Comparative Phylogeography of Microsnails from the Pacific Northwest

Authors: Smith, Megan L., Lang, Connor, Sneddon, David, Wallace, Jessica, Espíndola, Anahí, et al.

Source: Northwest Science, 96(1-2) : 117-132

Published By: Northwest Scientific Association

URL: https://doi.org/10.3955/046.096.0108

Megan L. Smith[1, 2], **Connor Lang**, **David Sneddon**[3], **Jessica Wallace**, Department of Evolution, Ecology, and Organismal Biology, Museum of Biological Diversity, The Ohio State University, 1315 Kinnear Road, Columbus, Ohio 43212

Anahí Espíndola, Department of Entomology, 4291 Fieldhouse Drive, University of Maryland, College Park, Maryland 20742-4454

Jack Sullivan, Department of Biological Sciences, and Institute for Bioinformatics and Evolutionary Studies (IBEST), Biological Sciences, University of Idaho, 875 Perimeter Drive, MS 3051, Moscow, ID 83844-3051.

Bryan C. Carstens, Department of Evolution, Ecology, and Organismal Biology, Museum of Biological Diversity, The Ohio State University, 1315 Kinnear Road, Columbus, Ohio 43212

Comparative Phylogeography of Microsnails from the Pacific Northwest

Abstract

Leaf-litter-dwelling invertebrates serve an important role in ecosystem function by breaking down nutrients and potentially acting as indicators of habitat quality. However, this community is understudied due to difficulties related to sampling and taxonomic identification. To explore this community, we sampled leaf litter from the coastal and Cascade ranges of the Pacific Northwest of North America and searched > 200 samples for micro-invertebrates. We removed and photographed more than 400 invertebrate specimens, sequenced a portion of the mitochondrial gene cytochrome oxidase I (COI) for 60 samples, and used COI and the BLASTn database to identify invertebrates. Using these sequences and environmental data from the collection localities, we investigated the phylogeographic history of the two best-sampled species of microsnails, the toothless column snail (*Columella edentula*) and the conical spot snail (*Punctum randolphii*). Results suggest that populations of these species from the coastal and Cascade ranges may have survived in a single refugium during the Pleistocene glacial cycles and recolonized the coastal and Cascade ranges during the Holocene. Our results add to the knowledge of species responses to the Pleistocene glacial cycles in the Pacific Northwest and suggest that future studies should aim to increase representation of micro-invertebrates, perhaps using metabarcoding techniques.

Keywords: micro-invertebrates, Pacific Northwest temperate rainforest, microsnails, phylogeography, barcoding

Introduction

The Pacific Northwest of North America (PNW) encompasses the largest extent of mesic, temperate forest in the world and is home to more than 150 endemic species (Nielson et al. 2001). These forests are dominated by western redcedar (*Thuja plicata*) and western hemlock (*Tsuga heterophylla*) (Daubenmire and Daubenmire 1968) and occur both inland, in the Northern Rocky Mountains of Montana, Idaho, and British Columbia, as well as in the Cascades and coastal ranges along the western coast from northern California to Alaska (Figure 1). The coastal and inland forests are largely separated by the arid Columbia Basin, although patches of forest are present to the south in the central Oregon highlands and to the north in the Okanogan highlands.

The history of the region is geologically and ecologically complex, with multiple forces shaping the distributions of mesic forest endemic taxa (reviewed in Brunsfeld et al. 2001). The earliest records of mesic forests in the region date to the middle Eocene in the northern Rocky Mountains of northeastern Washington (Graham 1999). With the orogeny of the Cascades (2 to 5 mya), the arid Columbia Basin formed separating the inland and coastal forests and leading to the

[1] Author to whom correspondence should be addressed. Email: Megan Smith, megansmth67@gmail.com
[2] Current address: Department of Biology, Indiana University, Jordan Hall, 1001 East 3rd Street, Bloomington, Indiana 47405
[3] Current address: Department of Biological Sciences, University of Idaho, 875 Perimeter Drive MS 3051, Moscow, Idaho 83844-3051

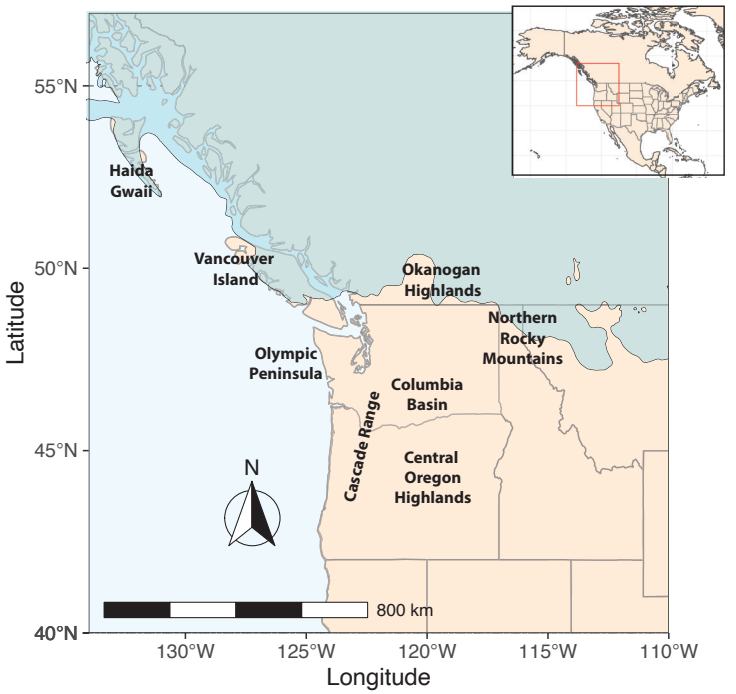

Figure 1. Map of the Pacific Northwest of North America, showing important locations, including potential refugia on Haida Gwaii Island, Vancouver Island, and the Olympic Peninsula. The light blue outline indicates the hypothesized extent of glaciers during the Last Glacial Maximum from Dyke et al. (2003).

current disjunct distribution. Later, during the Pleistocene, the region was heavily affected by 100,000 year glacial cycles during which large portions of the region were covered in ice (Delcourt and Delcourt 1993). During glacial cycles, many endemic species may have contracted their ranges to isolated refugia, expanding their ranges after glaciers retreated (Brunsfeld et al. 2001). Refugia have been proposed along the coast south of the glaciated regions, as well as farther north along the coast (e.g., on the Olympic Peninsula) and on islands (e.g., Vancouver Island, Haida Gwaii) (Brunsfeld et al. 2001; Figure 1). Inland refugia in the Northern Rocky Mountains have also been proposed, for example in the Clearwater River drainages (Brunsfeld et al. 2001).

A great deal of phylogeographic research has focused on this region, with contrasting patterns identified for rainforest endemics (e.g., Carstens et al. 2005). Several species of amphibians, including tailed frogs (*Ascaphus* spp.; Nielson et al. 2001, 2006) and salamanders (*Plethodon* and *Dicamptodon*; Carstens et al. 2004, 2005; Steele et al. 2005; Pelletier and Carstens 2016) appear to have survived in multiple, isolated refugia, and to have expanded to fill their current range following the Last Glacial Maximum (LGM). In contrast, some plants, including willow (*Salix*) and pine (*Pinus*), and mammals such as the water vole (*Microtus richardsoni*) show evidence of expansion from a single refugium after the LGM (Carstens et al. 2005). More recent work has focused on some of the region's invertebrates. Unlike amphibians, for which cryptic diversity associated with refugial structure has been found both within and between the Northern Rocky Mountains and the coastal and Cascade ranges, many of the invertebrates studied to date have completely lacked structure between these two disjunct patches of rainforest (e.g., the robust lancetooth snail [*Haplotrema vancouverense*], Smith et al. 2017; and taildropper slugs [*Prophysaon* spp.], Smith et al. 2018). Rather, refugial structure has been found in the Cascade Range, sometimes with deep splits between northern (Washington) and southern (Oregon) Cascades populations (Smith et al. 2018, Smith and Carstens 2020), as was previously suggested based on a study of plants from the region (Soltis et al. 1997). In other cases, there has been evidence of deep structure and multiple refugia, but the geographic locations of those refugia have been difficult to infer, as in the millipede *Chonaphe armata* (Espíndola et al. 2016).

