The role of multiple Pleistocene refugia in promoting diversification in the Pacific Northwest

Article *in* Molecular Ecology · July 2022

DOI: 10.1111/mec.16595

5 authors, including:

Megan Smith
Indiana University Bloomington
27 PUBLICATIONS **249** CITATIONS
SEE PROFILE

Jack Sullivan
University of Idaho
196 PUBLICATIONS **6,540** CITATIONS
SEE PROFILE

Smith Megan L (Orcid ID: 0000-0002-6362-9354)
Sullivan Jack (Orcid ID: 0000-0003-0216-6867)
Carstens Bryan Charles (Orcid ID: 0000-0002-1552-227X)

Article subject: Phylogeography

Title: The role of multiple Pleistocene refugia in promoting diversification in the Pacific Northwest.

Running title: Refugia and diversification in the Cascades

Authors: Megan L. Smith[1], Jessica Wallace[1], David C. Tank[2,3,4], Jack Sullivan[3,4], and Bryan C. Carstens[1]

[1]Department of Evolution, Ecology & Organismal Biology, The Ohio State University, 318 W. 12th Avenue, 300 Aronoff Labs, Columbus, OH 43210-1293, USA. [2]Department of Botany and Rocky Mountain Herbarium, University of Wyoming, 1000 E. University Ave., Laramie, WY 82071, USA, [3]Department of Biological Sciences, University of Idaho, 875 Perimeter Dr. MS 3051, Moscow, ID 83844-3051, USA. [4]Institute for Bioinformatics and Evolutionary Studies (IBEST), Biological Sciences, University of Idaho, 875 Perimeter Dr. MS 3051, Moscow, ID 83844-3051, USA.

Corresponding Author: Megan L. Smith, Department of Evolution, Ecology & Organismal Biology, The Ohio State University, 318 W. 12th Avenue, 300 Aronoff Labs, Columbus, OH 43210-1293, USA.

Correspondence: megansmth67@gmail.com

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.16595

Abstract

Pleistocene glacial cycles drastically changed the distributions of taxa endemic to temperate rainforests in the Pacific Northwest, with many experiencing reduced habitat suitability during glacial periods. In this study, we investigate whether glacial cycles promoted intraspecific divergence and whether subsequent range changes led to secondary contact and gene flow. For seven invertebrate species endemic to the PNW, we estimated Species Distribution Models (SDMs) and projected them onto current and historical climate conditions to assess how habitat suitability changed during glacial cycles. Using single nucleotide polymorphism (SNP) data from these species, we assessed population genetic structure and used a machine-learning approach to compare models with and without gene flow between populations upon secondary contact after the Last Glacial Maximum (LGM). Finally, we estimated divergence times and rates of gene flow between populations. SDMs suggest that there was less suitable habitat in the North Cascades and Northern Rocky Mountains during glacial compared to interglacial periods, resulting in reduced habitat suitability and habitat fragmentation during the LGM. Our genomic data identify population structure in all taxa, and support gene flow upon secondary contact in five of the seven taxa. Parameter estimates suggest that population divergences date to the later Pleistocene for most populations. Our results support a role of refugial dynamics in driving intraspecific divergence in the Cascades Range. In these invertebrates, population structure often does not correspond to current biogeographic or environmental barriers. Rather, population structure may reflect refugial lineages that have since expanded their ranges, often leading to secondary contact between once isolated lineages.

Keywords: Pacific Northwest, Pleistocene, phylogeography, refugia, secondary contact

Introduction

Throughout the Quaternary, the global climate has been characterized by dramatic fluctuations, with ice sheets advancing and receding in roughly 41-kyr and 100-kyr cycles. These drastic fluctuations have resulted in rapidly changing species distributions, with ranges likely undergoing repeated bouts of contraction and expansion to track suitable habitat (Hewitt 1996, 2000). Many temperate species likely would have been eliminated from northern parts of their ranges that were covered in ice during glaciation, with populations surviving south of glaciation in pockets of suitable habitat (refugia). During interglacial periods, these species likely expanded their ranges into previously uninhabitable regions, with subsequent glacial cycles leading to further contractions and expansions.

The effects of glacial cycles on genetic diversity have received substantial attention since the work of Hewitt (1996, 1999, 2000) and have become a central focus in phylogeographic investigations (e.g., Carstens & Richards, 2007; Graham et al., 2020; Myers et al., 2020; Stone et al., 2017; Wallis et al., 2016). Studies of temperate fauna in North America and Europe have highlighted several consistent patterns. First, temperate species often have reduced genetic diversity at northern latitudes, suggesting that these regions have been colonized relatively recently (Hewitt 2000; Kuchta & Tan, 2005; Soltis et al., 1997; Taberlet et al., 1998). Second, many species apparently survived in multiple, isolated glacial refugia (Shafer et al. 2010; Stone et al., 2017; Taberlet et al., 1998; Wallis et al., 2016). Finally, postglacial expansion has often led to secondary contact between refugial lineages (Hewitt 1996; Remington 1968; Taberlet et al. 1998).

Mesic, temperate rainforests characterize large parts of the Pacific Northwest of North America. These forests are characterized by the late-successional dominants western redcedar

(*Thuja plicata*) and western hemlock (*Tsuga heterophylla)* and are home to more than 150 temperate rainforest endemics (Nielson et al., 2001). Two disjunct segments of rainforest occupy the Northern Rocky Mountains and the coastal and Cascades ranges, and these two segments have been isolated since the orogeny of the Cascades range and subsequent xerification of the Columbia Basin 2-5 Myr (Figure 1). These rainforests were heavily affected by Quaternary glacial cycles, with the Laurentide ice sheet extending as far south as 40°N during the peak of glaciation (Figure 1). Temperate rainforest endemics may have been eliminated from northern portions of their ranges and survived south of glaciation or in isolated refugia.

Early phylogeographic work in the region tended to focus on whether refugial inland rainforests persisted throughout the Quaternary glacial cycles. Driven by pollen evidence suggesting the potential post-glacial re-establishment of inland rainforests, researchers aimed to evaluate whether there was evidence of persistence of inland refugia in several species. In several species of amphibians—for example, tailed frogs, *Ascaphus* (Nielson et al., 2001), *Plethodon* salamanders (Carstens et al., 2004), and *Dicamptodon* salamanders (Steele et al., 2005)—genetic data and model-based approaches support deep divergence between inland and coastal populations, suggesting that suitable habitat persisted in both inland and coastal forests throughout Quaternary glacial cycles. In other taxa, more complex histories are evident. For example, in the tree species *Alnus rubra* (Ruffley et al., 2018), *Thuja plicata*, and *Tsuga heterophylla* (Ruffley et al., 2022), genomic data suggest that both inland and coastal refugia persisted, but that there has been subsequent gene flow between these refugial populations. However, in many taxa that have been investigated to-date, including most invertebrates from the region, there is no support for deep divergence across the Columbia Basin (Carstens et al., 2005; Carstens et al., 2013; Espíndola et al., 2016; Smith et al., 2017, 2018; Smith & Carstens,

2020; Wilke & Duncan, 2004). Rather, inland rainforests appear to have been recolonized after the Pleistocene, suggesting the absence of a suitable refugium in inland rainforests during some glacial cycles. Within the coastal and Cascades ranges, refugia have been proposed in several locations in the South Cascades including the Klamath-Siskyou Mountains, the Olympic Peninsula, Queen Charlotte Islands, and Vancouver Island (Brunsfeld et al., 2001; Shafer et al. 2010), with several species apparently having occupied multiple refugia during the Quaternary (Brunsfeld et al., 2001; Shafer et al. 2010). Supporting the importance of multiple refugial populations, several plants exhibit N-S population structure (Soltis et al., 1997), and multiple populations have been found in the region in several invertebrates (Smith et al., 2018; Wilke & Duncan, 2004). The presence of multiple putative refugia in this region may have led to complex post-glacial colonization routes with the potential for secondary contact and admixture between once isolated refugial lineages.

Early investigations into phylogeographic patterns in the PNW relied on analyses of mitochondrial DNA and estimates of lineage divergence; such estimates are inaccurate and lack precision when made from single-locus data (Edwards & Beerli, 2000). Recent advances in data collection have increased the availability of SNP data within described species and species complexes (Garrick et al., 2015) and these data enable investigations into complex evolutionary scenarios, including isolation in Pleistocene refugia and expansion and gene flow between previously isolated lineages (e.g., Nevado et al., 2018). Recent studies that have investigated PNW endemics using genomic data have highlighted complex histories including multiple refugia and gene flow between refugial lineages (Ruffley et al., 2018; Ruffley et al., 2022; Smith and Carstens, 2020). Here, we revisit this classic question using SNP data from multiple

terrestrial invertebrates that inhabit temperate rainforests in the Pacific Northwest of North America.

