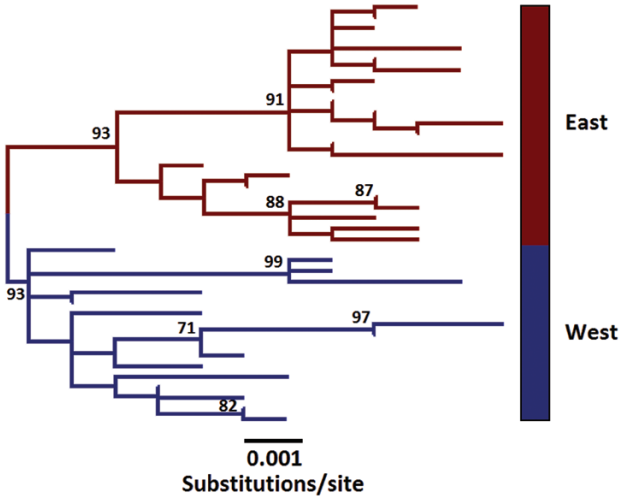

Figure 5. Best maximum likelihood phylogeny for *A. nimapuna* based on mtDNA haplotypes. Bootstrap values on nodes indicate relationships that are well-supported (≥ 70). BAPS clusters that correspond to the best partition of the data are indicated by different coloured bars.

Uncorrected nucleotide diversity (π) was 0.0076 while the Watterson estimator (θ) was higher at 0.012 (Supporting Information, Fig. S6), although this difference is not significant: Tajima's D = -1.22 ($P > 0.10$). We repeated Tajima's D for both western and eastern clades which showed D = -0.95 ($P > 0.10$) and D = -1.01 ($P > 0.10$), respectively. We compared the distribution curves of π (Supporting Information, Fig. S7) for the *A. nimapuna* data set to that of a simulated data set (of a single, neutrally evolving population) produced to have an identical mean. The two distribution curves displayed identical means and qualitatively similar variabilities, suggesting that *A. nimapuna* is composed of a single population. However, BAPS clustering analysis returned the best partition of the data as two clusters (K = 2) that correspond to the eastern and western phylogroups (Fig. 6).

The Bayesian skyline plot for all *A. nimapuna* individuals indicated a steadily increasing trend (Supporting Information, Fig. S8). When we repeated the skyline analysis for the two BAPS clusters, the eastern group showed a similarly increasing trend; however, the trend for the western clade is more or less stable without abrupt change. Lastly, the best supported scenario evaluated using DIYABC was scenario 2 (logistic approach PP = 0.80; Fig. 1A) in which there was a signal of population increase for both eastern and western clusters during the $t2 - 0$ interval. The posterior probabilities of remaining scenarios were as follows: scenario 1, PP = 0.04; scenario 3, PP = 0.01; scenario 4, PP = 0.03; and scenario 5, PP = 0.12.

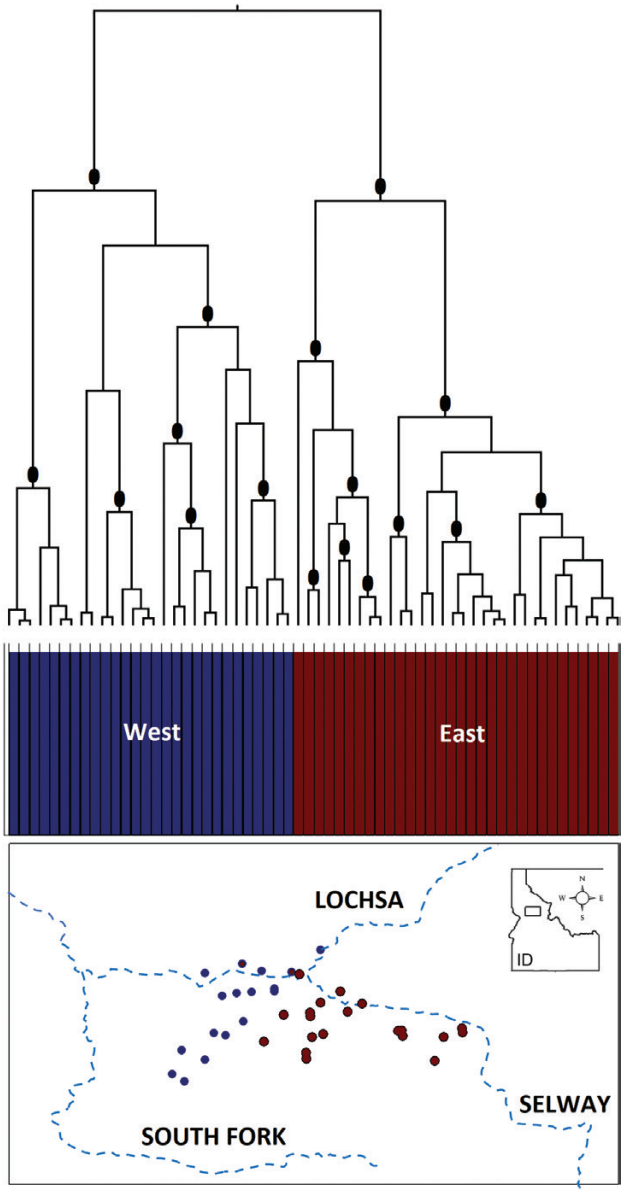

Figure 6. Bayesian phylogeny (constant population size tree prior) based on the mtDNA *A. nimapuna* all alleles data set. Circles on branches indicate Bayesian posterior probability ≥ 0.90. BAPS clusters that correspond to the best partition of the data are indicated by different coloured bars. Map showing the distribution of the *A. nimapuna* sampling localities and two mtDNA clades along the Lochsa, Selway and South Fork rivers in the Clearwater River Drainage of Idaho County, ID.