Despite the recognized abundance of and increased interest in the invertebrates of the region, most of the small invertebrates that occupy the leaf litter have not been investigated in the PNW tem-

perate forests. Leaf-litter-dwelling invertebrates are important components of the forest ecosystem and serve an important ecological role by breaking down leaf litter (Edwards et al. 1973). Amongst the detritivores expected to be present in most temperate forest leaf litter are microsnails, small terrestrial snails that are often < 2 mm in diameter. In addition to their role in nutrient breakdown, microsnails may serve as strong indicators of habitat quality and land use (Douglas et al. 2013); despite this, they are understudied even in comparison to other terrestrial gastropods. Due to their minute size, and a combination of phenotypic plasticity in some groups and morphological stasis in others, microsnails are difficult to identify based on morphological characters (Weigand et al. 2011). The mitochondrial gene cytochrome oxidase I (COI; the classical DNA barcoding locus for animals) has been widely used for species identification and delimitation (e.g., Hebert et al. 2003, Weigand et al. 2011). The use of DNA barcoding could aid in rapid identification of microsnails and other leaf-litter-dwelling invertebrates, a potentially useful first step to understanding the composition of leaf-litter communities.

Further, DNA barcoding may reveal previously unrecognized cryptic diversity, particularly in regions like the PNW, where cryptic diversity is widespread in larger invertebrate and vertebrate lineages. Thus, in addition to using DNA barcoding to identify invertebrates sampled from the leaf litter, we took a more in-depth look at the phylogeography of two microsnails. The conical spot snail (*Punctum randolphii* Dall 1895) is found in western Washington, northwestern Oregon, the Idaho Panhandle, the Blue and Wallowa mountains of eastern Oregon, and in southwestern British Columbia (Burke 2013), with a range mirroring many other temperate rainforest endemics from the PNW. The distribution and taxonomy of the second species, the toothless column snail (*Columella edentula*) (Draparnaud 1805), has been difficult to determine. This species was originally described as *Pupa edentula* by Draparnaud (1805), with a type locality in France. In 1840, *P. simplex* (Gould 1840) was described from a single locality in Cambridge, Massachusetts, and it was noted that this species was nearly identical to *P. edentula*. The genus was later changed to *Columella.* Burke (2013) references *C. edentula* as a synonym for *C. simplex*, which is suggested to be an "unresolved" species (Turgeon et al. 1998). According to Burke (2013), *C. edentula* is distributed in North America, ranging from Alaska, through Canada and the northern United States, and south through much of the PNW and into California (Burke 2013). Though this species is not present in the most arid areas, it is thought to be present in more xeric areas than the typical temperate rainforest endemics. No previous work has investigated the phylogeographic history or potential for undescribed cryptic diversity in these two species.

In this study, we set out to generate barcode sequences for micro-invertebrates present in the PNW and to investigate the evolutionary history of two abundant microsnails using phylogeographic analyses. First, we collected and then searched leaf-litter samples, and removed all micro-invertebrates. Second, for a subset of these invertebrates, we sequenced a portion of the mitochondrial gene COI to identify the organisms. For two species of microsnails found in the leaf litter (*P. randolphii* and *C. edentula*) we combined sequence data from this study with data from GenBank to investigate whether cryptic diversity is evident in this group, or whether genetic data point to recent expansion from a single ice age refugium in the Cascade Range.

Methods

Sampling

We collected samples of leaf litter from throughout the Cascades and coastal ranges (Figure 2). At each site, we collected leaf-litter samples from one or more locations, depending on the size, heterogeneity, and composition of the sites. We collected samples from 117 locations from 41 unique sites, with a median of 3 samples per site (minimum = 1, maximum = 7, mean = 2.85). We tended to collect samples from more locations at larger sites with more macro-invertebrates present, as we hoped to maximize the number of invertebrates sampled. From each sampling location, we collected two samples: one for use in determining leaf-litter

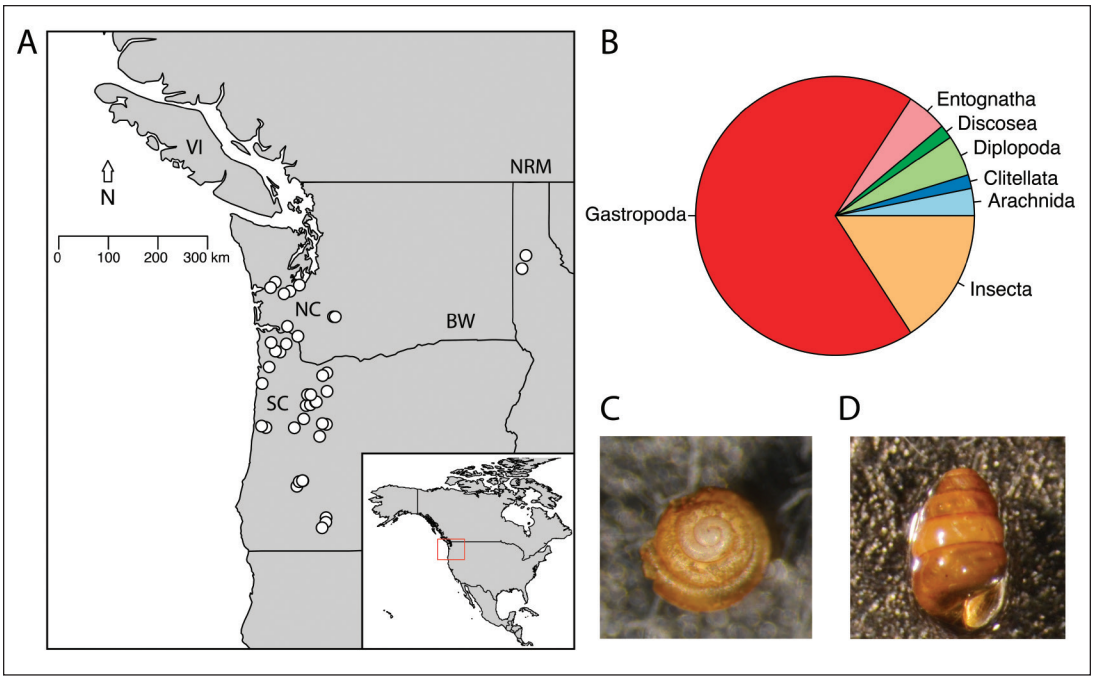

Figure 2. Sampling of leaf litter and invertebrates from the Pacific Northwest of North America in 2017. A) Map of sampling localities across the Pacific Northwest. NC: northern (Washington) Cascades, BW: Blue and Wallowa mountains, SC: southern (Oregon) Cascades, NRM: Northern Rocky Mountains, VI: Vancouver Island. B) Pie chart showing the classes of the sequenced invertebrates, based on Basic Local Alignment Search Tool (BLASTn) results. C) Photograph of a *Punctum randolphii* sample sequenced in this study. D) Photograph of *Columella edentula*-PNW sample sequenced in this study.

pH (pH samples) and one for use in measuring leaf-litter moisture content (moisture-content samples) because pH and leaf-litter moisture content have been shown to be of importance in the assembly of land snail communities (Barker and Mayhill 1999, Hylander et al. 2005). Leaves were collected from near locations where macroinvertebrates—specifically, terrestrial gastropods—were sampled, and pH samples included deciduous leaves and coniferous needles. The pH samples were collected in 709-mL Tupperware containers. For the moisture-content samples, we collected ten deciduous leaves from the forest floor; these ten leaves were weighed immediately upon collection and again later, after drying, to assess moisture content of the leaf litter at the site following Hadley (2009). All leaf-litter samples were frozen immediately to prevent degradation of organisms. Later, leaf-litter samples were thawed, and each leaf-litter sample was sorted. The pH samples were searched for 45 minutes under light and 10× magnification, and for moisture samples, each leaf was searched carefully under light and 10× magnification until all leaves had been searched. Each organism that we found was initially identified to broad taxonomic categories (i.e., Class) and placed in 95% ethanol. After removing micro-invertebrates from the leaf litter, all leaf-litter samples were dried in a dehydrator for 150 minutes at 95 °C. The moisture-content samples were again weighed, the final weight was subtracted from the original weight, and leaf-litter moisture content was estimated as the proportion of the dry weight for each sample (Hadley 2009). The pH samples were ground using a mortar and pestle and sent to the STAR laboratory at Ohio State University to measure pH.