In this study, we aim to test 1) whether populations survived in one or multiple refugia throughout Quaternary glacial cycles and where these refugia were located, 2) whether refugial lineages have subsequently come into secondary contact, and 3) if potential secondary contact has led to gene flow between these once isolated lineages. We focus on seven invertebrates endemic to the temperate rainforests of the PNW: five species of congeneric taildropper slugs (*Prophysaon* spp.), a carnivorous land snail (*Haplotrema vancouverense*), and a millipede (*Chonaphe armata*). We chose these taxa because previous work has highlighted the presence of multiple populations within the coastal and Cascades ranges. The six gastropods studied do not harbor deep divergence across the Columbia Basin (Smith et al., 2017, 2018; Smith & Carstens, 2020), but mitochondrial data suggest divergence within the coastal and Cascades ranges (Smith et al., 2018; Espíndola et al. 2016). Genomic data from *P. andersoni* suggest the presence of multiple refugia, and secondary contact and gene flow between refugial populations (Smith and Carstens, 2020). Previous work on the millipede *C. armata* led to equivocal support for traditional models of deep divergence or recent dispersal across the Columbia Basin, and suggested that, while multiple refugia were present, novel demographic models may be needed to better understand refugial structure and dynamics (Espíndola et al., 2016). Here, we use Species Distribution Models (SDMs) to explore the potential effects of glacial cycles on these species' ranges and to evaluate whether species experienced reduced habitat suitability and increased habitat fragmentation during glaciation, which we expect if glaciation led to divergence and species have subsequently expanded their ranges. We then use restriction-digest associated sequencing to collect genomic-scale data, and population clustering to identify

putative refugial populations. We use a machine-learning approach to compare models that vary in the number of refugia and in the presence or absence of gene flow between refugial populations. Finally, to better understand the timing of divergence and the magnitude of gene flow and to thus understand whether Pleistocene glacial cycles led to intraspecific divergence, we estimate divergence times and migration rates within each species.

Materials and Methods

Geographic data and Species Distribution Models

To evaluate how species ranges changed during glacial cycles we estimated current and historic SDMs. We downloaded locality data from published papers (Espíndola et al., 2016; Smith et al., 2017, 2018; Smith & Carstens, 2020) and from the Global Biodiversity Information Facility (GBIF) (Supporting Table S1; Harvard University, 2019; iNaturalist.org, 2018; Grant & Maier , 2018; European Molecular Biology Laboratory, 2019; Lopez, 2018; Roberts, 2018; Menard, 2018; Norton, 2018; Uribe & Villaronga, 2018; Paulay & Brown, 2019; Wheeler & McIntosh, 2018; Gagnon & Shorthouse, 2019; Academy of Natural Sciences, 2019) and combined them with localities from this study. We also downloaded range maps from the IUCN when possible. Using custom R scripts and the R packages rgdal v1.5-16 (Bivand et al. 2019), maps v3.3.0 (Brownrigg, 2018), mapdata v2.3.0 (Brownrigg, 2018), and rgeos v0.5-2 (Bivand & Rundel, 2019), we plotted localities and ranges and manually removed points that fell substantially outside of the Pacific Northwest, given that the focal taxa are endemic to this region and points substantially outside the region are likely erroneous. Given difficulties in identifying these taxa,

we suspect that removing these likely inaccurate occurrences should improve our SDMs. However, for *Prophysaon* ssp, we did not include any GBIF data, following Smith et al. (2017), as field identifications are often incorrect. For these species, we used only localities from this study, Smith et al. (2017), and Wilke & Duncan (2004). We thinned all occurrences to exclude points within 1km of other occurrences. After locality data were curated, we used the R package 'biomod2' v3.3-7 (Thuiller et al. 2019) to build SDMs. Since our area of interest is the rainforests of the PNW, but we are interested in hindcasting SDMs into the past, we chose a study area broader than our focal area by cropping the study area to a grid spanning -140 to -110 degrees Longitude and 40 to 60 degrees Latitude (to include areas that possibly supported rainforests in the past but that no longer do). We downloaded bioclimatic data from seven historic climate scenarios (late-Holocene, mid-Holocene, early-Holocene, Younger Dryas Stadial, Bølling-Allerød, Heinrich Stadial, LGM, and the Last Interglacial) at 2.5 minute resolution on September 11, 2020 from http://www.paleoclim.org/ (CHELSA, Brown et al. 2018). We downloaded the current data at the same resolution (CHELSA, Brown et al. 2018); while higher resolution layers are available for the current and LGM, the highest resolution available for other time frames used here is 2.5 minute resolution. We chose nine bioclimatic variables that were not highly correlated ($r < 0.7$) throughout the study area (mean diurnal range, isothermality, temperature seasonality, mean temperature of wettest quarter, mean temperature of driest quarter, mean temperature of warmest quarter, precipitation of wettest month, precipitation seasonality, precipitation of warmest quarter). When choosing between two correlated variables, we selected variables that we suspected to be biologically important for the focal taxa. For each species, we sampled pseudo-absences from a disk with a maximum distance of 10,000 meters from occurrence points. We sampled ten times as many pseudoabsenses as occurrence points

after thinning. We then used three modelling approaches: Maximum Entropy implemented in maxent, General Linear Models, and Random Forests. We used 80 percent of our data to build the models, and 20 percent for testing. We ran five replicates to evaluate model accuracy using cross-validation, and three permutations per variable to evaluate variable importance. 'biomod2' (Thuiller et al. 2019) assesses variable importance by shuffling a single variable, making predictions on the shuffled data set, and calculating a correlation between original and shuffled predictions. We used the Receiving Operator Characteristic (ROC) to evaluate models and rescaled all models to facilitate combining these models downstream. Then, we used an ensemble approach. We ignored models with a ROC less than 0.75, and, otherwise, combined models using a mean weighted approach based on ROC score. We forecast this ensemble model onto current and historic conditions and averaged the five replicates to generate a single prediction for each time frame.

We constructed ten binary maps for each species to quantify the change in habitat suitability from the present to the LGM. We used four thresholds (300, 500, 700, and 900 of maximum suitability 1000) to generate binary presence and absence maps both in the present and at the LGM. We also considered a threshold based on the 95% highest density interval (HDI) of occurrence points in the present. Specifically, we used the R package HDInterval v0.1,3 (Meredith & Kruschke, 2016) to find the 95% HDI of suitability scores for our occurrence points in the current SDM projection. We then used the lower interval as a minimum threshold to convert SDMs into binary presence/absence maps. For each threshold, we calculated the number of suitable raster cells in both time frames and calculated the ratio of suitable cells at the LGM to suitable cells at the present. To quantify the change in habitat fragmentation from the LGM to

the present, we used the same eight binary maps constructed above. We used the rasterToPolygons function in the R package 'raster' v3.0-12 (Hijmans, 2020) to transform each binary map into a set of polygons. Then, if there were at least two polygons at a time period and binary threshold, we used the functions knearneigh and knn2nb from the R package 'spdep' v1.1-5 (Bivand & Wong, 2018) to calculate the median nearest neighbor distance for each species, time period, and threshold. We then compared these distances between the current and LGM models to assess whether habitat fragmentation increased or decreased from the LGM to the present.

Extrapolating SDMs to unsampled parts of the current range, or to alternative climate scenarios could lead to inaccuracies when environments are not similar across sampled and unsampled temporal and spatial axes. To assess whether climate conditions were analogous across space and time, we used multivariate environmental similarity surface analyses (MESS; Elith et al. 2010). To construct MESS maps for each species, we used the R package dismo (Hijmans et al., 2021). We extracted environmental data for the present at each occurrence point, after conducting spatial thinning based on a 1km threshold. Then, we used the mess function to construct the MESS map for current and LGM climate conditions.

Genomic data: collection and processing

Reduced-representation sequence data were collected from five *Prophysaon* species using restriction-digest protocols to assess refugial structure and gene flow. We then added data from two additional species, a carnivorous land snail (*Haplotrema*; Smith et al. 2017) and a millipede

(Espindola et al. 2016). For the five *Prophysaon* taxa, a modified version of the GBS protocol (Elshire et al., 2011; Smith & Carstens, 2020) was used to generate a sequencing library. Details of data processing in *P. andersoni* are described in Smith and Carstens (2020). Details for other species are below. Data from all species were clustered independently during read clustering and alignment steps. Sequencing was conducted on an Illumina NextSeq at Novagene. Data processing in *Prophysaon* relied on the fastxtoolkit (Gordon & Hannon, 2010) to filter raw reads for adapters and to trim the ends. Due to quality issues in the 150 bp reads and the reverse reads, we trimmed all reads to 50 bp prior to analysis and discarded the reverse reads from the paired-end sequencing. Quality was assessed using FastQC (Andrews, 2010) to ensure that reads with adapter contamination were removed or trimmed. We used ipyrad v0.7.24 (Eaton & Overcast, 2016) to demultiplex, filter, and align the reads. We used statistical base calling, with a minimum depth of six reads per sample per locus and a maximum depth of 10,000 reads. We filtered reads that were shorter than 35 base pairs after trimming to remove low quality base pairs and barcodes. For clustering within and between samples, we used a threshold of 0.85. We allowed no more than five Ns and no more than eight heterozygous bases in the consensus sequence. Data processing in *H. vancouverense* and *C. armata* was conducted in ipyrad under similar settings, and is described fully in the Supporting Information. We required that loci were sequenced from 50 percent of individuals to be included in downstream analyses.