SPECIES DISTRIBUTION MODELS

For both current and LGM time periods, we report the projected (predicted) habitat suitability for *A. k. occidentalis* (Supporting Information, Fig. S9) and *A. nimapuna* (Supporting Information, Fig. S10) at two different thresholds of suitability

(0.5 and 0.75). The *A. k. occidentalis* prediction for current conditions is largely consistent with the known distribution, showing the species occupies the coniferous ecosystem of the NRMs. Results for the LGM suggests a palaeodistribution largely contained within the current range of the species, though areas of high suitability (> 0.75) are projected to be more fragmented. The *A. nimapuna* prediction for current conditions is largely restricted to the Clearwater drainage in Idaho County, ID, and the immediate surrounding areas. Results for the LGM suggests a palaeodistribution slightly more widespread and encompassing the current range of the species.

DISCUSSION

The inland temperate rainforests of Idaho, Montana and British Columbia are home to a large array of endemic taxa (Brunsfeld *et al.*, 2001; DellaSala, 2011). Some of these have phylogenetic connections to Pacific coastal species (e.g. tailed frogs, *Ascaphus montanus*), some are relictual species that only occur in the inland rainforest (e.g. pygmy slugs, *Kootenaia burkei*), yet others have more complex biogeographic histories. Here, we examined two taxa in the tiger snail genus *Anguispira*, one of which (*A. kochi occidentalis*) has biogeographic affinities to eastern congeners and the other (*A. nimapuna*) exhibits a classic narrow endemic pattern associated with the single drainage of the Clearwater River.

Interspecific phylogeny

Phylogenetic analysis of the mtDNA indicates that *Anguispira* may not be monophyletic, and that *A. nimapuna* and *A. k. kochi*+*A. k. occidentalis* are differentiated from the remaining *Anguispira* species, the latter of which form a statistically supported clade in all three phylogenetic reconstruction methodologies (Fig 2; Supporting Information, Figs S1, S2). Indeed, the most consistent result from these analyses is that eastern *Anguispira* species (apart from subspecific *A. k. kochi*) form a strongly supported clade. Only in the ML (Fig. 2) reconstruction do all *Anguispira* species form a monophyletic group, with *A. kochi* and then *A. nimapuna* splitting off deeply; however, monophyly is not statistically supported (< 70 bootstrap support). In the other two analyses, *A. nimapuna* and *A. kochi* cluster with *Discus* species (Supporting Information, Figs S1, S2). This low phylogenetic resolution is likely explained by a relatively short period of rapid evolution deep in the tree, resulting in short internal branches relative to the longer branches that lead to terminal taxa.

The distribution of *A. kochi* is disjunct, consisting of two components that extend from southern Ontario to Tennessee in the east and from southern British Columbia to Oregon in the west. The western form, *A. k. occidentalis* (Von Martens, 1882), was described subsequent to the eastern type based on shell morphology (coloration, thickness and size), and Pilsbry (1948) noted that while some western specimens are not discernible from eastern specimens, most could be easily separated. Although the western and eastern forms display superficially indistinguishable shell characteristics and morphology, Pilsbry (1948) maintained the subspecific status of *A. k. occidentalis* because of a separation of over 2000 km with no connections. These two subspecies are sisters in all three trees presented in this study, but with varied statistical support. That is, Bayesian posterior probability support is high (0.98; Supporting Information, Fig. S2) but bootstrap support is < 70 in both the MP (Supporting Information, Fig. S1) and ML (Fig. 2) reconstructions. This relationship between *A. k. kochi* and *A. k. occidentalis* is of particular interest given the large midcontinental gap between the two distributions. Several other taxonomic groups (with similarly low dispersal capacities) share this disjunct pattern of occurrence in both eastern and western North America, including salamanders (Vieites *et al.*, 2007) and spiders (Hendrixson & Bond, 2007), as well as other snails (Emberton, 1994). These distributions hint at an originally continuous, transcontinental range with subsequent vicariance due to increasing aridity in the interior (Emberton & Roth, 1994). Indeed, a major vicariant event associated with the Western Interior Seaway (Cretaceous period; 145–66 Mya) followed by the decline of forest biomes and spread of grasslands (Roberts & Kirschbaum, 1995) may be the most parsimonious scenario for diversification at higher ranks, although divergence at lower levels may be more likely explained by more recent transcontinental movements. However, the former explanation is plausible for the *A. k. kochi*/*A. k. occidentalis* split given it is about as deep as the basal divergence for the clade comprising all the other eastern species. Regardless, *A. k. kochi* and *A. k. occidentalis* have been in isolation long enough to have achieved reciprocal monophyly which takes, on average, xN_e generations [where x is the inheritance scaler (Moritz, 1994; Allendorf & Luikart, 2009)]. Given that these two groups are independently evolving lineages that are geographically isolated, we believe that elevating *A. k. occidentalis* to species status is warranted in order to facilitate more efficient conservation planning or management.