At all sites where leaf litter was collected, we also collected elevation, slope aspect and angle, the proportion of deciduous versus coniferous

trees, and canopy cover, as these variables are also thought to be important for terrestrial gastropod communities (Barker and Mayhill 1999, Hylander et al. 2005). To measure elevation and to record latitude and longitude, we used a Garmin eTrex GPS. We measured slope degree and aspect using a clinometer and compass, respectively. To estimate the proportion of coniferous versus deciduous trees, we identified the ten closest trees to the point of collection and recorded the proportion that were coniferous. We used a densiometer (Forestry Suppliers Spherical Crown Densiometer, Convex, Rapid City, SD) to estimate canopy cover at each point of collection following manufacturer's standard protocols.

DNA Extraction and Sequencing

DNA was extracted using DNeasy Blood and Tissue kits (Qiagen, Germantown, MD). Since destructive sampling was necessary to obtain DNA due to the small size of the invertebrates, photos were taken of each invertebrate prior to DNA extraction using a Nikon D3100 digital SLR camera and a Leica M80 dissecting scope. We then added 180 uL of Buffer ATL and 20 uL of Proteinase K to the invertebrates and placed them into a 56 °C water bath for 90 minutes. Samples were removed from the water bath, and any remaining shell or exoskeleton was removed and stored in 95% ethanol. The remainder of the DNA extraction was performed using the manufacturer's standard protocol (Qiagen). A 710-bp portion of the COI gene was amplified via PCR using the HCO2198 and LCO1490 primers (Folmer 1994), cleaned using ExoSAP-IT (Thermo Fisher Scientific, Waltham, MA), and sequenced at the Ohio State University Biomedical Research Tower. Forward and reverse chromatograms were assembled in Geneious Prime v2019.0.4 (Kearse et al. 2012). Invertebrates were identified by searching the NCBI nucleotide database using the Basic Local Alignment Search Tool (BLASTn) program (McGinnis and Madden 2004) implemented within Geneious. We searched the nucleotide collection, retrieved matching regions, and allowed a maximum of 100 hits.

Based on the number of specimens that were collected and identified using BLASTn, we selected two microsnail species, *C. edentula* and *P. randolphii,* for downstream analysis. In addition to the data collected here, we downloaded 60 sequences of *Columella* from GenBank (Ansart et al. 2014; Hebert et al. 2014; Telfer et al. 2015; Harl et al. 2017; DeWaard 2017a, 2017b), including 12 identified as *C. edentula* as well as sequences from the Barcode of Life Data System (BOLD; Ratnasingham and Hebert 2007). Most of these BOLD sequences were products of the International Barcode of Life Initiative, and some were identified only as Order Stylommatophora, but grouped with other *Columella* sequences. Additionally, a sequence from Chonrinidae (*Granaria frumentum illyrica*) was downloaded from GenBank and used as an outgroup, as this was the closest outgroup available (Harl et al. 2017). We retrieved 15 available *Punctum* COI sequences from GenBank in March 2019, including eight samples identified as *P. randolphii* (Hebert et al. 2014; DeWaard, 2017a, 2017b). We downloaded additional sequences from the genus *Punctum* from BOLD. Some were identified only as Order Stylommatophora but grouped with other *Punctum* sequences. Additionally, a sequence from one species within Punctidae, but not from *Punctum* (i.e., *Paralaoma servilis*), was downloaded from GenBank and used as an outgroup (DeWaard 2017b). Sequence alignment was performed using the MUSCLE alignment algorithm (Edgar 2004) for *Punctum* and *Columella* separately, implemented within Geneious using default settings.

Gene Tree Estimation

We used the AutoModel function in Phylogenetic Analysis Using Parsimony* (PAUP* and other methods) v4.0a (Swofford 2001) to select the best model of sequence evolution for *Punctum* and *Columella,* separately. As a starting tree, we used a neighbor-joining tree calculated from LogDet/paralinear distances. We considered seven substitution schemes, gamma rate variation across sites (+G), and a proportion of invariable sites (+I) and compared models using corrected Akaike information criteria corrected for small sample

size (AICc), Bayesian information criteria (BIC), and decision theory (DT).

We used GARLI v2.0.1 (Zwickl 2006) to estimate the maximum-likelihood gene tree for *Punctum* and for *Columella*, using the model of nucleotide substitution selected using DT in PAUP*. Stepwise addition using a random sequence order and a maximum likelihood (ML) criterion was used to generate a starting tree, and state frequencies, the shape of the gamma distribution, and substitution rates were estimated from the data. When gamma rate variation was included, we used four categories. We conducted two runs and stopped each run when the likelihood did not improve for 20,000 generations, with an improvement defined as a change of more than 0.01 units. Then 100 bootstrap replicates were performed using the same settings as above to assess nodal support. To estimate the posterior distribution of gene trees and parameters, we used MrBayes v3.2.7a (Ronquist et al. 2012). We used the model of nucleotide substitution selected based on DT in PAUP*. We conducted two independent runs with 5,000,000 generations and four chains per run. When gamma rate variation was included, four categories were used, as in the ML analyses in GARLI. For all other settings and priors, default settings were used. We discarded the initial 25 percent of runs as burn-in, and evaluated convergence using the average deviation of split frequencies.

Estimating the Number of Refugia

We used approximate Bayesian computation (ABC) to estimate the number of refugia. We performed this analysis both with and without *a priori* population assignments based on geography. For all analyses aimed to estimate the number of refugia, we used only samples that grouped with *C. edentula* from the PNW in the inferred gene trees (*C. edentula*-PNW, see Results). For *P. randolphii*, we considered only *P. randolphii* from the green clade in Figure 3, as we considered the other samples to likely be misidentified (see Results and Discussion).

To assess the number of refugial populations without making *a priori* assumptions about individual assignments to particular populations, we used ABC to estimate the number of refugia during the LGM for each species while integrating over population assignments (Supplemental Figure S1). First, we calculated the Euclidean distance between samples using the python package 'SciPy' (Lindblad and Kisner 2013). Next, we simulated 100,000 datasets for each species. For each simulation, we first assigned individuals to one of three populations using K-Means clustering with randomization. Individuals were initially assigned to populations based on K-Means clustering implemented in the python package 'Scikit-learn' (Pedregosa et al. 2011). To ensure some random noise in this process (i.e., to prevent population assignments from being constant across simulations), we used only a single iteration of K-Means clustering in this step. Following initial assignment, individuals were reassigned to a population with a probability of 0.2. We required that at least two individuals were sampled from each population. This clustering approach was designed to create a balance between taking geographic information into account when testing for refugial structure and relying completely on geographic information to determine population assignments.