Population structure and summary statistics

To examine population structure, we used STRUCTURE v2.3.4 (Pritchard et al., 2000). For input, we used unlinked SNPs from ipyrad (.ustr files). For *P. andersoni,* we used the results

from Smith and Carstens (2020). For all other species, STRUCTURE analyses were conducted as follows. Analyses were run for K values from one to 10 with five replicates per K. The first 100,000 generations were discarded as burn-in, followed by 500,000 samples. We used Structure Harvester (Earl, 2012) to justify a K value based on likelihood scores and the Delta K method (Evanno et al., 2005).

For each population inferred in STRUCTURE, we calculated π in the R package 'PopGenome' (Pfeifer et al., 2014). We also calculated π for populations defined by geographic regions (North Cascades, South Cascades, Northern Rocky Mountains, and Blue and Wallowa Mountains). We calculated F_{ST} between all STRUCTURE populations.

Assessing support for secondary gene flow between refugial populations

To evaluate the number of populations that survived the LGM and assess whether secondary contact and gene flow occurred between refugial populations after the LGM, we used the R package 'delimitR' v2.0.0 (Smith and Carstens, 2020). 'delimitR' uses machine learning and coalescent simulations conducted by fastsimcoal2 (Excoffier et al. 2013) to compare demographic models using the Site Frequency Spectrum as a summary statistic. We used the K as determined from STRUCTURE results as the maximum number of refugial populations and assigned individuals to the population for which they had the highest coancestry coefficient. Then, we used custom Python scripts (available at github.com/meganlsmith/BuildmSFS) to construct a multidimensional Site Frequency Spectrum (mSFS), with population assignments based on Structure results. We required that all SNPs used in the construction of the mSFS be

biallelic and sampled in at least 50 percent of the individuals in each population. For SNPs sampled in more than 50 percent of individuals in a population, we randomly down-sampled alleles, following Satler and Carstens (2017). We sampled only a single SNP from each locus. This allowed us to build a SFS from a matrix without missing data. For the purpose of model selection, we did not consider the monomorphic cell of the SFS.

We considered all possible divergent topologies between populations for all possible combinations of populations. For populations for which there was evidence of admixture in STRUCTURE results, we allowed for the possibility of secondary contact and gene flow after the LGM. When gene flow upon secondary contact occurred, it was modelled as beginning sometime in the past and continuing until the present. We did not model the possibility of multiple pulses of secondary contact, both because this would result in a very large model space, and because such models can be very difficult to distinguish. In *P. vanattae,* where there was minimal evidence of admixture, we allowed for the potential for gene flow upon secondary contact between two populations that were both geographically proximate and had no biogeographic barriers in that part of their range. We also considered models that included expansion after the LGM. When expansion was included, it was modeled as occurring in all populations but the rate of expansion was allowed to vary among populations. We considered only a single migration event per model to limit the model space. While this approach is approximate, it nevertheless allows us to test the prediction that gene flow has occurred upon secondary contact in the focal taxa. Full information on the priors used in each analysis is available in the Supporting Information (Supporting Table S2). In general, we used broad priors for population sizes, divergence times, migration rates and expansion rates. For *P. andersoni,* we

used the models, simulations, and results from Smith and Carstens (2020). For all other species, we simulated 10000 SFS under each model in fastsimcoal2 (fsc2; Excoffier et al., 2013) using functions from the R package 'delimitR' and summarized the SFS by binning (four classes per population). For model selection in 'delimitR', we used 500 decision trees in the RF classifier. We recorded out-of-bag error rates (a measure of accuracy of RF classifiers based on internal cross-validation), the selected model, and the approximated posterior probability of the selected model for each species. Out-of-bag error rates are calculated by considering some element of the prior and classifying that element using only decision trees constructed without reference to the element. They measure the frequency of correct classification of elements in the prior and should provide an accurate assessment of power. To approximate the posterior probability, 'delimitR' implements the approach of Pudlo et al., (2016) and regresses against these out-of-bag error rates under the selected model.

Finally, we estimated current population sizes, divergence times, and migration rates and their confidence intervals under the best model in fsc2 (Excoffier et al., 2013) for each species. We used the Expectation-Conditional Maximization (ECM) algorithm with 100000 simulations to estimate the expected mSFS and 40 ECM cycles to estimate parameters. For this analysis, we considered linked SNPs and the monomorphic cell of the SFS to improve parameter estimation. To consider the monomorphic cell in parameter estimation, it was necessary to include mutation rates in the models. We used broad priors for mutation rates in all species (U(1e-9,1e-7)). To construct mSFS using linked SNPs and the monomorphic cell, we used a custom python script with the same down-sampling method as described when constructing the mSFS with only linked SNPs (available at https://github.com/meganlsmith/BuildmSFS). We used 100 non-

parametric bootstrap replicates to build 95% density intervals around parameter estimates by repeating the downsampling procedure 100 times and performing parameter estimation on each downsampled replicate. We bounded all priors in this analysis. We used the R package 'HDInteval' v0.2.0 (Meredith and Kruschke 2016) to construct 95% CI for each parameter, and we calculated the median value for each parameter. To convert divergence times into years, we used a generation time of one year, as suspected for *Prophysaon* (COSEWIC 2006), and common in many other terrestrial gastropods.

Results

Geographic data and Species Distribution Models

Generally, SDM performance was moderate, with average AUC scores ranging from 0.640 (*P. foliolatum*) to 0.807 (*H. vancouverense*). For all species except *P. foliolatum*, all three algorithms considered were included in the ensemble model (Table 1). For *P. foliolatum*, only a single algorithm had a high enough score to be included in the ensemble model (the GLM algorithm). Average ROC scores across replicates and modelling strategies included in the ensemble model for each species are reported in Table 1, and average ROC scores for each modelling strategy and each species are reported in Supporting Table S3. While these values are relatively low compared to some studies, we note that we employed spatial thinning, and sampled a limited number of pseudoabsenses in an attempt to get more accurate assessments of model performance. While this may lead to lower ROC scores, we argue that these scores are a more accurate representation of the amount of confidence we should have in the SDMs. For all

species, suitable habitat was predicted in the Cascades and Coastal ranges, and for some taxa, suitable habitat was predicted inland in the Northern Rocky Mountains. For all taxa known to have present distributions in the Northern Rocky Mountains except *P. coeruleum*, suitable habitat was clearly predicted in the inland rainforests.

In general, suitable habitat was shifted south and towards the coast during the colder periods of the Last Glacial Maximum and the Heinrich Stadial and shifted north and inland towards the present as climates warmed (Figure 2, Supporting Figures S1-S7). These results highlight the dynamic nature of species ranges. Although it becomes more difficult to predict climate further into the past, it is likely the case that species' ranges repeatedly contracted and expanded in response to glacial cycles. SDMs across several time periods are shown in the Supporting Information (Supporting Figures S1-S7). Since the most striking differences are generally between the LGM and the present, and since climatic oscillations became increasingly dramatic towards the present (Hewitt 2000), we focus on this time period in the main text and attempt to quantify the changes between the LGM and the present. While species undoubtedly went through numerous cycles of range contraction and expansion throughout the Quaternary, the changes since the LGM should be among the most severe and should offer an informative picture of the potential effects of these cycles on the distributions of the focal taxa.

All species except *P. foliolatum* have more suitable habitat available in the present than during the LGM (Table 1). Under the HDI threshold, the ratio of suitable habitat in the current to the LGM ranged from 0.023 in *P. vanattae* to 2.99 in *P. foliolatum*. Notably, our confidence in the model for *P. foliolatum* is generally low, as only a single algorithm could be included in the

ensemble model, the current and predicted past ranges are rather different, and only a small number of occurrence points were available after thinning (Table 1). Evidence for habitat fragmentation during the LGM was mixed and detecting the significance of differences is not straight-forward. However, for four of the seven species studied, all thresholds showed higher habitat fragmentation during the LGM than during the present (Table 1; Supporting Table S4). Across all thresholds, *P. andersoni* and *P. foliolatum* showed higher current fragmentation than inferred for the LGM, the opposite of the expected pattern. Again, our confidence in results from *P. foliolatum* is limited. *Chonaphe armata* showed different trends depending on the threshold used, suggesting that results for this species are equivocal.

MESS maps indicated that extrapolation was more prevalent in some species than in others (Supporting Figure S8). For the most heavily-sampled taxa (*P. andersoni* and *H. vancouverense*), MESS maps indicated relatively high similarity across current and LGM time periods, and the most negative scores (indicating dissimilarity) were north of the extent of the glaciers in non-coastal portions of the species range during the LGM. Since this region was almost certainly uninhabitable for these species when glaciers were present, and predicted as unsuitable by the SDMs, we do not suspect that dissimilarity had a meaningful effect on these models. On the other hand, for *P. foliolatum* (the species with the lowest ROC scores), MESS maps indicated substantial dissimilarity outside the species current range in the present, and throughout much of the species' range in the LGM. Other species tended to fall in between these two extremes, with areas north of the glacial extent in the LGM being the least analogous.