A. nimapuna is the only *Anguispira* species other than *A. k. occidentalis* that is uniquely distributed in western North America and was described in a paper

alongside *Discus marmorensis*, an extremely narrow endemic known only from the confluence of John Day Creek and the Salmon River in Idaho (Baker, 1932). Interestingly, Umiński (1963) suggested the possibility of a closer relationship between *A. nimapuna* and *D. marmorensis* (and proposed changing *D. marmorensis* to *A. marmorensis*) given similarities in the shape, aperture and sculpture of the shell, as well as geographical distribution; both species inhabit mountainous regions of north-central Idaho. Although *A. nimapuna* and *D. marmorensis* cluster together in our Bayesian analysis (Supporting Information, Fig. S2), this relationship has < 0.90 PP support. In the other analyses, *A. nimapuna* clusters with other *Discus* species in the MP reconstruction (Supporting Information, Fig. S1) and is sister to all other *Anguispira* species in the ML reconstruction (Fig. 2). That *A. nimapuna* is a member of the Discidae is supported; however, like *A. kochi*, the taxon appears to be only distantly related to other *Anguispira* species. That is, if it is not sister to the rest of *Anguispira*, it is a probable *Discus* representative. However, this is not comprehensive because there are several *Discus* species that are not represented in our data set, including two taxa native to western North America—*D. brunsoni* [lake disc (Berry, 1955)] and *Discus shimekii* [striate disc (Pilsbry, 1890)]—as well as up to three other named *Discus* species in North America and Eurasia.

Results regarding only eastern *Anguispira* (apart from *A. k. kochi*) generally agreed with that of Haskell & Pan (2013) in the grouping patterns and their levels of support. There is a well-supported clade of carinate limestone specialists as follows: {[(*Anguispira alabama*, *Anguispira picta*), Dry Creek], *A. cumberlandiana*}. Specimens of *Anguispira alabama* were collected from Alabama and Tennessee whereas specimens of the other three clades were all collected in Tennessee. The Dry Creek Population originates from a rare population of snails recently discovered (2008) in Tennessee in an area north-west of the *A. picta* range (see Haskell & Pan, 2013). *A. cumberlandiana* is recovered as two additional clades, one of which contains *A. cumberlandiana* individuals while the other contains the subspecific *A. c. columba*. All three *A. cumberlandiana* clades are from Tennessee but do not appear to be closely related. Specimens of *A. fergusoni*, an Atlantic coastal plain species, form a strongly supported clade. *Anguispira jessica* are recovered as three distantly related clades, one including only *A. jessica* individuals (from North Carolina) whereas the other two each contain *Anguispira jessica* individuals are recovered and *A. alternata* individuals, one from Virginia and the other from West Virginia. Additional specimens identified as *A. alternata* are recovered as members of several different clusters. The largest *A. alternata* clade contains representatives broadly distributed from the central and eastern United States and Ontario. A second *Anguispira alternata* clade is sister to *Anguispira columbiana columba*, and these two clades are sister to an *Anguispira strongylodes* clade, all three of which were collected from Tennessee and form a highly supported monophyletic group. A third well-supported cluster contains several *A. alternata* specimens as well as *A. macneilli*, *A. mordax* and additional *A. strongylodes*. Collection localities from this group are broadly distributed across the southern United States. Given identifications were based on shell morphology, it is likely that some distantly related clades have converged on a similar external shell morphology.

A. K. OCCIDENTALIS PHYLOGEOGRAPHY

The current distribution of *A. k. occidentalis* spans both unglaciated and glaciated regions of the NRM [the Last Glacial Period occurred from ~115 000–11 700 years ago (Pielou, 1991)], and our data suggest that both historical isolation (e.g. the deep split between the Clearwater clade and all remaining clades) and more recent expansions (e.g. the NE clade) have shaped the spatial distribution and genetic structure of this species. Genetic diversity is structured at both deep and shallow scales, with there being several deeply divergent mtDNA lineages that, in turn, exhibit shallower coalescent events among samples. The deepest split separates the Clearwater clade from all other clades. Despite being from a small geographic area (samples collected around the confluence of the Lochsa, Clearwater and Selway rivers), the Clearwater clade has a relatively high level of diversity (13 haplotypes identified from 15 samples). The depth of coalescence within the Clearwater clade is also comparable to the interallelic divergences between the remaining phylogroups (Figs 3, 4) which, coupled with the predicted demographic stability from the Skyline analysis (Supporting Information, Fig. 8), suggest long-term stability of both population and habitat. In addition to the Clearwater clade, there is a highly supported 'expanding margins' clade that includes a single haplotype from Idaho's panhandle (C881), a highly supported Oregon clade, and a multitude of individuals from northern Idaho, Montana and British Columbia, the lattermost of which exhibits a collapsed, comb-like structure indicative of a recent demographic expansion (Figs 3, 4). BAPS inferred an additional three clusters: Selkirk, WSR and Central (Fig. 4). The Selkirk clade is geographically separated from neighbouring lineages by the Priest River basin whereas the WSR group is separated from ESR individuals by the Salmon River. Finally, within the

Central BAPS cluster, three clades (WCDA, CCDA and ECDA) are distributed allopatrically across the Coeur d'Alene Mountains of Idaho and Montana but any putative physiographic barriers effecting them is not immediately apparent (Figs. 4, 6).