Based on these assignments, data were simulated under seven models (about 14,300 datasets under each model). Model 1 included a single population expanding from a refugium after the LGM. Models 2 to 4 included two populations expanding from isolated refugia after the LGM and differed in which populations were collapsed. Models 5 to 7 included three populations expanding from isolated refugia after the LGM and differed in topology. Simulated data matched the observed data in terms of the number of segregating sites, which was calculated using a custom python script. We ignored sites with missing data when calculating the number of segregating sites. Growth rates were drawn from uniform U(1,50) priors. Divergence times between populations and the timing of expansion were also drawn from uniform priors. The minimum bound for the divergence time priors was based on a time of 115,000 years ago (beginning of last glacial cycle), a population size of 1,000,000, and species-specific generation lengths (from the literature). For the generation

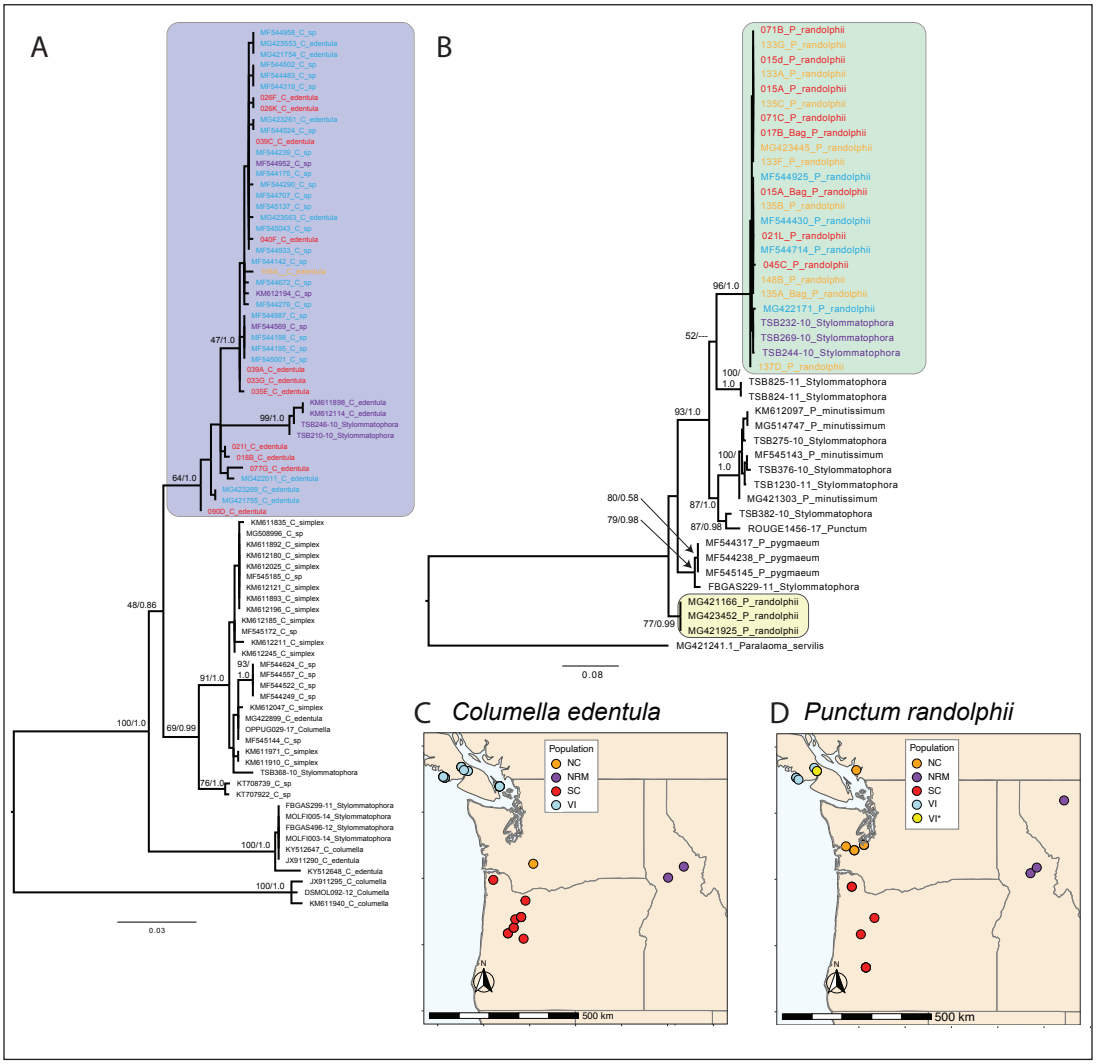

Figure 3. Gene trees for all *Columella* and *Punctum* sequences from this study, GenBank, and Barcode of Life Data System (BOLD), and sampling localities in the Pacific Northwest of North America (PNW). A) Maximum likelihood estimate (MLE) of the cytochrome oxidase I (COI) gene tree for *Columella.* Numbers on the nodes represent bootstrap values (BS) / posterior probabilities (pp). Samples within *C. edentula*-PNW are color coded corresponding to the map in C; B) MLE of the COI gene tree for *Punctum.* Numbers on the nodes represent bootstrap values (BS) / posterior probabilities (pp). Samples of *P. randolphii* are color coded corresponding to the map in D; C) Sampling map for *C. edentula*-PNW, including GenBank samples and samples from this study. These are only the samples from the group highlighted in blue on the COI tree in A. Colors correspond to sampling regions; D) Sampling map for *P. randolphii,* including GenBank samples and samples from this study. The yellow samples are the three divergent samples from GenBank identified as *P. randolphii.*

length parameter, we used one year as the generation time, given that many terrestrial gastropods reach sexual maturity and have a generation time of one year. The upper bound for the divergence time priors was based on a time of 5,000,000 years ago, a population size of 10,000, and a generation length of one year. The minimum and maximum bounds on expansion times (the time when expansion stopped) were calculated similarly, with the minimum time based on 10,000 years ago and the maximum time based on 20,000 years ago. We required that the first divergence time predate the

second, and that expansion times predated divergence times. We used Hudson's ms (Hudson 2002) to simulate data and calculated summary statistics using a Perl script (Pelletier and Carstens 2014).

We calculated π and Tajima's D (Tajima 1989) from the empirical data using the R package 'Popgenome' (Pfeifer et al. 2014), and we plotted histograms of simulated and observed data to verify that our observed data fell within the range of the simulated data for the two summary statistics. Finally, we used the R package 'abc' (Csilléry et al. 2012) to assess power and perform model selection. Prior to assessing power, we collapsed our models to one (model 1), two (models 2 to 4), or three (models 5 to 7) refugia models. To assess power, we used the function cv4abc. Briefly, we considered 100 simulated datasets from each of the three models (one, two, or three refugia) and treated those datasets as pseudo-observed data. We then used a simple rejection step, a tolerance of 0.001, and the summary statistics π and Tajima's D to find the simulated datasets closest to these pseudo-observed datasets. Finally, we calculated the percentage of time the best model was selected when data were simulated under each model. Next, we used the postpr function to calculate the posterior probability of each of the three models for the empirical data. As before, we used a simple rejection step and a threshold of 0.001, with the summary statistics π and Tajima's D.

To estimate the number of refugia in a more traditional framework, we performed ABC with individuals assigned to populations based on geographic groupings. For *C. edentula*-PNW, we considered three populations: the South Cascades, Vancouver Island, and the Northern Rocky Mountains. We omitted a single sample (MLS2017F-106A) from this analysis, because it was the only sample collected from the northern Cascades. We considered models with one, two, or three refugial populations, allowing for all possible collapsed populations and all possible topologies. To account for different possible topologies, we simulated 100,000 datasets, but weighted our simulations such that the two and three refugia models (each consisting of three submodels) were three times as likely to be chosen for any particular simulation compared to the single refugium model. Priors, cross-validation, and model selection were as above. For *P. randolphii*, we considered four populations: the northern Cascades, the southern Cascades, Vancouver Island, and the Northern Rocky Mountains. We considered models with one, two, three, or four refugial populations, and simulated 100,000 datasets weighting by the number of possible topologies such that each potential topology had a roughly equal number of simulated datasets. Priors, cross-validation, and model selection were as above. We performed all ABC analyses with and without samples identified only to Order Stylommatophora to evaluate whether taxonomic uncertainty and misidentifications could affect our results. For *C. edentula*-PNW, we also removed all samples that grouped with the focal clade in our mitochondrial gene tree but were only identified as *C.* spp. In the case of *P. randolphii*, this resulted in removing all samples from the NRM, and thus the analysis using *a priori* assignments based on geography considered only one, two, or three refugia.