Next-Generation sequence data: collection and processing

For the five *Prophysaon* species sequenced using GBS, we obtained between 6,855 (*P. vanattae*) and 27,339 (*P. foliolatum*) loci (average = 17,131 loci) with the requirement that loci be present in >50% of all individuals (Table 2). For *H. vancouverense*, 3,746 loci were sequenced in at least half of the samples (Table 2). For *C. armata*, 135 loci were sequenced in at least half of the samples (Table 2).

Estimating the number of populations

For *P. andersoni*, we followed Smith and Carstens (2020) in using four as the maximum number of populations in downstream analyses, because in STRUCTURE analyses, the likelihood steadily increased up until K reached four. For *P. coeruleum, P. dubium, P. foliolatum,* and *C. armata,* there was a peak in ΔK when K reached two, so we used two as the maximum number of populations downstream. For *P. vanattae*, likelihood scores and ΔK strongly suggested that K = three. For *H. vancouverense*, there was a clear peak in ΔK when K reached four (Supporting Figures 9 and 10).

For *P. andersoni,* three populations were unique to the South Cascades (SC) and the fourth extended through the North Cascades (NC) and the Northern Rocky Mountains (NRM); admixture was limited to the SC (Figure 3a; Smith and Carstens, 2019). For *P. coeruleum,* all individuals sampled from the NRM belong to a single population that also includes individuals from south of Eugene, OR. Populations from Mt. Hood comprise a second population, and there are admixed individuals in SE Oregon and Eugene (Figure 3b; Supporting Table S5). For *P.*

dubium, both populations occur in the SC, and one spans the SC, NC, and the NRM; there is limited evidence of admixture (Figure 3c; Supporting Table S6). For *P. foliolatum*, one population is exclusively found in and around Olympia, one occurs near the Columbia River, and admixed individuals are found in Olympia (Figure 3d; Supporting Table S7). For *P. vanattae*, one population is found near the Columbia River, in the NC, and in the Blue and Wallowa Mountains. A second occurs in the NC, and a third occurs in the SC; there is limited evidence of admixture (Figure 3e; Supporting Table S8). For *C. armata*, the two populations are widespread with some geographic overlap and evidence of admixture in the NRM (Figure 3f; Supporting Table S9). For *H. vancouverense,* two populations occur in the NRM. One population occurs in the SC, NC, and Haida Gwaii, and one population occurs in the NC and on Vancouver Island. There is evidence of admixture in all populations except the northern population in the NRM (Figure 3g; Supporting Table S10).

For four of the seven species (*P. coeruleum, P. dubium, P. foliolatum, P. andersoni*), genetic diversity (π) was higher in the SC than in any other sampled region (Supporting Table S11). Notably, there were only two samples available from the SC for *C. armata*, and genetic diversity was 0 in this region. For *H. vancouverense* and *P. vanattae*, genetic diversity was highest in the North Cascades. Nucleotide diversity values for the STRUCTURE populations are reported in Supporting Table S12). F_{ST} values in species with only two populations ranged from 0.067 in *C. armata* to 0.791 in *P. dubium*. All F_{ST} values were high for *P. vanattae* (0.804-0.921), and values ranged widely in *P. andersoni* and *H. vancouverense* (Supporting Table S13).

Assessing support for secondary contact between refugial populations

For five of the seven taxa, we found support for gene flow upon secondary contact, whereas for *P. dubium* and *P. vanattae*, the best model was a divergence-only model (Figure 4; Table 3). For all species, the best model included the maximum number of populations evaluated, as suggested by STRUCTURE results. Generally, for species with two populations, error rates were lowest, with the exception of *C. armata*, for which we had very limited data (Table 3; error rate = 0.23). For all species of *Prophysaon* with only two populations (*P. coeruleum, P. dubium,* and *P. foliolatum*), error rates ranged from 0.02 to 0.05.

For *P. andersoni*, errors were most often between models that differed only in the presence or absence of expansion. The best model was a four-population model with secondary contact (pp=0.689), and all models receiving more than 20 votes included secondary contact. Results are described in more detail in Smith and Carstens (2020).

For *P. coeruleum*, error rates were greater than ten percent only for the two-population model with secondary contact and expansion (error rate = 0.11; Supporting Table S14), which was most often mistaken for the two-population model with secondary contact and no expansion or the two-population divergence-only model. The best model was the two-population model with secondary contact and no expansion (pp=0.686), a model with a low error rate (0.06), which was most often mistaken for the same model with expansion in simulations; the only other models receiving votes were the same model with expansion and the two-population divergence-only model (97 and 76 of the 500 votes, respectively; Supporting Table S15).

For *P. dubium*, error rates were highest for the two-population model with secondary contact and expansion (error rate = 0.13; Supporting Table S16), which was most often mistaken for the same model without secondary contact. The best model was the two-population divergence-only model (pp=0.784; Supporting Table S17), and the same model with expansion received the second highest number of votes.

For *P. foliolatum*, error rates were below ten percent across all models (Supporting Table S18). The best model was the secondary-contact model without expansion (pp=0.528), and the model receiving the second highest number of votes was the divergence-only model (167/500 votes; Supporting Table S19).

For *P. vanattae*, error rates were above fifteen percent for two models: the single-population model and the single-population model with expansion, which were most often mistaken for each other (Supporting Table S20). Four models had error rates above ten percent but below fifteen percent (Supporting Table S20). The best model was the three-population divergence-only model (pp=0.706). Three other models received ~ 100 votes: two three-population divergence-only models that differed in topology and one three population model with secondary contact between population 1 and population 2 (Supporting Table S21). Given the similar number of model votes amongst these three models, there is substantial uncertainty regarding the best model for *P. vanattae*.

For *C. armata*, error rates were above ten percent for four models: the divergence-only model, the secondary-contact model, the divergence-only model with expansion, and the secondary-

contact model with expansion (Supporting Table S22). The best model was the secondary-contact model with expansion (pp=0.675; Supporting Table S23), but given the high error rates and limited number of loci in this species, our confidence in this result is low.

Error rates were highest for *H. vancouverense* (0.56) but were highest between models that differed only in the presence or absence of expansion (Supporting Table S24). Five models received ten or more votes. All nine of these models included four populations and secondary contact (Supporting Table S25). The best model included secondary contact between Population 1 and Population 2 and population expansion (pp = 0.436), and the model receiving the second highest number of votes differed from the best model only in the absence of population expansion.

Estimates and confidence intervals for all parameters are reported in Supporting Tables S26 and S27. Population divergences were all estimated to have occurred during the Pleistocene, with a maximum median estimated divergence time of 421,672 years in *P. vanattae,* and a minimum median estimated divergence time 39,405 years in *P. foliolatum* (Figure 5). Parameter estimates suggest that species that began to diverge earlier may have lower migration rates in the present, but this is not a strong inference. *P. coeruleum* and *P. vanattae* had the most ancient divergence times (> 100,000 years; Table S22). *P. dubium* and *C. armata* had moderate divergence time estimates > 50,000 years ago, and other divergence time estimates were < 50,000 years ago. Since the best models for *P. dubium* and *P. vanattae* did not include migration, migration estimates were zero. Apart from these two species, *P. coeruleum* had the lowest estimated migration rate. All species with recent divergence times (<50,000 generations ago), had non-zero

migration rates. Considering 95% HDI on parameters, divergence time estimates ranged from 22,910 to 1,657,738 years ago. The maximum bounds exceeded 500,000 years before the present for *P. vanattae*, *P. coeruleum*, and one relationship within *H. vancouverense*.

Discussion

Pleistocene glacial cycles shape intraspecific diversity in the Pacific Northwest

Our results support the hypothesis that Pleistocene glacial cycles were a key factor in driving divergence within invertebrates endemic to the temperate rainforests of the PNW. As expected, genetic diversity is generally higher in the southern and coastal portions of species' ranges, and there is generally evidence of multiple populations in these portions of each species' range (Figure 3), suggesting that these areas are likely locations of refugia. For all taxa, we find evidence of population divergence during the Pleistocene using model-based approaches. For all but two species (*P. vanattae* and *P. coeruleum*) and one divergence in *H. vancouverense*, divergence-time estimates and confidence intervals are less than 500 kya (Figure 5, Table S22), suggesting that most divergences occurred in the later Pleistocene. Further, the lineages uncovered in this investigation were typically parapatric or sympatric in their geographic distributions (Figure 3), suggesting that population genetic structure does not result from biogeographic barriers but rather from historic isolation in multiple refugia.