The overall phylogeographic pattern in *A. k. occidentalis*—considerable differentiation across a small geographic range, as well as mixing of phylogroup lineages—is often reported for other terrestrial snails (Thomaz *et al.*, 1996), including another western United States snail group, *Oreohelix* (Dempsey *et al.*, 2019, 2020; Linscott *et al.*, 2020). Compared to other taxa, land snails exhibit high mtDNA substitution rates (Thomaz *et al.*, 1996), and highly divergent lineages separated by short spatiotemporal events are often found (Chiba, 1999; Pinceel *et al.*, 2005). We suggest that the mtDNA patterns observed in *A. k. occidentalis* are likely the result of a combination of high mtDNA substitution rates and prolonged isolation and *in situ* diversification of refugial populations and subsequent expansion from multiple refugia. Indeed, the lack of resolution of relationships among deeper branches of the *A. k. occidentalis* phylogeny, and the short internodes separating those splits, is consistent with a near-simultaneous fragmentation of a widespread ancestor. Our ABC analyses provide additional support for a near-simultaneous origin for the BAPS clusters, as well as the subsequent increase of N_e of the 'expanding margin' clade.

Taken together, these results strongly support, at a minimum, a NRM dual refugia hypothesis (Brunsfeld *et al.*, 2001). That is, advancing glaciers likely forced the mesic ecosystem into refugia located to the south or in canyons deep enough to offer climatic insulation (Daubenmire, 1975), thereby fragmenting a previously contiguous distribution of populations into multiple compartmentalized refugia and fostering divergence among groups. Indeed, our SDM analysis predicted areas of high habitat suitability were more fragmented during the LGM when compared to current conditions (Supporting Information, Fig. S9), and several studies have shown that during past glaciations species' distributions were not uniformly shifted to the south, but that multiple refugia occurred proximal to the ice sheets (Stewart & Lister, 2001). Genetic data from *Hemphillia* slugs (Rankin *et al.*, 2019), Rocky Mountain tailed frog [*Ascaphus montanus* (Nielson *et al.*, 2001; Metzger *et al.*, 2015)] and Constance's bittercress [*Cardamine constancei* (Brunsfeld & Sullivan, 2005)] are all consistent with expansion from multiple refugia; however, there are other species that exhibit evidence of expansion from a single refugium [e.g. Idaho giant salamanders, *Dicamptodon aterrimus* (Carstens *et al.*, 2005)].

A. NIMAPUNA PHYLOGEOGRAPHY

A. nimapuna is a narrow-range endemic species that occurs only around the watershed of the South Fork of the Clearwater River and upstream into the Lochsa and Selway drainages, Idaho County, Idaho (Fig. 6). Despite this relatively small area (our sampling spans ~60 km east to west and ~30 km north to south) phylogenetic analyses revealed two geographically structured and highly supported mtDNA lineages. However, the net nucleotide diversity between these two phylogroups (0.00432 subs/site) is on par with intralineage nucleotide diversity (0.00495^{West} subs/site and 0.00594^{East} subs/site), in which case we assume the two groups have been diverging for only ~N_e generations (Moritz, 1994; Allendorf & Luikart, 2009). Indeed, the variance of π for all samples is qualitatively similar to that expected for a single, neutrally evolving population (Supporting Information, Fig. S7). If the western and eastern phylogroups are indeed real, then it is likely they have only very recently diverged.

BAPS also recovered the western and eastern clusters as the best partition of the data, and subsequent DIYABC analysis of those clusters indicated this species has experienced an increase in population size, which is perplexing given its narrowly restricted range. Although this result was only partially supported by the skyline analyses in that the western group did not exhibit a strong signal of increasing N_e (Supporting Information, Fig. S8). A stable and perhaps increasing demographic history would be predicted for a species inhabiting an area of habitat stability through fluctuating climates, and it is not unreasonable to assume *A. nimapuna* has occupied its current habitat throughout the Quaternary climate fluctuations given our SDM analysis predicted the Clearwater drainage to have suitable habitat for both LGM and current conditions (Supporting Information, Fig. S10). Several additional lines of evidence suggest the Clearwater served as a refugium during past glacials (Daubenmire, 1952, 1978; Detling, 1968; Shafer *et al.*, 2010). That is, the river canyons of the Clearwater basin are some of the northernmost canyons that were free of glacial ice and probably served as a refugium for many extant species with high moisture-heat requirements. This moist-canyon ecosystem contains many plants and animals endemic to the NRMs, as well as regional endemics whose distributions are closely tied to the Clearwater drainage such as the bank monkeyflower [*Mimulus clivicola* (Lorain, 1992)], several species of byrrhid beetles (Johnson, 1987) and the snail *Allogona lombardi* (Burke, 2013). It is also interesting to note that a widespread species like *A. k. occidentalis* contains a distinct (all private haplotypes) Clearwater clade with relatively high diversity, suggesting this is a long-established population as well.

High concentrations of narrow endemics can be linked to the accumulation of distinct evolutionary units within long-term stable environments (Molina-Venegas *et al.*, 2017) and are sometimes indicative of refugia that have experienced long-term stability of climate and habitats (e.g. Jetz *et al.*, 2004; Ohlemüller *et al.*, 2008; Harrison & Noss, 2017). However, although this may explain why the Clearwater refugium is an endemic hotspot, it does not explain why some species with highly restricted ranges (*A. nimapuna*) have close relatives with widespread distributions (*A. k. occidentalis*). Given the distinctive conditions of the Clearwater drainage (a unique climate resulting from a combination of a low-elevation canyon, mountainous terrain, and high precipitation and moderate temperatures similar to Pacific coastal climates), it may be that *A. nimapuna* is adapted to the particular conditions of the Clearwater drainage and has not colonized new adjacent sites with more stressful conditions. However, the observation that *A. k. occidentalis* contains a unique and diverse Clearwater clade may also suggest colonization to and from this immediate area is difficult. Indeed, the refugium ecosystem occupies a V-shaped valley and slopes rise abruptly from the valley to form steep breaklands and upstream and downstream migration pathways may prove to be too challenging for organisms with low dispersal ability (Lichthardt & Moseley, 1994). In any case, *A. nimapuna* is an exceptionally rare species that occupies habitat that is itself rare and sensitive. Species with these attributes should be allotted conservation priority, which is amongst the reasons this species is classified as a Species of Greatest Conservation Need by the U.S. state of Idaho (IDFG, 2017).