Species Distribution Models

To estimate species distribution models for *P. randolphii* and *C. edentula*-PNW, we used the R package 'biomod2' (Thuiller et al. 2013). We downloaded climate data from the current (30 arc-seconds) and from the LGM (2.5 minutes) from worldclim (Hijmans et al. 2005). We limited the extent of these data to between –150 and –100 degrees longitude and between 35 and 65 degrees north latitude. We excluded highly correlated variables ($r > 0.7$) and used seven variables in our models (bio4, bio5, bio8, bio14, bio15, bio18, bio19). We used occurrence data from our collections and from GenBank. We used 10,000 pseudoabsences drawn at random from the background. To estimate the species distribution models (SDMs), we used an ensemble approach, combining models from four modelling methods: 1) Maxent; 2) general linear models (GLM), 3) random forest (RF); and 4) generalized boosting models (GBM). We used 80 percent of the data for training and 20 percent for testing. Five runs were used to evaluate models, and three runs were used to estimate variable importance. Models

were evaluated using ROC (receiver operator characteristic) and rescaled to allow projections to different time periods. When building the ensemble models, we ignored models with a ROC less than 0.85, and weighted models by ROC. Then, models were projected onto both the present climate layers and onto the LGM climate layers.

Results

Sampling

We sorted through 105 pH samples and 117 moisture samples and searched for micro-invertebrates. In total, we removed 401 invertebrates from the leaf-litter samples (Figure 2a; Supplemental Table S1). Of these, the majority (349) were found in the pH samples, as expected, since these contained larger amounts of leaf litter. Of the sampled invertebrates, 109 were microsnails.

DNA Extraction and Sequencing

We sequenced 60 invertebrates from the leaf litter, including 39 snails. Of the other 21 sequenced specimens, 18 were identified as belonging to the Phylum Arthropoda, with one sample each from Annelida and Amoebozoa. Of the 18 Arthropoda, 10 belonged to Class Insecta and the others belonged to Classes Arachnida, Diplopoda, and Entognatha (Supplemental Table S2; Figure 2b). The quality and quantity of DNA that we were able to extract from these samples was low, likely because samples were collected from leaf litter that was stored for several weeks in the field after collection (though the leaf litter was frozen during this time), some individual snails were damaged and had died prior to sampling, and all samples were very small (i.e., microsnails were < 2 mm in diameter). Of the microsnails we were able to sequence, we extracted a minimum of 9.3 ng of DNA, a median of 102 ng of DNA, and a maximum of 2,295 ng of DNA in 150 uL. The low levels of DNA extracted from these samples placed a practical limitation on the amount of data we were able to obtain in this study (Supplemental Table S2) and limited our data to high copy mtDNA sequences. Including our samples plus those from GenBank and BOLD, we obtained 83 *Columella* sequences, 24 of which were initially identified as *C. edentula* (Supplemental Table S3, Figure 3c). The *Columella* sequences collected for this study ranged from 602 to 652 bp in length. We obtained 42 *Punctum* sequences, including 24 sequences identified as *P. randolphii* (Supplemental Table S4, Figure 3d). The *P. randolphii* sequences collected in this study ranged from 526 to 655 bp in length.

Gene Tree Estimation

For *Columella*, the best model of nucleotide substitution was HKY+I+G, based on AICc, BIC, and DT. The gene tree estimated in GARLI showed deep divergence between *Columella* and the outgroup, making it difficult to discern relationships within *Columella* (Supplemental Figure S2). Because of the lack of differentiation within *Columella*, we estimated gene trees for the ingroup only and rooted using midpoint rooting. The best model of nucleotide substitution without the outgroup was HKY+I+G, based on AICc, BIC, and DT. The ΔAICc between the HKY+I+G and the simplest model (JC) was 398, the ΔBIC was 381, and the ΔDT was 0.112. The ΔAICc between HKY+I+G and the second-best model (K81uf+I+G) was 2.27, and the ΔBIC was 5.18. The second-best model according to DT criteria was TrN+I+G, and the ΔDT between this model and HKY+I+G was 0.000840. The MLE of the midpoint rooted gene tree included a clade comprised of *C. edentula* and *C.* spp. samples, as well as two samples from BOLD identified only as Stylommatophora, collected from the PNW, in the range described by Burke (2013; Figure 3a). This clade had low to moderate support (BS = 64) in the bootstrap analysis, but high support (posterior probabilities [pp] = 1.0) in the Bayesian analysis. Additionally, both ML and Bayesian analyses support a clade containing *C. simplex*, *C.* spp., a single *C. edentula* sample, and one Stylommatophora sample, which includes samples from the PNW and from the eastern United States. A third group includes one *C. columella* sample (from Austria), two *C. edentula* samples (one from Austria, one from unknown locality), and four samples from BOLD identified only as Order Stylommatophora from Germany and Finland. A final clade includes three *C. columella* samples (two from Canada,

one from unknown locality). For downstream analyses, we considered only those samples grouping in the *C. edentula*-PNW clade (*C. edentula*-PNW, Figure 3a, blue group). Due to the complex and unclear taxonomy of *Columella*, it was difficult to determine the specific status of this and other clades, but given the divergence between clades, and given that the all samples collected from the focal area of this study (Cascades and coastal ranges of the PNW) belonged to a single clade, we chose this clade as the unit of analysis moving forward (hereafter, *C. edentula*-PNW).

For *Punctum,* the best model of nucleotide substitution was HKY+G based on AICc, BIC, and DT. The ΔAICc between the HKY+G and the simplest model (JC) was 402, the ΔBIC was 382, and the ΔDT was 0.527. The ΔAICc between HKY+G and the second-best model (HKY+I+G) was 0.382, and the ΔBIC was 4.19. The second-best model according to DT criteria was K81uf+G, and the ΔDT between this model and HKY+G was 0.00661. We used this model to estimate the MLE of the gene tree in GARLI and the posterior distribution of gene trees in MrBayes. The MLE of the gene tree showed five main groups (Figure 3b): *P. randolphii* (green), *P. minutissimum, P. pygmaeum, P. randolphii* (yellow), and a group with two sequences identified only as Stylommatophora from California. All five groups received high bootstrap support and posterior probabilities (Figure 3b). We considered only samples from the main *P. randolphii* clade to be *P. randolphii* (*P. randolphii* sensu stricto), as the other three *P. randolphii* samples from GenBank were highly divergent. Given that these samples were divergent from other *P. randolphii* samples, we conservatively treat them as not belonging to *P. randolphii* sensu stricto in downstream analyses. These samples may represent undescribed microsnail diversity from the region or may belong to a described species for which no sequenced samples were available.

Estimating the Number of Refugia

For *C. edentula*-PNW, the observed data included 26 segregating sites (Table 1). When integrating over individual assignments to populations, we selected the best model during cross-validation 95, 79, and 84% of the time for the one, two, and three refugia models, and the best model was a single refugium model (pp = 0.95), with strong support over the second-best model (two refugia, BF = 19). When using *a priori* assignment of individuals to geographic populations, we selected the best model during cross-validation 97, 98, and 94% of the time for the one, two, and three refugia models, respectively, and the best model was a single refugium model (pp = 1.0). We conducted the same analysis, but removed all sequences downloaded from GenBank or BOLD that were in the focal clade in our mitochondrial gene tree but were not identified as *C. edentula* in the corresponding databases (24 samples), and results were qualitatively consistent, with the single refugium model always supported as the best model.