This interpretation is consistent with the results of the SDMs, which show reduced habitat suitability in all species except *P. foliolatum* and increased habitat fragmentation for four species

during the LGM, suggesting a potential mechanism for Pleistocene divergence in some taxa. Results for SDMs, particularly for *P. foliolatum,* should be interpreted with caution. For several species (*P. coeruleum, P. foliolatum, and C. armata*) at least one modelling algorithm had an average ROC <0.7, suggesting limited discriminatory power (Table 1; Supporting Table S3). Particularly, for *P. foliolatum*, only the GLM algorithm had an average ROC > 0.7. Despite this, our SDMs show a pattern of habitat contraction during the LGM for all but *P. foliolatum*. While divergences pre-date the LGM, the results of the SDMs likely point to a general trend of habitat contractions during the height of glacial cycles, followed by habitat expansion, a process that may have driven divergence in the focal taxa. When extrapolating SDMs to novel spatial or temporal extents, it is important to consider whether analogous climates exist across time and space (Elith et al., 2010). We used MESS maps to assess this. These results highlight that our confidence should be higher for some species (*P. andersoni, H. vancouverense*), where climates appear to be largely analogous, than for others, like *P. foliolatum*. In general, the least analogous climates are north of the glacial extent during the LGM. Since this region was almost certainly uninhabitable during the LGM, and was predicted as such by the SDMs, it is unlikely that this extrapolation substantially impacted results. However, dissimilar climates were more widespread for some taxa, and warrant caution when interpreting these results. This lack of analogous climates could contribute to the apparent reduced habitat suitability and increased habitat fragmentation during the LGM.

Our results suggest that several taxa were isolated in multiple refugia followed by post-glacial expansion and secondary contact. Though previous work has supported secondary contact between inland and coastal populations of wind-pollinated trees (*Alnus rubra, Thuja plicata,*

Tsuga heterophylla) with high dispersal ability (Ruffley et al., 2018; Ruffley et al., 2022), most published investigations into separate inland and coastal refugial lineages in the PNW have been limited in their ability to understand the effects of dynamic ranges and secondary contact because of limited dispersal across the Columbia Basin. By focusing on Cascades and Coastal distributions, we uncovered a complex history of isolation and secondary contact as well as population structure that does not correspond to current biogeographic or environmental barriers. Specifically, we found parapatric lineages in several taxa (Figure 3), and our model-selection results support gene flow upon secondary contact between previously isolated lineages in five of the seven taxa studied (Figure 4, Table 3). Gene flow in these lineages results from dynamic range shifts that put previously isolated lineages into secondary contact. Notably, for the species with the deepest divergences, *P. vanattae*, there was no support for secondary contact, suggesting, perhaps, that more time in isolation and increased divergence prevented gene flow upon secondary contact in this species.

Species limits in terrestrial invertebrates from the Pacific Northwest

Our results highlight the potential for undescribed diversity within these species. Particularly for *P. vanattae,* where we find evidence of deep divergence and no current gene flow, future work should investigate potential morphological and ecological differences in this taxon and should evaluate whether additional species should be described. For other species, evaluating species status for recently diverged lineages with evidence of gene flow requires additional ecological and morphological data. Previous work has identified undescribed lineages in *P. andersoni* (Smith & Carstens, 2020) and highlighted ecological differences between lineages within this

species. Future work in *P. andersoni* and the other species we investigated should evaluate ecological and morphological divergence because these data may provide further insight into whether the lineages found here represent distinct species or whether gene flow will eventually lead to population fusion. Given the myriad potential outcomes of gene flow upon secondary contact, and the difficulties of evaluating complex model spaces (see below), care should be taken when making taxonomic decisions based on genomic data alone, particularly when there is support for gene flow between lineages.

High-throughput sequencing, machine learning, and complex model spaces

Exploring complex model space is difficult, even while facilitated by the increasing availability of genomic data. Our results demonstrate adequate statistical power for differentiating among a relatively small set of models (~6 models) using a moderate amount of genetic data by current standards (>=1,000 SNPs). However, with fewer SNPs, as was the case for *C. armata*, error rates increase dramatically, and with larger model spaces, for example for *P. vanattae* (24 models) or *H. vancouverense* (348 models), our power to distinguish among similar models is diminished. This suggests a limit to the evolutionary processes amongst which we can distinguish, despite the fact that a very small subset of these processes was explored here. As more and more genomic data become available, the importance of understanding the limits of the processes we can infer and of incorporating the impacts of other factors structuring genomic variation (e.g., linked selection; Cruickshank & Hahn, 2014; Ewing & Jensen, 2016), variable migration rates across the genome (Roux et al., 2016), and sequencing error (Shafer et al., 2017)) will become increasingly important. In the case of the effects of complex glacial cycles, such as those

explored here, true models are likely far more complex than those included in the model set. It seems likely that multiple bouts of range contractions and expansions structured intraspecific diversity in these taxa. However, evaluating such complex models would not only lead to an intractably large model space, but also would likely lead to issues in model identifiability, as later events likely obscure the effects of previous ones on genetic diversity (Harpending et al. 1998). This could mislead inference; for example, several periods of gene flow may be a better explanation than simple secondary contact after the LGM, but the models explored here would not allow us to distinguish amongst these scenarios. Future work should aim to better understand these limits and how our inferences are likely to be misled when certain processes are ignored, an inevitability in model-based approaches.

Conclusions

Here, we investigate the effects Pleistocene glacial cycles on structuring intraspecific diversity in several terrestrial invertebrates from the Pacific Northwest. We find support for a complex history with lineages isolated in multiple refugia in the Cascades and Coastal ranges. In several species, our results support gene flow between refugial lineages upon secondary contact subsequent to glacial cycles, highlighting the complex effects of glacial cycles on intraspecific diversification. More broadly, the growing availability of genomic data (Garrick et al. 2015), sequence databases (e.g., Deck et al. 2017), and powerful analytical techniques like the machine-learning approach used here suggest that phylogeography is entering a period of renaissance in which complex model sets can be considered to better understand the factors driving diversification.

Acknowledgements: MLS was funded by an NSF GRFP (DGE-1343012) and NSF PRFB (DBI-2009989), and this work was funded by NSF (DEB1457519 to BCC, DEB1457726 to JS & DCT, and DBI-1560116 to BCC) and through grants to MLS through the Society for the Study of Evolution and the Society of Systematic Biologists. We thank Michael Lucid and Joel Sauder of the Idaho Department of Fish and Game for providing samples. We thank the Idaho Department of Fish and Game, the Washington Department of Fish and Wildlife and the California Department of Fish and Wildlife for providing permits. We thank Katherine Field, Andrew Rankin, Megan Ruffley, Ben Stone, Tara Pelletier, and Anahí Espíndola for assistance with fieldwork for this study.

References

Academy of Natural Sciences. MAL. Occurrence dataset https://doi.org/10.15468/xp1dhx accessed via GBIF.org on 2019-01-12.

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

Batchelor, C. L., Dowdeswell, J. A., & Pietras, J. T. (2014). Evidence for multiple Quaternary ice advances and fan development from the Amundsen Gulf cross-shelf trough and slope, Canadian Beaufort Sea margin. *Marine and Petroleum Geology*, *52*, 125–143. https://doi.org/10.1016/j.marpetgeo.2013.11.005

Bivand, R. S. & Wong, D. W. S. (2018). Comparing implementations of global and local indicators of spatial association. *TEST, 27*(3), 716-748. https://doi.org/10.1007/s11749-018-0599-x

Bivand, R., Keitt, T., & Rowlingson, B. (2019). rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.4-8. https://CRAN.R-project.org/package=rgdal.

Bivand, R. & Rundel, C. (2019). rgeos: Interface to Geometry Engine - Open Source ('GEOS'). R package version 0.5-2. https://CRAN.R-project.org/package=rgeos

Brown, J. L., Hill, D. J., Dolan, A. M., Carnaval, A. C., & Haywood, A. M. (2018). PaleoClim, high spatial resolution paleoclimate surfaces for global land areas. *Nature – Scientific Data*. *5*, 180254. https://doi.org/10.1038/sdata.2018.254

Brownrigg, R. (2018). Original S code by Becker, R.A., Wilks, A.R. R version by Brownrigg, R. Enhancements by Minka, T.P. & Deckmyn, A. maps: Draw Geographical Maps. R package version 3.3.0. https://CRAN.R-projet.org/package=maps.

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

Carstens, B. C., Stevenson, A. L., Degenhardt, J. D., & Sullivan, J. (2004). Testing nested phylogenetic and phylogeographic hypotheses in the *Plethodon vandykei* species group. *Systematic Biology, 53*(5), 781-792. https://doi.org/10.1080/10635150490522296

Carstens, B. C., Brunsfeld, S. J., Demboski, J. R., Good, J. M., & Sullivan, J. (2005). Investigating the evolutionary history of the Pacific Northwest mesic forest ecosystem: hypothesis testing within a comparative phylogeographic framework. *Evolution*, *59*(8), 1639–1652. https://doi.org/10.1111/j.0014-3820.2005.tb01815.x

Carstens, B. C., & Richards, C. L. (2007). Integrating coalescent and ecological niche modeling in comparative phylogeography. *Evolution, 61*(6), 1439-1454. https://doi.org/10.1111/j.1558-5646.2007.00117.x

Carstens, B. C., Brennan, R. S., Chua, V., Duffie, C. V., Harvey, M. G., Koch, R. A., McMahan, C. D., Nelson, B. J., Newman, C. E., Satler, J. D., Seeholzer, G., Posbic, K., Tank, D. C., & Sullivan, J. (2013). Model selection as a tool for phylogeographic inference: an example from the willow *Salix melanopsis*. *Molecular Ecology, 22*(15), 4014-4028. https://doi.org/10.1111/mec.12347

COSEWIC. (2006). COSEWIC assessment and status report on the Blue-grey Taildropper slug *Prophysaon coeruleum* in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa. vi + 25 pp. http://www.registrelep-sararegistry.gc.ca/default_e.cfm.

Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, *23*(13), 3133–3157. https://doi.org/10.1111/mec.12796

Deck, J., Gaither, M. R., Ewing, R., Bird, C. E., Davies, N., Meyer, C., Riginos, C., Toonen, R. J., & Crandall, E. D. (2017). The Genomic Observatories Metadatabase (GeOMe): A new repository for field and sampling event metadata associated with genetic samples. *PLoS Biology, 15*(8), e2002925. https://doi.org/10.1371/journal.pbio.2002925

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Eaton, D. A., & Overcast, I. (2016). ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics, 36*(8), 2592-2594. https://doi.org/10.1093/bioinformatics/btz966

Edwards, S., & Beerli, P. (2000). Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution*, *54*(6), 1839-1854. https://doi.org/10.1111/j.0014-3820.2000.tb01231.x

Elith, J., Kearney, M., & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution, 1*(4), 330-342. https://doi.org/10.1111/j.2041-210X.2010.00036.x

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS One*, *6*(5), e19379. https://doi.org/10.1371/journal.pone.0019379

Espíndola, A., Ruffley, M., Smith, M. L., Carstens, B. C., Tank, D. C., & Sullivan, J. (2016). Identifying cryptic diversity with predictive phylogeography. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1841), 20161529. https://doi.org/10.1098/rspb.2016.1529

European Molecular Biology Laboratory (EMBL). (2019). Geographically tagged INSDC sequences. Occurrence dataset https://doi.org/10.15468/cndomv accessed via GBIF.org on 2019-01-12.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology 14*(8), 2611-2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Ewing, G. B., & Jensen, J. D. (2016). The consequences of not accounting for background selection in demographic inference. *Molecular Ecology*, *25*(1), 135–141. https://doi.org/10.1111/mec.13390

Excoffier, L., Dupanloup, I., Huerta-Sánchez, E., Sousa, V. C., & Foll, M. (2013). Robust demographic inference from genomic and SNP data. *PLoS Genetics*, *9*(10). https://doi.org/10.1371/journal.pgen.1003905

Gagnon J, Shorthouse D (2019). Canadian Museum of Nature Mollusc Collection. Version 1.16. Canadian Museum of Nature. Occurrence dataset https://doi.org/10.15468/esfi97 accessed via GBIF.org on 2019-01-12.

Garrick, R. C., Bonatelli, I. A., Hyseni, C., Morales, A., Pelletier, T. A., Perez, M. F., Rice, E., Satler, J. D., Symula, R. E., & Thomé, M. T. C. (2015). The evolution of phylogeographic data sets. *Molecular Ecology*, *24*(6), 1164–1171. https://doi.org/10.1111/mec.13108

Gordon, A. and Hannon, G.J. (2010). Fastx-toolkit. FASTQ/A short-reads preprocessing tools (unpublished). Available online at: http://hannonlab.cshl.edu/fastx_toolkit

Graham, M., Santibáñez-López, C. E., Derkarabetian, S., & Hendrixon, B. E. (2020). Pleistocene persistence and expansion in tarantulas on the Colorado Plateau and the effects of missing data on phylogeographical inferences from RADseq. *Molecular Ecology*, *29*(19), 3684-3701. https://doi.org/10.1111/mec.15588

Grant, S., & Maier, C. (2018). Field Museum of Natural History (Zoology) Insect, Arachnid and Myriapod Collection. Version 12.9. Field Museum. Occurrence dataset https://doi.org/10.15468/0ywfpc accessed via GBIF.org on 2019-01-12.

Harpending, H. C., Batzer, M. A., Gurven, M., Jorde, L. B., Rogers, A. R., & Sherry, S. T. (1998). Genetic traces of ancient demography. *Proceedings of the National Academy of Sciences, 95*, 1961-1967. https://doi.org/10.1073/pnas.95.4.1961

Harvard University M, Morris P J (2019). Museum of Comparative Zoology, Harvard University. Version 162.135. Museum of Comparative Zoology, Harvard University. Occurrence dataset https://doi.org/10.15468/p5rupv accessed via GBIF.org on 2019-01-12.

Hewitt, G. M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linnean Society, 58*(3), 247-276. https://doi.org/10.1111/j.1095-8312.1996.tb01434.x

Hewitt, G. M. (1999). Post-glacial recolonization of European biota. *Biological Journal of the Linnean Society, 68*(1-2), 87-112. https://doi.org/10.1111/j.1095-8312.1999.tb01160.x

Hewitt, G. M. (2000) The genetic legacy of the Quaternary ice ages. *Nature*, *405*, 907-913. https://doi.org/10.1038/35016000

Hijmans, R. J. (2020). raster: Geographic Data Analysis and Modeling. R package version 3.0-12. https://CRAN.R-project.org/package=raster

Hijmans, R. J., Phillips, S., Leathwick, J., & Elith, J. (2021) dismo: Species Distribution Modeling. R package version 1.3-5. https://CRAN.R-project.org/package=dismo

iNaturalist.org (2018). iNaturalist Research-grade Observations. Occurrence dataset https://doi.org/10.15468/ab3s5x accessed via GBIF.org on 2019-01-12.

Kuchta, S. R., & Tan, A. M. (2005). Isolation by distance and post-glacial range expansion in the rough-skinned newt, *Taricha granulosa*. *Molecular Ecology, 14*(1), 225-244. https://doi.org/10.1111/j.1365-294X.2004.02388.x

Lopez, A. (2018). UAM Invertebrate Collection (Arctos). Version 31.22. University of Alaska Museum of the North. Occurrence dataset https://doi.org/10.15468/wrvy1y accessed via GBIF.org on 2019-01-12.

Meredith, M., & Kruschke, J. (2016). HDInterval: highest (posterior) density intervals. *R package version 0.1*, *3*.

Menard, K., King, P. (2018). Recent Invertebrates Specimens. Version 101.62. Sam Noble Oklahoma Museum of Natural History. Occurrence dataset https://doi.org/10.15468/glxcep accessed via GBIF.org on 2019-01-12.

Myers, E. A., McKelvy, A. D., & Burbrink, F. T. (2020). Biogeographic barriers, Pleistocene refugia, and climatic gradients in the southeastern Nearctic drive diversification in cornsnakes *(Pantherophis guttatus* complex). *Molecular Ecology, 29*(4), 797-811. https://doi.org/10.1111/mec.15358

Norton, B. (2018). NCSM Mollusks Collection. Version 1.2. North Carolina State Museum of Natural Sciences. Occurrence dataset https://doi.org/10.15468/unormg accessed via GBIF.org on 2019-01-12.

Nevado, B., Contreras-Ortiz, N., Hughes, C., & Filatov, D. A. (2018). Pleistocene glacial cycles drive isolation, gene flow and speciation in the high-elevation Andes. *New Phytologist*, *219*(2), 779–793. https://doi.org/10.1111/nph.15243

Nielson, M., Lohman, K., & Sullivan, J. (2001). Phylogeography of the tailed frog (*Ascaphus truei*): implications for the biogeography of the Pacific Northwest. *Evolution*, *55*(1), 147–160. https://doi.org/10.1111/j.0014-3820.2001.tb01280.x

Paulay, G., Brown, W. (2019). UF Invertebrate Zoology. Florida Museum of Natural History. Occurrence dataset https://doi.org/10.15468/sm6qo6 accessed via GBIF.org on 2019-01-12.

Pfeifer, B., et al. (2014). PopGenome: An efficient swiss army knife for population genomic analyses in R. *Molecular Biology and Evolution, 31*(7), 1929-1936. https://doi.org/10.1093/molbev/msu136

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

Pudlo, P., Marin, J. M., Estoup, A., Cornuet, J. M., Gautier, M., & Robert, C., P. (2016). Reliable model choice via random forests. *Bioinformatics, 32*(6), 859-866. https://doi.org/10.1093/bioinformatics/btv684

Rand, A. (1948). Glaciation, an isolating factor in speciation. *Evolution*, 314–321. https://doi.org/10.2307/2405522

Remington, C. L. (1968). Suture-zones of hybrid interaction between recently joined biotas. In *Evolutionary biology* (pp. 321-428). Springer.

Roberts, D. (2018). CHAS Malacology Collection (Arctos). Version 13.20. Chicago Academy of Sciences. Occurrence dataset https://doi.org/10.15468/tk35ga accessed via GBIF.org on 2019-01-12.