CONCLUSION

Both *A. k. occidentalis* and *A. nimapuna* represent unique taxa that are genetically and/or geographically distinct from their congeners and are part of a larger group of endemic molluscs that display restricted geographic distributions in the NRM. Several of these are highly localized in small areas such as the Clearwater drainage of north-central Idaho. A combination of intrinsic and extrinsic factors, such as molluscs having poor dispersal ability or limited accidental dispersal opportunities, geologic and historic factors, and long-term permanency of some NRM habitats likely have contributed to the high diversity and regional endemism in the area. There is an increasing need to obtain data on the genetic and spatial structure of endemic taxa in areas such as the NRM due to the threat of habitat fragmentation (both natural and induced by human activity), which may be exponentially greater to minute organisms that maintain small ranges and have little public appeal.

ACKNOWLEDGEMENTS

This research was supported by the National Science Foundation grant DEB-1457519 (to J.S.), the Institute for Bioinformatics and Evolutionary Studies (IBEST) at the University of Idaho, which is supported by National Institute of Health (NIH) NCRR 1P20RR016454-01, NIH NCRR 1P20RR016448-01 and NSF EPS-809935, and a University of Idaho graduate fellowship in Bioinformatics and Computational Biology to A.M.R. Collection of field samples was supported by U.S. Fish and Wildlife Service Wildlife and Sport Fish and Restoration Program Competitive State Wildlife Grant (F12AP01101). We thank the many employees and volunteers of the Multi-species Baseline Initiative project (https://idfg.idaho.gov/baseline) for providing specimens. Additional tissue samples were graciously provided by the California Academy of Sciences, Carnegie Museum of Natural History, Delaware Museum of Natural History, Field Museum of Natural History, Montana Natural Heritage Program, Royal British Columbia Museum and University of Florida Museum. We thank the following individuals for donating specimens: Amy Van Devender, Bill Leonard, Casey Richart, Dan Dourson, Joel Sauder, John Slapcinsky, Mason Linscott, Ron Caldwell and Stephanie Clark. Special thanks go to Lindsay McCoy for preparing the figures. Lastly, we thank three reviewers for the helpful comments that improved the manuscript. The authors declare no conflicts of interest.

REFERENCES

Allendorf FW, **Luikart G. 2009.** *Conservation and the genetics of populations.* Hoboken, New Jersey: John Wiley & Sons.

Baker HB. 1932. New land snails from Idaho and eastern Oregon. *The Nautilus* **45:** 82–87.

Berry SS. 1955. An important new land-snail from the Mission Range, Montana. *Bulletin of the Southern California Academy of Sciences* **54:** 17–19.

Bouckaert R, **Heled J**, **Kühnert D**, **Vaughan T**, **Wu CH**, **Xie D**, **Suchard MA**, **Rambaut A**, **Drummond AJ. 2014.** BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Computational Biology* **10:** e1003537.

Brunsfeld SJ, **Sullivan J. 2005.** A multi-compartmented glacial refugium in the northern Rocky Mountains: evidence from the phylogeography of *Cardamine constancei* (Brassicaceae). *Conservation Genetics* **6:** 895–904.

Brunsfeld SJ, **Sullivan J**, **Soltis DE**, **Soltis PS. 2001.** Comparative phylogeography of northwestern North

America: a synthesis. *Special Publication-British Ecological Society* **14:** 319–340.

Burke TE. 2013. *Land snails and slugs of the Pacific Northwest.* Corvalis, Oregon: Oregon State University Press.

Carstens BC, **Degenhardt JD**, **Stevenson AL**, **Sullivan J. 2005.** Accounting for coalescent stochasticity in testing phylogeographical hypotheses: modelling Pleistocene population structure in the Idaho giant salamander *Dicamptodon aterrimus*. *Molecular Ecology* **14:** 255–265.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* **17:** 540–552.

Chiba S. 1999. Accelerated evolution of land snails *Mandarina* in the oceanic Bonin Islands: evidence from mitochondrial DNA sequences. *Evolution* **53:** 460–471.

Clutts SA. 2008. *Systematics of* Anguispira *(Pulmonata: Discidae) based on molecular and morphometric data.* MS Dissertation, Southern Illinois University.

Collins WD. 2006. The community climate system model version 3 (CCSM3). *Journal of Climate* **19:** 2122–2143.

Corander J, **Marttinen P**, **Sirén J**, **Tang J. 2008.** Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* **9:** 1–14.

Cornuet JM, **Ravigné V**, **Estoup A. 2010.** Inference on population history and model checking using DNA sequence and microsatellite data with the software DIYABC (v1. 0). *BMC Bioinformatics* **11:** 401.

Cornuet JM, **Santos F**, **Beaumont MA**, **Robert CP**, **Marin JM**, **Balding DJ**, **Guillemaud T**, **Estoup A. 2008.** Inferring population history with DIY ABC: a user-friendly approach to approximate Bayesian computation. *Bioinformatics* **24:** 2713–2719.

Daubenmire R. 1952. Plant geography of Idaho. In: Davis RJ, ed. *Flora of Idaho.* Provo: Brigham Young University Press, 1–17.