TABLE 1. Summary statistics for *Columella edentula*-PNW, *C. edentula*-PNW with northern Cascades samples omitted (no NC), and *Punctum randolphii* samples from the Pacific Northwest of North America. S is the number of segregating sites, π is nucleotide diversity, and Tajima's D is a test statistic based on allele frequencies that is often used to detect deviations from neutrality and population size changes.

Species	S	π	Tajima's D
C. edentula-PNW	26	4.45	–0.74
C. edentula-PNW (no NC)	26	4.46	–0.75
P. randolphii	6	0.71	–1.69

For *Punctum*, the observed data included six segregating sites (Table 1). Without using *a priori* assignments of individuals to populations, we selected the best model 86, 89, and 77% of the time for the one, two, and three refugia models, respectively, during cross-validation, and the best model was the single refugium model (pp = 0.98), with high support over the second-best model (two refugia, BF = 49.00). When we used *a priori* assignments of individuals to populations, we selected the best model 87, 93, 62, and 59% of the time for the one, two, three, and four refugia models, respectively. Though these results seem poor for the three and four refugia models, they were mistaken for similar models. Specifically,

the three-refugia model was most often mistaken as either a two- or four-refugia model, and the four-refugia model was most often mistaken for the three-refugia model. The best model was a single refugium model (pp = 0.97), and it had strong support over the second-best model (two refugia; BF = 32.3). We conducted the same analysis, but removed all sequences downloaded from GenBank or BOLD that were only identified to the Order Stylommatophora (three samples), and results were qualitatively consistent, with the single refugium model always supported as the best model, with strong support as quantified by Bayes factors.

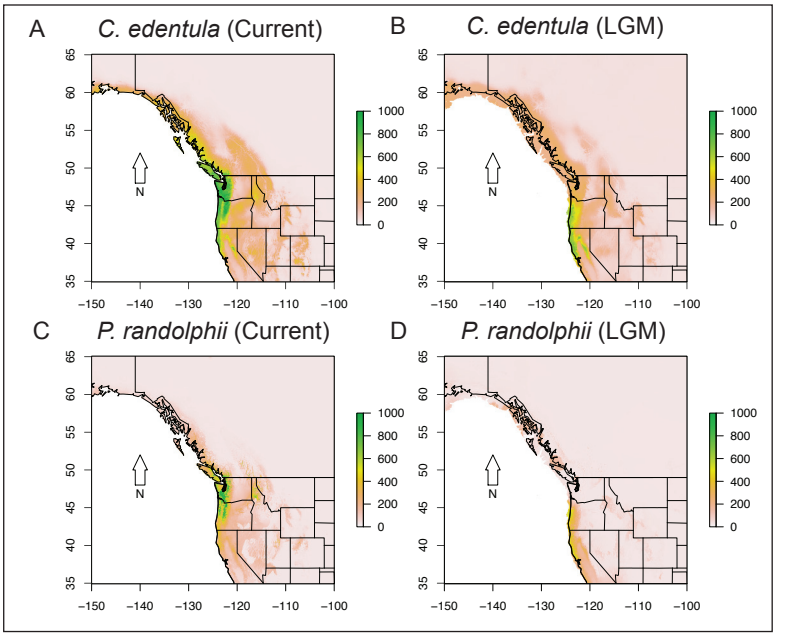

Figure 4. Species distribution models reconstructed for *Punctum randolphii* and *Columella edentula*-PNW from North America, with colors showing ecological suitability for each, as specified in the legend. A suitability of 0 indicates that the habitat is not predicted to be suitable for the species, while a suitability of 1000 is the highest possible suitability. X-axes in figures represent degrees north latitude and y-axes represent degrees longitude. A) *C. edentula*-PNW in the present; B) *C. edentula*-PNW during the Last Glacial Maximum (LGM); C) *P. randolphii* in the present; D) *P. randolphii* in the Last Glacial Maximum.

Species Distribution Models

For *C. edentula*-PNW, the average ROC score was 0.89 (Supplemental Table S5), and for *P. randolphii,* the average ROC score was 0.96 (Supplemental Table S6). For both species, we recovered the current range of the spaces with high predicted suitable climatic habitat, both in the coastal and Cascade ranges, and moderate-to-high predicted suitable climatic habitat in the Northern Rocky Mountains (Figure 4). During the LGM, we hindcasted suitable habitat primarily on the coast and farther south, and we observed less suitable habitat overall (Figure 4). This suggests that suitable habitat has expanded since the LGM for both *P. randolphii* and *C. edentula*-PNW.

Discussion

Sampling of Invertebrates

We were able to sample a wide array of invertebrates by collecting leaf-litter samples from the temperate rainforests of the coastal and Cascade ranges (Figure 2). While this sampling approach required many hours spent in the lab searching through leaf-litter samples, we were able to identify the majority of samples using the mitochondrial marker COI (Supplemental Table S2). This result suggests that metabarcoding may be a viable approach for identifying micro-invertebrates from temperate rainforest communities. It is similar in some aspects to the use of environmental DNA (eDNA), which has been demonstrated as a valuable tool for studying biodiversity (reviewed in Bohmann et al. 2014). Metabarcoding of leaf-litter samples has previously been suggested as a viable way to assess biodiversity in forests in China and Vietnam with a focus on identifying arthropods (Yang et al. 2014). Metabarcoding could greatly enhance our knowledge of these understudied leaf-litter-dwelling invertebrates and could expand the

phylogeographic knowledge of the PNW to a much broader array of taxa. Our work builds upon the current database of reference sequences and photos for future work using environmental sampling techniques. Notably, photographing such small specimens, even using a dissecting scope, was rather difficult, given the difficulty of arranging the specimens appropriately for photographs. Dedicated equipment for imaging specimens about 2 mm in size would greatly improve the quality of photos, and new techniques like microCT scanning offer potential approaches to collect more useful phenotypic data from such specimens. The ecological data we collected here allowed us to make predictions about optimal sampling localities for metabarcoding studies and may help in focusing sampling efforts for future work in this region aimed at collecting the species studied here.

Pacific Northwest Phylogeography

The two microsnail species, *P. randolphii* and *C. edentula*-PNW, have ranges that are similar to those of larger invertebrates and invertebrate taxa from the PNW (Burke 2013), but the phylogeographic patterns found here differ profoundly from other invertebrates that have been studied from the region. Despite the limitations inherent to a single genetic marker (discussed below), we can begin to understand how the Pleistocene glacial cycles affected the coastal distributions of these species. The lack of deep structure observed in the gene trees across the coastal ranges of these samples suggests a lack of population structure (Figure 3). Furthermore, both ABC analyses suggest the presence of a single refugium during the Pleistocene glacial cycles for *C. edentula*-PNW and *P. randolphii*.

While we expected population structure to be present in the taxa based on assumed limited dispersal capacity and the lack of previous work in the group, there are a few potential explanations for the lack of divergence. First, consistent with previous suggestions that microsnails may be good indicators of habitat quality (Douglas et al. 2013), the microsnails in our sample may be particularly vulnerable to habitat disturbance. If these microsnails are sensitive to disturbances, Pleistocene glacial cycles may have eliminated them from large portions of their ranges, and they may have recently recolonized much of the region. This interpretation is supported by SDMs, which indicate that suitable habitat was substantially contracted for these snails during the LGM (Figure 4). Further, our ABC results support a history of recent population expansion from a single refugium during the LGM, which is consistent with the lack of divergence shown by the gene trees (Figure 3). However, an alternative interpretation is also possible: microsnails may be less dispersal limited than initially expected. Microsnails have been reported to be dispersed via the digestive tracts of birds (Wada et al. 2012), and terrestrial slugs have been found in bird feathers (Pearce et al. 2012), so animal-facilitated dispersal may lead to high levels of migration across the ranges of microsnails, and thus, to a lack of genetic structure.