Roux, C., Fraisse, C., Romiguier, J., Anciaux, Y., Galtier, N., & Bierne, N. (2016). Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biology*, *14*(12), e2000234. https://doi.org/10.1371/journal.pbio.2000234

Ruffley, M., Smith, M. L., Espíndola, A., Carstens, B. C., Sullivan, J., & Tank, D. C. (2018). Combining allele frequency and tree-based approaches improves phylogeographic inference from natural history collections. *Molecular Ecology*, *27*(4), 1012–1024. https://doi.org/10.1111/mec.14491

Ruffley, M., Smith, M. L., Espíndola, A., Turck, D., Mitchell, N., Carstens, B. C., Sullivan, J., & Tank, D. C. (2022). Genomic evidence of an ancient inland temperate rainforest. *Molecular Ecology* (In Press).

Satler, Jordan D., & Carstens, B.C. (2017). Do ecological communities disperse across biogeographic barriers as a unit? *Molecular Ecology 26*(13), 3533-3545. https://doi.org/10.1111/mec.14137

Shafer, A. B., Cullingham, C. I., Cote, S. D., & Coltman, D. W. (2010). Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, *19*(21), 4589–4621. https://doi.org/10.1111/j.1365-294X.2010.04828.x

Schafer, A. B., Peart, C. R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C. W., & Wolf, B. W. (2017). Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*, *8*(8), 907-919. https://doi.org/10.1111/2041-210X.12700

Smith, M. L., & Carstens, B. C. (2020). Process-based species delimitation leads to identification of more biologically relevant species. *Evolution*, *74*(2), 216–229. https://doi.org/10.1111/evo.13878

Smith, M. L., Ruffley, M., Espíndola, A., Tank, D. C., Sullivan, J., & Carstens, B. C. (2017). Demographic model selection using random forests and the site frequency spectrum. *Molecular Ecology*, *26*(17), 4562–4573. https://doi.org/10.1111/mec.14223

Smith, M. L., Ruffley, M., Rankin, A. M., Espíndola, A., Tank, D. C., Sullivan, J., & Carstens, B. C. (2018). Testing for the presence of cryptic diversity in tail-dropper slugs (*Prophysaon*) using molecular data. *Biological Journal of the Linnean Society*, *124*(3), 518–532. https://doi.org/10.1093/biolinnean/bly067

Soltis, D. E., Gitzendanner, M. A., Strenge, D. D., & Soltis, P. S. (1997). Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. *Plant Systematics and Evolution*, *206*(1–4), 353–373. https://doi.org/10.1007/BF00987957

Steele, C. A., Carstens, B. C., Storfer, A., & Sullivan, J. (2005). Testing hypotheses of speciation timing in *Dicamptodon copei* and *Dicamptodon aterrimus* (Caudata: Dicamptodontidae). *Molecular Phylogenetics and Evolution*, *36*(1), 90–100. https://doi.org/10.1016/j.ympev.2004.12.001

Stone, G. N., White, S. C., Csóka, G., Melika, G., Mutun, S., Pénzes, Z., Sadeghi, S., Schönrogge, K., Tavakoli, M., & Nicholls, J. A. (2017). Tournament ABC analysis of the western Palearctic population history of an oak gall wasp, *Synergus umbraculus*. *Molecular Ecology, 26*(23), 6685-6703. https://doi.org/10.1111/mec.14372

Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., & Cosson, J. F. (1998). Comprative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology 7*(4):453-464. https://doi.org/10.1046/j.1365-294x.1998.00289.x

Thuiller, W., Georges, D., Engler, R. & Breiner, F. (2019). biomod2: Ensemble Platform for Species Distribution Modeling. R package version 3.3-7.1. https://CRAN.R-project.org/package=biomod2

Uribe, F., Agulló Villaronga, J. (2018). Museu de Ciències Naturals de Barcelona: MCNB-Malac. Museu de Ciències Naturals de Barcelona. Occurrence dataset https://doi.org/10.15468/pnkuwh accessed via GBIF.org on 2019-01-12.

Wallis, G. P., Waters, J. M., Upton, P., & Craw, D. (2016). Transverse alpine speciation driven by glaciation. *Trends in Ecology and Evolution 31*(12), 916-926. https://doi.org/10.1016/j.tree.2016.08.009

Wheeler, E. & McIntosh, H. (2018). Royal BC Museum - Invertebrates Collection. Version 1.1. Royal British Columbia Museum. Occurrence dataset https://doi.org/10.5886/zh7n1e accessed via GBIF.org on 2019-01-12.

Wilke, T., & Duncan, N. (2004). Phylogeographical patterns in the American Pacific Northwest: lessons from the arionid slug *Prophysaon coeruleum*. *Molecular Ecology*, *13*(8), 2303–2315. https://doi.org/10.1111/j.1365-294X.2004.02234.x

Data Accessibility: Raw reads for all species with new GBS or ddRAD data have been deposited in the NCBI Short Read Archive under project PRJNA847541. Custom scripts and input files are available in the Dryad Digital Repository (https://datadryad.org/stash/share/zeLGY68qeXTj-1se-2bM7hZ8KF_3I7A9ezN0m8p3zj8). Original files containing GBIF IDs downloaded from GBIF for *H. vancouverense* and *C. armata* are also available on Dryad.

Author contributions: MLS, DCT, JS, and BCC conceived the ideas behind the research. MLS and JW collected data, and MLS analyzed data. MLS led the writing with assistance from BCC. All authors read, edited, and approved a final version of the manuscript.

Tables

Table 1: Results from SDMs. ROC scores are the mean ROC scores across replicates and modelling strategies included in the ensemble model. The suitability (S) is the ratio of habitat suitability at the LGM to habitat suitability at the present at various thresholds. Nearest neighbor distance is the difference in median nearest neighbor distance between the current and the LGM at the 500 threshold. Negative values imply increased fragmentation during the LGM.

Species	*ROC*	*S (300)*	*S (500)*	*S (700)*	*S (900)*	*S (HDI)*	*Nearest neighbor distance*	*Algorithms*	*Samples after thinning*
P. andersoni	0.769	0.290	0.720	0.603	0.000	0.701	0.370	All	56
P. coeruleum	0.836	0.168	0.812	0.661	0.065	0.158	-0.324	All	41
P. dubium	0.817	0.337	0.380	0.170	0.000	0.313	-0.254	All	27
P. foliolatum	0.872	4.355	2.556	3.141	5.135	2.993	0.663	GLM	13
P. vanattae	0.842	0.716	0.026	0.012	0.000	0.023	-0.313	All	28
C. armata	0.883	0.746	0.758	0.062	0.000	0.346	0.329	All	29
H. vancouverense	0.829	0.121	0.014	0.000	0.000	0.105	-0.226	All	306

Table 2: Next-generation sequencing results. The number of samples sequenced, the number of loci sequenced in at least half of individuals, the number of unlinked SNPs used to assess population structure in STRUCTURE, and the number of unlinked SNPs used to compare models in *delimitR* are reported.

Species	*# Samples*	*# Loci sequenced in 50% of individuals*	*# Unlinked SNPs for STRUCTURE*	*# Unlinked SNPs for delimitR*
P. andersoni	88	18625	18527	8611
P. coeruleum	17	12454	10176	7507
P. dubium	48	20381	13492	8440
P. foliolatum	49	27339	25893	17902
P. vanattae	44	6855	6760	1754
C. armata	35	135	125	54
H. vancouverense	76	3746	3381	1213

Table 3: Results of demographic model selection in *delimitR*. We report error rates, the best model, and the posterior probability of the best model.

Species	error rate	best model	posterior probability
P. andersoni	0.05	four populations, secondary contact	0.689
P. coeruleum	0.04	two populations, secondary contact	0.686
P. dubium	0.04	two populations, divergence only	0.784
P. foliolatum	0.02	two populations, secondary contact	0.528
P. vanattae	0.10	three populations	0.706
C. armata	0.23	two populations, secondary contact, expansion	0.675
H. vancouverense	0.56	four populations, secondary contact, expansion	0.436

Figure Legends

Figure 1: A) Map of the Pacific Northwest (PNW) showing major features and the extent of the ice sheet (blue). B-F) Photos of study species by M.L. Smith; B) *P. andersoni;* C) *P. vanattae;* D) *H. vancouverense*; E) *P. foliolatum*; F) *P. dubium*.

Figure 2: Species Distribution Models (SDMs): Habitat suitability scores for each species across time periods. a-g) SDMs in the current and Last Glacial Maximum (LGM) for each species. h) Scale for maps. Suitability scores range from 0 to 1000.

Figure 3: Results of population clustering in STRUCTURE. Collections are plotted on maps depicting portions of the Pacific Northwest that were largely unglaciated (c.f. Fig. 1 for a broader perspective). Colors correspond to the population with the highest membership proportion, and insets show individual membership to each cluster. Individual membership proportions are shown only for admixed individuals. a) *P. andersoni*: Admixed individuals were found only in the South Cascades, and q-values for all of these individuals are displayed in the inset bar chart. b) *P. coeruleum*. Admixed individuals were found only in southwestern Oregon, and q-values for all of these individuals are displayed in the inset bar chart. c) *P. dubium*: Only two individuals showed evidence of mixed ancestry, and the inset bar chart shows q-values for these samples. d) *P. foliolatum*: Only samples from a single locality showed evidence of mixed ancestry, and the inset bar chart displays q-values at this locality. e) *P. vanattae*: A single individual showed limited evidence of admixture, and the q-values for this individual are displayed in the pie chart. f) *C. armata*: There was only evidence of mixed ancestry in the Northern Rocky Mountains, and the inset bar chart shows q-values of samples from this region. g) *H. vancouverense*: All populations except the northern samples from the Northern Rocky Mountains showed evidence of admixture, and q-values are displayed on inset bar charts. Shading represents topography.