Daubenmire R. 1975. Floristic plant geography of eastern Washington and northern Idaho. *Journal of Biogeography* **2:** 1–18.

Daubenmire R. 1978. *Plant geography with special reference to North America.* New York: Academic Press.

DellaSala DA, **ed. 2011.** *Temperate and boreal rainforests of the world: ecology and conservation.* Washington, D.C.: Island Press.

Dempsey ZW, **Burg TM**, **Goater CP. 2019.** Found, forgotten, and found again: systematics and distribution of Cooper's Rocky Mountain snail (*Oreohelix cooperi*) on a sky island in the Canadian Prairies. *Canadian Journal of Zoology* **97:** 833–840.

Dempsey ZW, **Goater CP**, **Burg TM. 2020.** Living on the edge: comparative phylogeography and phylogenetics of *Oreohelix* land snails at their range edge in Western Canada. *BMC Evolutionary Biology* **20:** 1–13.

Detling LE. 1968. *Historical background of the flora of the Pacific Northwest. Bulletin No. 13*. Eugene: Museum of Natural History, University of Oregon.

Emberton KC. 1991a. Ecology of a shell convergence between subfamilies of polygyrid land snails. *Biological Journal of the Linnean Society* **44:** 105–120.

Emberton KC. 1991b. Polygyrid relations: a phylogenetic analysis of 17 subfamilies of land snails (Mollusca: Gastropoda: Stylommatophora). *Zoological Journal of the Linnean Society* **103:** 207–224.

Emberton KC. 1994. Polygyrid land snail phylogeny: external sperm exchange, early North American biogeography, iterative shell evolution. *Biological Journal of the Linnean Society* **52:** 241–271.

Emberton KC. 1995. Sympatric convergence and environmental correlation between two land-snail species. *Evolution* **49:** 469–475.

Emberton KC, **Roth B. 1994.** "Extralimital" Land Mollusks (Gastropoda) from the Deep River Formation, Montana: Evidence for Mesic Medial Tertiary Climate. *Proceedings of the Academy of Natural Sciences of Philadelphia* **145:** 93–106.

Excoffier L, **Laval G**, **Schneider S. 2007.** Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1:** 47–50.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39:** 783–791.

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

Fend SV, **Gustafson DL. 2001.** A new lumbriculid genus and species from North America (Clitellata, Lumbriculidae). In: Pilar Rodriguez and Piet F. M. Verdonschot, eds. *Aquatic Oligochaete Biology. Vol. 8*. Dordrecht: Springer, 13–22.

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

Geller JB, **Walton ED**, **Grosholz ED**, **Ruiz GM. 1997.** Cryptic invasions of the crab *Carcinus* detected by molecular phylogeography. *Molecular Ecology* **6:** 901–906.

Harrison S, **Noss R. 2017.** Endemism hotspots are linked to stable climatic refugia. *Annals of Botany* **119:** 207–214.

Haskell DG, **Pan JW. 2013.** Phylogenetic analysis of threatened and range-restricted limestone specialists in the land snail genus *Anguispira*. *Conservation Genetics* **14:** 671–682.

Hendrixson BE, **Bond JE. 2007.** Molecular phylogeny and biogeography of an ancient Holarctic lineage of mygalomorph spiders (Araneae: Antrodiaetidae: *Antrodiaetus*). *Molecular Phylogenetics and Evolution* **42:** 738–755.

Hijmans RJ, **Cameron SE**, **Parra JL**, **Jones PG**, **Jarvis A. 2005.** Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology: A Journal of the Royal Meteorological Society* **25:** 1965–1978.

Hillis DM, **Mable BK**, **Moritz C. 1996.** Applications of molecular systematics: the state of the field and a look to the future. In: *Molecular systematics.* Sunderland: Sinauer Associates, Inc., 515–543.

Ho SY, **Shapiro B. 2011.** Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources* **11:** 423–434.

Idaho Department of Fish and Game. **2017**. *Idaho state wildlife action plan.* In: Dixon R, ed. Boise: Idaho Department of Fish and Game.

Jetz W, **Rahbek C**, **Colwell RK. 2004.** The coincidence of rarity and richness and the potential signature of history in centres of endemism. *Ecology Letters* **7:** 1180–1191.

Johnson PJ. 1987. Larval taxonomy, biology, and biogeographyof the genera of North American Byrrhidae. M.S. thesis, University of Idaho, Moscow, ID.

Lichthardt J, **Moseley RK. 1994.** Ecosystem analysis and conservation planning for the Clearwater refugium, Clearwater and Nez Perce National Forests. *Unpublished report for the Nez Perce and Clearwater National Forests*. Boise: Idaho Department of Fish and Game, Conservation Data Center.

Linscott TM, **Weaver K**, **Morales V**, **Parent CE. 2020.** Assessing species number and genetic diversity of the Mountainsnails (Oreohelicidae). *Conservation Genetics* **21:** 971–985.

Lorain C. 1992. *Conservation strategy for* Mimulus clivicola *(bank monkeyflower)*. Unpublished report. Boise: Idaho Department of Fish and Game, Conservation Data Center.

Lucid MK, **Ehlers S**, **Robinson L**, **Cushman SA. 2018.** Beer, brains, and brawn as tools to describe terrestrial gastropod species richness on a montane landscape. *Ecosphere* **9:** e02535.