Finally, the lack of population structure uncovered in this study may also be an artefact of limited sampling. Though our sampling did greatly expand the data available for these two microsnails, there are portions of each species' range that remain under-sampled. For *P. randolphii,* the inland rainforests were not well sampled, and it is possible that a second refugial population could have survived in these ranges and was not recovered in our dataset. There is also potential for a second clade of *P. randolphii* consisting of the three GenBank samples from Vancouver Island that form a separate clade in the COI gene tree (Figure 3b). However, given that these sequences were generated from samples from an unpublished study, it is difficult to verify their identity as *P. randolphii*, and we suspect that these samples may represent undescribed diversity from the region. Future work should aim to collect additional samples from this region to evaluate whether undescribed diversity is present.

For *C. edentula*-PNW, we lack samples from more inland stretches of the species range, as well as for the most southern stretches and more arid portions of the species range. For this species, it is likely that our sampling indicates only a

single refugium within the Cascade Range, and not a single glacial refugium for the species as a whole. Further, the taxonomy and distribution of *C. edentula* has been difficult to resolve, and it is unclear whether the specimens collected here represent *C. edentula sensu stricto* or a different species. Future work should aim to collect *Columella* more broadly and to resolve taxonomic issues in this genus.

Despite the uncertain taxonomy of our focal species specifically and microsnails in general, we suspect that it is unlikely that misidentified specimens have altered our results. Given the low level of diversity observed across the sampled populations of the two species, we view it unlikely that our samples inadvertently included other species. Further, if our analyses did include such samples, we would suspect that this would bias us away from inferring a single refugium by inflating our estimates of diversity. However, we consistently inferred a single refugium. Further, when we removed samples identified only to the level of order or genus from our analyses, our results remained consistent, suggesting that taxonomic uncertainty and misidentification were unlikely to have misled our core analyses. Even after acknowledging these shortcomings, our sampling and the data collected here supported only a single Cascades refugium for both species, a pattern that contrasts with results from previously studied invertebrates (e.g., Smith et al. 2017, Smith and Carstens 2020).

Despite the promise of eDNA and metabarcoding, there are inherent limitations to studies based on a single gene. The lack of recombination of the mitochondrial genome, its reduced effective population size and generally high mutation rates both make mitochondrial DNA a powerful marker for detecting population structure and result in a limited view of the phylogeographic history of a species (Ballard and Whitlock 2004). Additionally, selection and introgression could affect phylogeographic inference made based on mitochondrial DNA alone (Ballard and Whitlock 2004). By using a mitochondrial marker, we observed only a single realization of the coalescent process, so additional genetic loci would improve inferences of the phylogeographic histories of these species. However, the data we were able to collect in this case were limited due to the difficulties of sampling micro-invertebrates and of extracting high-quality DNA from these invertebrates. Despite this, this study provides a substantial increase in the amount of data available for this system. Future work should aim to take advantage of constantly improving sequencing technologies to sequence nuclear markers from low-quality DNA, or to extract higher quality DNA from small and often damaged samples.

Our results add considerably to our knowledge of species responses to glacial cycles in the PNW. Other invertebrates have exhibited signals of multiple refugia in the coastal and Cascade ranges (Smith et al. 2018), and a lack of structure as extreme as that shown here has not been shown in invertebrates from the region, to our knowledge. By sampling leaf litter, we were able to study these previously unstudied microsnails and to begin to gather samples of many other leaf-litter-dwelling invertebrates. Continuing to expand sampling for such invertebrates, via metabarcoding sampling or other methods, will contribute substantially to our understanding of how communities responded to historical climatic changes.

Data Accessibility

All sequences are available on the BOLD (Smith et al. 2021a). Reference photos are available on FigShare (Smith et al. 2021b).

Acknowledgments

MLS was funded by a NSF GRFP (DGE-1343012), and this work was funded by NSF (DEB1457519 and DEB1457726.). DS was funded by the Undergraduate Research Scholarship from the College of Arts and Sciences Honors Program at Ohio State University. We would like to thank members of the Carstens Lab for comments that improved the manuscript prior to publication. We would also like to thank Katherine Field and Andrew Rankin for assistance with fieldwork for this study and Sarah Blausey for assistance with labwork.

Literature Cited

Ansart, A., A. Guiller, O. Moine, M. C. Martin, and L. Madec. 2014. Is cold hardiness size-constrained? A comparative approach in land snails. Evolutionary Ecology 28:471-493.

Ballard, J. W. O., and M. C. Whitlock. 2004. The incomplete natural history of mitochondria. Molecular Ecology 13:729-744.

Barker, G. M., and P. C. Mayhill. 1999. Patterns of diversity and habitat relationships in terrestrial mollusc communities of the Pukeamaru Ecological District, northeastern New Zealand. Journal of Biogeography, 26:215-238.

Bohmann, K., A. Evans, M. T. P. Gilbert, G. R. Carvalho, S. Creer, M. Knapp, D. W. Yu, and M. De Bruyn. 2014. Environmental DNA for wildlife biology and biodiversity monitoring. Trends in Ecology & Evolution 29:358-367.

Brunsfeld, S. J., J. Sullivan, D. E. Soltis, and P. S. Soltis. 2001. Comparative phylogeography of northwestern North America: a synthesis. Special Publication-British Ecological Society 14:319-340.

Burke, T. 2013. Land Snails and Slugs of the Pacific Northwest. Oregon State University Press, Corvallis.

Carstens, B. C., S. J. Brunsfeld, J. R. Demboski, J. M. Good, and J. Sullivan. 2005. Investigating the evolutionary history of the Pacific Northwest mesic forest ecosystem: hypothesis testing within a comparative phylogeographic framework. Evolution 59:1639-1652.

Carstens, B. C., A. L. Stevenson, J. D. Degenhardt, and J. Sullivan. 2004. Testing nested phylogenetic and phylogeographic hypotheses in the *Plethodon vandykei* species group. Systematic Biology 53:781-792.

Csilléry, K., O. François, and M. G. B. Blum. 2012. abc: an R package for approximate Bayesian computation (ABC). Methods in Ecology and Evolution 3:475-479.

Daubenmire, R., and J. B. Daubenmire. 1968. Forest vegetation of eastern Washington and northern Idaho. Technical Bulletin 60. Washington State University, College of Agriculture, Washington Agricultural Experiment Station, Pullman.

Delcourt, P. A., and H. R. Delcourt. 1993. Paleoclimates, paleovegetation, and paleofloras during the late Quaternary. *In* Flora of North America Editorial Committee (editors), Flora of North America, Vol. 1, Oxford University Press, New York. Pp.71-94.

Dewaard, J. R. 2017a. Canadian National Parks Data Release. Available online at http://www.boldsystems.org/ (accessed 05 May 2020).

Dewaard, J. R. 2017b. BIOUG Archive GGBN Data Release. Available online at http://www.boldsystems.org/ (accessed 05 May 2020).

Douglas, D. D., D. R. Brown, and N. Pederson. 2013. Land snail diversity can reflect degrees of anthropogenic disturbance. Ecosphere 4:1-14.

Draparnaud, J. P. R. 1805. Histoire naturelle des mollusques terrestres et fluviatiles de la France. Levrault, Plassan, Paris.

Dyke, A. S., A. Moore, and L. Robertson. 2003. Deglaciation of North America, Open File 1574, Natural Resources Canada, Ottawa.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research 32:1792-1797.

Edwards, C. A., D. E. Reichle, and D. A. Crossley. 1973. The role of soil invertebrates in turnover of organic matter and nutrients. *In* D. E. Reichle (editor), Analysis of Temperate Forest Ecosystems, Springer Science and Buisiness Media. Pp. 147-172.