Lucid MK, **Rankin A**, **Espindola A**, **Chichester L**, **Ehlers S**, **Robinson L**, **Sullivan J. 2018.** Taxonomy and biogeography of *Hemphillia* (Gastropoda: Pulmonata: Arionidae) in North American rainforests, with description of a new species (Skade's jumping-slug, *Hemphillia skadei* sp. nov.). *Canadian Journal of Zoology* **96:** 305–316.

Martens E von. 1882. Ueber von den Gebrüdern Krause in Amerika gesammelte Conchylien. *Sitzungsberichte der Gesellschaft Naturforschender Freunde zu Berlin* (November): 138–143.

Merritt TJS, **Shi L**, **Chase MC**, **Rex MA**, **Etter RJ**, **Quattro JM. 1998.** Universal cytochrome b primers facilitate intraspecific studies in molluscan taxa. *Molecular Marine Biology and Biotechnology* **7:** 7–11.

Metzger G, **Espindola A**, **Waits LP**, **Sullivan J. 2015.** Genetic structure across broad spatial and temporal scales: Rocky Mountain tailed frogs (*Ascaphus montanus*; Anura: Ascaphidae) in the inland temperate rainforest. *Journal of Heredity* **106:** 700–710.

Minin V, **Abdo Z**, **Joyce P**, **Sullivan J. 2003.** Performance-based selection of likelihood models for phylogeny estimation. *Systematic Biology* **52:** 674–683.

Molina-Venegas R, **Aparicio A**, **Lavergne S**, **Arroyo J. 2017.** Climatic and topographical correlates of plant palaeo- and neoendemism in a Mediterranean biodiversity hotspot. *Annals of Botany* **119:** 229–238.

Moritz C. 1994. Applications of mitochondrial DNA analysis in conservation: a critical review. *Molecular Ecology* **3:** 401–411.

NatureServe. **2017.** NatureServe Explorer: An online encyclopedia of life [web application]. Version 7.1. NatureServe, Arlington, Virginia. Available at: http://explorer.natureserve.org.

Nei M, **Jin L. 1989.** Variances of the average numbers of nucleotide substitutions within and between populations. *Molecular Biology and Evolution* **6:** 290–300.

Nei M, **Li WH. 1979.** Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences* **76:** 5269–5273.

Newell RL, **Baumann RW**, **Stanford JA. 2008.** Stoneflies of Glacier National Park and Flathead River basin, Montana. In: Hauer FR, Stanford JA, Newell RL, eds. *International advances in the ecology, zoogeography and systematics of mayflies and stoneflies. Entomology Vol. 128*, 173–186. Berkeley, California: University of California Press.

Nielson M, **Lohman K**, **Sullivan J. 2001.** Phylogeography of the tailed frog (*Ascaphus truei*): implications for the biogeography of the Pacific Northwest. *Evolution* **55:** 147–160.

Nordsieck H. 1986. The system of the Stylommatophora (Gastropoda), with special regard to the systematic position of the Clausiliidae, II: Importance of the shell and distribution. *Archiv für Molluskenkunde der Senckenbergischen Naturforschenden Gesellschaft* **117:** 93–116.

Ohlemüller R, **Anderson BJ**, **Araújo MB**, **Butchart SH**, **Kudrna O**, **Ridgely RS**, **Thomas CD. 2008.** The coincidence of climatic and species rarity: high risk to small-range species from climate change. *Biology Letters* **4:** 568–572.

Peterson AT. 2011. *Ecological niches and geographic distributions (MPB-49).* New Jersey: Princeton University Press.

Pfeiffer L. 1845 (1846). Description of thirty-six new species of *Helix*, belonging to the collection of H. Cuming, Esq. *Proceedings of the Zoological Society of London* **1845:** 126–133.

Pielou EC. 1991. *After the ice age. The return of life to deglaciated North America.* Chicago: University of Chicago Press.

Pilsbry HA. 1890. Nautilus, 4:3. *Proceedings of the Academy of Natural Sciences of Philadelphia*. **42:** 297, pl. 5, figs. 9–11; reprinted in Nautilus, 5:39.

Pilsbry HA. 1948. Land Mollusca of North America. *Academy of Natural Sciences of Philadelphia, Philadelphia Monograph* **2:** 521–1113.

Pinceel J, **Jordaens K**, **Backeljau T. 2005.** Extreme mtDNA divergences in a terrestrial slug (Gastropoda, Pulmonata, Arionidae): accelerated evolution, allopatric divergence and secondary contact. *Journal of Evolutionary Biology* **18:** 1264–1280.

Rankin AM, **Wilke T**, **Lucid M**, **Leonard W**, **Espíndola A**, **Smith ML**, **... & Sullivan J. 2019.** Complex interplay of ancient vicariance and recent patterns of geographical speciation in north-western North American temperate rainforests explains the phylogeny of jumping slugs (Hemphillia spp.). *Biological Journal of the Linnean Society* **127:** 876–889.

R Core Team. **2014.** *R: a language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing. Available at: http://www.R-project.org/

Rambaut A, **Drummond AJ. 2007.** *Tracer v1.4.* Available at: http://beast.bio.ed.ac.uk/Tracer

Ritchie AM, **Lo N**, **Ho SY. 2017.** The impact of the tree prior on molecular dating of data sets containing a mixture of inter-and intraspecies sampling. *Systematic Biology* **66:** 413–425.

Roberts LNR, **Kirschbaum MA. 1995.** Paleogeography and the Late Cretaceous of the Western Interior of middle North America; coal distribution and sediment accumulation. *U.S. Geological Survey Professional Paper* **1561:** 115.

Shafer AB, **Cullingham CI**, **Cote SD**, **Coltman DW. 2010.** Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology* **19:** 4589–4621.

Stagliano DM. 2016. Mayflies (Insecta: Ephemeroptera) of conservation concern in Montana: status updates and management needs. *Western North American Naturalist* **76:** 441–451.

Stagliano DM, **Stephens GM**, **Bosworth WR. 2007.** *Aquatic invertebrate species of concern on USFS Northern Region lands. Report prepared for USDA Forest Service, Northern Region, Missoula, MT.* Boise and Helena: Montana Natural Heritage Program and Idaho Conservation Data Center. Agreement number 05-CS-11015600-036. 95 pp. + appendixes.

Stark BP, **Gustafson DL. 2004.** New species and records of *Soliperla* Ricker, 1952 from western North America (Insecta, Plecoptera, Peltoperlidae). *SPIXIANA* **27:** 97–105.

Stewart JR, **Lister AM. 2001.** Cryptic northern refugia and the origins of the modern biota. *Trends in Ecology & Evolution* **16:** 608–613.

Swofford DL. 2003. *PAUP: phylogenetic analysis using parsimony. Mac Version 3. 1. 1*. Available at: https://paup.phylosolutions.com/.

Tajima F. 1989a. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123:** 585–595.

Tajima F. 1989b. The effect of change in population size on DNA polymorphism. *Genetics* **123:** 597–601.

Tamura K, **Stecher G**, **Peterson D**, **Filipski A**, **Kumar S. 2013.** MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30:** 2725–2729.

Thomaz D, **Guiller A**, **Clarke B. 1996.** Extreme divergence of mitochondrial DNA within species of pulmonate land snails. *Proceedings of the Royal Society of London. Series B* **263:** 363–368.

Thuiller W, **Lafourcade B**, **Engler R**, **Araújo MB. 2009.** BIOMOD–a platform for ensemble forecasting of species distributions. *Ecography* **32:** 369–373.

Uminski T. 1963. Taxonomy of *Anguispira marmorensis* (H.B. Baker, 1932) and notes on the taxonomy of the genera *Anguispira* (Morse) and *Discus* (Fitzinger) (Gastropoda, Endodontidae). *Annales Zoologici Musei Polonici Historiae Naturalis* **21:** 81–91.

Vieites DR, **Min MS**, **Wake DB. 2007.** Rapid diversification and dispersal during periods of global warming by plethodontid salamanders. *Proceedings of the National Academy of Sciences* **104:** 19903–19907.

Wahlund S. 1928. Zusammensetzung von Populationen und Korrelationserscheinungen vom Standpunkt der Vererbungslehre aus betrachtet. *Hereditas* **11:** 65–106.

Wakeley J. 1996a. The variance of pairwise nucleotide differences in two populations with migration. *Theoretical Population Biology* **49:** 39–57.

Wakeley J. 1996b. Distinguishing migration from isolation using the variance of pairwise differences. *Theoretical Population Biology* **49:** 369–386.

Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology* **7:** 256–276.

Zwickl DJ. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph. D. Thesis, The University of Texas, Austin.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Data S1. Details of specimen IDs, localities, museum catalogue numbers and GenBank accession numbers.
Figure S1. Maximum parsimony phylogeny of the interspecific *COI-cytb*-16S haplotypes data set. Numbers above the branches represent bootstrap values (500 replicates).
Figure S2. Bayesian phylogeny of the interspecific *COI-cytb*-16S haplotypes data set. Numbers above the branches represent posterior probabilities.
Figure S3. Uncorrected nucleotide diversity (π) and Watterson estimator (θ) for the *COI-cytb*-16S *A. k. occidentalis* data set plotted against sample size.
Figure S4. The distribution curve of π for the *COI-cytb*-16S *A. k. occidentalis* data set compared to that of a simulated data set (of a single, neutrally evolving population) produced to have an identical mean.
Figure S5. Bayesian skyline plots for all *A. k. occidentalis* alleles (N = 120) and BAPS clusters (Expanding margins, N = 63; Central, N = 28; and Clearwater, N = 15) showing effective population size (scaled by mutation rate) through time. Note the time scale is given in substitution per site, which can be converted to units of time using a molecular clock calibration. Black lines indicate the median value while grey dashed lines denote the 0.95 highest posterior probability intervals.
Figure S6. Uncorrected nucleotide diversity (π) and Watterson estimator (θ) for the *COI-cytb*-16S *A. nimapuna* data set plotted against sample size.
Figure S7. The distribution curves of π for the *COI-cytb*-16S *A. nimapuna* data set compared to that of a simulated data set (of a single, neutrally evolving population) produced to have an identical mean.

Figure S8. Bayesian skyline plots for all *A. nimapuna* alleles ($N = 60$) and BAPS clusters (East, $N = 32$; and West, $N = 28$) showing effective population size (scaled by mutation rate) through time. Note the time scale is given in substitution per site, which can be converted to units of time using a molecular clock calibration. Black lines indicate the median value while grey dashed lines denote the 0.95 highest posterior probability interval.
Figure S9. Species distribution models (SDM) showing the projected distribution of *A. k. occidentalis* under (A) the last glacial maximum (LGM) and (B) current climatic conditions. Distributions are shown at 0.5 (orange) and 0.75 (red) probability thresholds.
Figure S10. Species distribution models (SDM) showing the projected distribution of *A. nimapuna* under (A) the last glacial maximum (LGM) and (B) current climatic conditions. Distributions are shown at 0.5 (orange) and 0.75 (red) probability thresholds.