Espíndola, A. E., M. Ruffley, M. L. Smith, B. C. Carstens, D. C. Tank, and J. Sullivan. 2016. Identifying cryptic diversity with predictive phylogeography. Proceedings of the Royal Society B: Biological Sciences 283:20161529.

Folmer, O. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3:294-299.

Gould, A. A. 1840. Monograph of the species of *Pupa* found in the United States; with figures. Boston Journal of Natural History 3:395-404.

Graham, A. 1999. Late Cretaceous and Cenozoic History of North American Vegetation. Oxford University Press, New York.

Hadley, J. 2009. Leaf Litter Moisture Content at Harvard Forest HEM and LPH Towers 2006. Harvard Forest Data Archive: HF157 (v.12). Available online at https://doi.org/10.6073/pasta/9ff1822986ab67c1695f4b69453ef886 (accessed 05 May 2020).

Harl, J., E. Haring, T. Asami, M. Sittenhaler, H. Sattmann, and B. Páll-Gergely. 2017. Molecular systematics of the land snail family Orculidae reveal polyphyly and deep splits within the clade Orthurethra (Gastropoda: Pulmonata). Zoological Journal of the Linnean Society 181:778-794.

Hebert, P. D. N, A. Cywinska, S. L. Ball, and J. R. Dewaard. 2003. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London B: Biological Sciences 270:313-321.

Hebert, P. D. N., S. Ratnasingham, E. V. Zakharov, A. C. Telfer, V. Levesque-Beaudin, M. A. Milton, S. Pedersen, P. Jannetta, and J. R. Dewaard. 2014. Barcoding Canada Data Release. Available online at http://www.boldsystems.org/ (accessed 05 May 2020).

Hijmans, R.J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 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.

Hudson, R. R. 2002. Generating samples under a Wright–Fisher neutral model of genetic variation. Bioinformatics 18:337-338.

Hylander, K., C. Nilsson, B. Gunnar Jonsson, and T. Göthner. 2005. Differences in habitat quality explain nestedness in a land snail meta-community. Oikos 108:351-361.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Cooper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, and A. Drummond. 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647-1649.

Lindblad, T., and J. M. Kisner. 2013. NumPy, SciPy and Python image library. *In* T. Lindblad and J. M. Kisner (editors), Image Processing Using Pulse-Coupld Neural Networks, Springer, Berlin, Heidelberg. Pp. 35-56.

McGinnis, S., and T. L. Madden. 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. Nucleic acids research 32:W20-W25.

Nielson, M., K. Lohman, C. H. Daugherty, F. W. Allendorf, K. L. Knudsen, J. Sullivan, D. J. Ellis, B. T. Firth, and I. Belan. 2006. Allozyme and mitochondrial DNA variation in the tailed frog (Anura: *Ascaphus*): the influence of geography and gene flow. Herpetologica 62:235-258.

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

Pearce, T. A., R. S. Mulvihill, and K. A. Porter. 2012. Land slugs (Gastropoda: Pulmonata) on birds demonstrate dispersal potential. Nautilus 126:38-40.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, and M. P. E. Duchesnay. 2011. Scikit-learn: machine learning in Python. Journal of Machine Learning Research 12:2825-2830.

Pelletier, T. A. and B. C. Carstens. 2014. Model choice for phylogeographic inference using a large set of models. Molecular Ecology 23:3028-3043.

Pelletier, T. A. and B. C. Carstens. 2016. Comparing range evolution in two western *Plethodon* salamanders: glacial refugia, competition, ecological niches, and spatial sorting. Journal of Biogeography 43:2237-2249.

Pfeifer, B., U. Wittelsbürger, S. E. Ramos-Onsins and M. J. Lercher. 2014. PopGenome: an efficient Swiss army knife for population genomic analyses in R. Molecular Biology and Evolution 31:1929-1936.

Ratnasingham, S., and P. D. N. Hebert. 2007. BOLD: The barcode of life data system. Molecular Ecology Notes 7:355-364.

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

Smith, M. L., and B. C. Carstens. 2020. Process-based species delimitation leads to identification of more biologically relevant species. Evolution 74:216-229.

Smith, M. L., C. Lang, D. Sneddon, J. Wallace, A. Espíndola, J. Sullivan, and B.C. Carstens. 2021a. Data from Comparative Phylogeography of Microsnails from the Pacific Northwest. BOLD. Dataset. dx.doi.org/10.5883/DS-LLINVERT

Smith, M. L., C. Lang, D. Sneddon, J. Wallace, A. Espíndola, J. Sullivan, and B.C. Carstens. 2021b. Data from Comparative Phylogeography of Microsnails from the Pacific Northwest. figshare. Dataset. https://doi.org/10.6084/m9.figshare.17027201.v4

Smith, M. L., M. Ruffley, A. Espíndola, D. C. Tank, J. Sullivan, and B. C. Carstens. 2017. Demographic model selection using random forests and the site frequency spectrum. Molecular Ecology 26:4562-4573.

Smith, M. L., M. Ruffley, A. M. Rankin, A. Espíndola, D. C. Tank, J. Sullivan, and B. C. Carstens. 2018. Testing for the presence of cryptic diversity in tail-dropper slugs (*Prophysaon*) using molecular data. Biological Journal of the Linnean Society 124:518-532.

Soltis, D. E., M. A. Gitzendanner, D. D. Strenge, and P. S. Soltis. 1997. Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. Plant Systematics and Evolution 206:353-373.

Steele, C. A., B. C. Carstens, A. Storfer, and J. Sullivan. 2005. Testing hypotheses of speciation timing in *Dicamptodon copei* and *Dicamptodon aterrimu*s (Caudata: Dicamptodontidae). Molecular Phylogenetics and Evolution 36:90-100.

Swofford, D. L. 2001. PAUP*: Phylogenetic analysis using parsimony (and other methods) 4.0. B5. Sinauer Associates, Sunderland, MA.

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

Telfer, A. C., M. R. Young, J. Quinn, K. Perez, C. N. Sobel, J. E. Sones, V. Levesque-Beaudin, R. Derbyshire, J. Fernandez-Triana, R. Rougerie, et al. 2015. Biodiversity inventories in high gear: DNA barcoding facilitates a rapid biotic survey of a temperate nature reserve. Biodiversity Data Journal 3:E6313.

Thuiller, W., D. Georges, R. Engler, and F. Breiner. 2013. biomod2: ensemble platform for species distribution modeling, R package version 2. Available online at https://cran.r-project.org/web/packages/biomod2/index.html (accessed 01 March 2020).

Turgeon, D. D., F. F. Quinn, A. E. Bogan, E. V. Coan, F. G. Hochberg, W. G. Lyons, P. M. Mikkelsen, R. J. Neves, C. F. E. Roper, G. Rosenberg, B. Roth, A. Scheltema, F. G. Thompson, M. Vecchione, and J. D. Williams. 1998. Common and Scientific Names of Aquatic Invetebrates from the United States and Canada: Mollusks. 2nd ed. American Fisheries Society Publication 26. Bethesda, MD.

Wada, S., K. Kawakami, and S. Chiba. 2012. Snails can survive passage through a bird's digestive system. Journal of Biogeography 39:69-73.

Weigand, A. M., A. Jochum, M. Pfenninger, D. Steinke, and A. Klussmann-Kolb. 2011. A new approach to an old conundrum—DNA barcoding sheds new light on phenotypic plasticity and morphological stasis in microsnails (Gastropoda, Pulmonata, Carychiidae). Molecular Ecology Resources 11:255-265.

Yang, C., X. Wang, J. A. Miller, M. De Blécourt, Y. Ji, C. Yang, R. D. Harrison, and D. W. Yu. 2014. Using metabarcoding to ask if easily collected soil and leaf-litter samples can be used as a general biodiversity indicator. Ecological Indicators 46:379-389.

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.

Supplemental material available online at http://www.bioone.org/loi/nwsc

Received 27 July 2020
Accepted 11 April 2022