Figure 4: Results of demographic model selection in delimitR. Population colors correspond to colors on the maps in Figure 3. Arrows indicate the presence of gene flow upon secondary contact. Expanding branch tips (triangles) indicate population expansion.

Figure 5: Divergence times estimated using the composite likelihood approach implemented in fastsimcoal2. Median divergence times estimated under the best model for each species, plus 95% HDI from bootstrap replicates. Scale is compared to the dates of glacial-interglacial cycles: Pre-Illinoian, Illinoian, Sangamonian Interstadial (S), Wisconsin (W), and Holocene (H) (Batchelor et al., 2014).

Figures

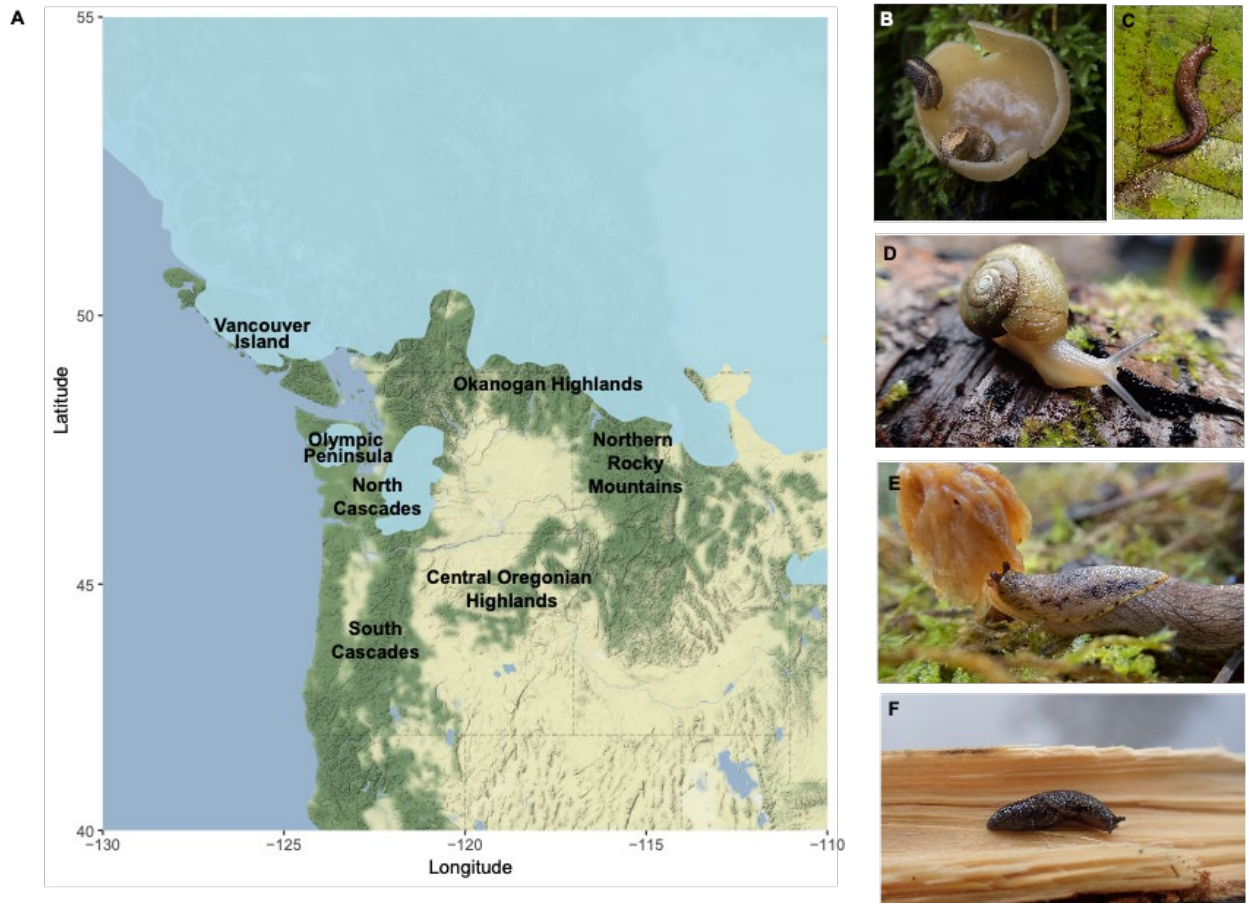

Figure 1: A) Map of the Pacific Northwest (PNW) showing major features and the extent of the ice sheet (blue). B-F) Photos of study species by M.L. Smith; B) *P. andersoni;* C) *P. vanattae;* D) *H. vancouverense*; E) *P. foliolatum*; F) *P. dubium*.

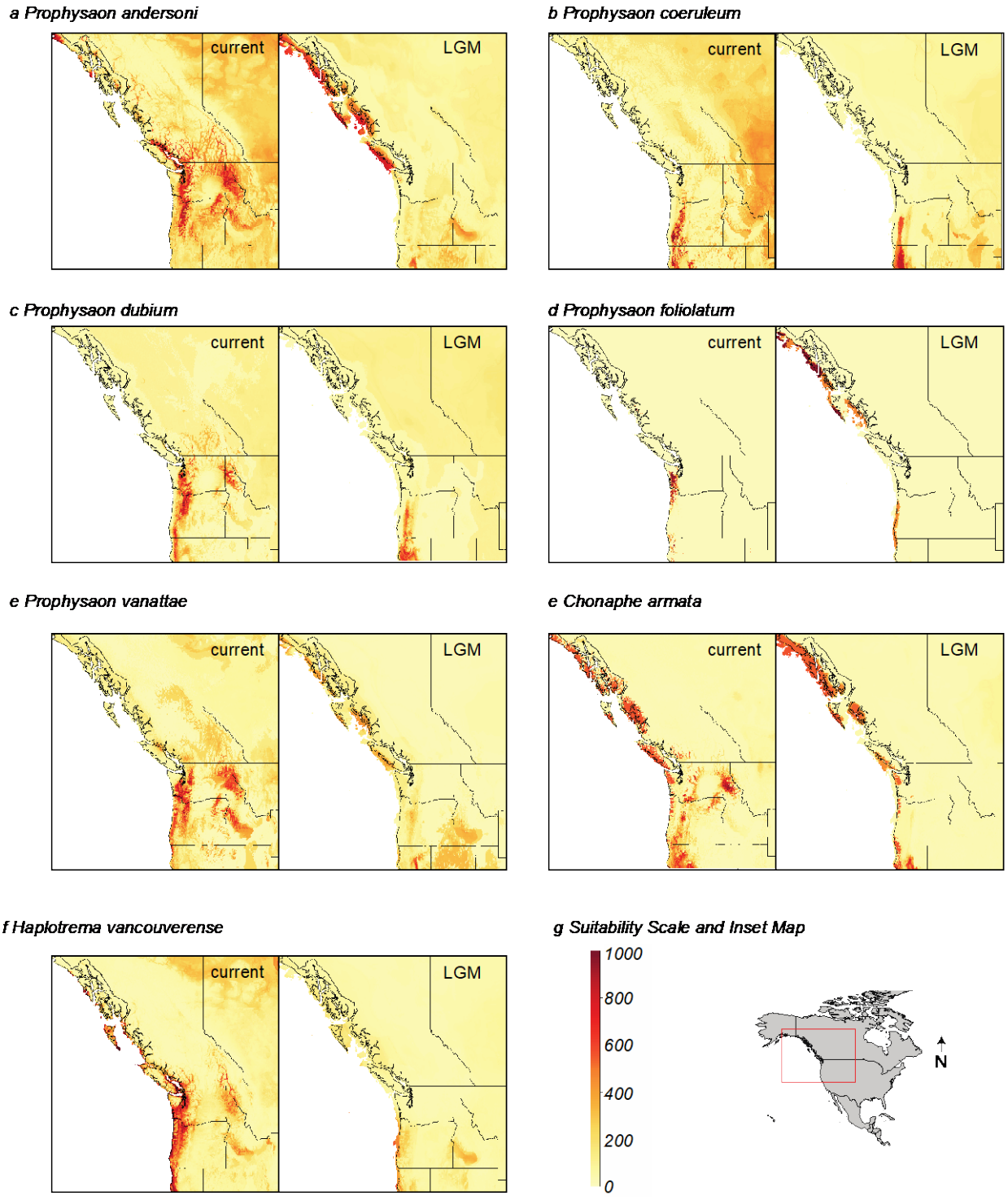

Figure 2: Species Distribution Models (SDMs): Habitat suitability scores for each species across time periods. a-g) SDMs in the current and Last Glacial Maximum (LGM) for each species. h) Scale for maps. Suitability scores range from 0 to 1000.

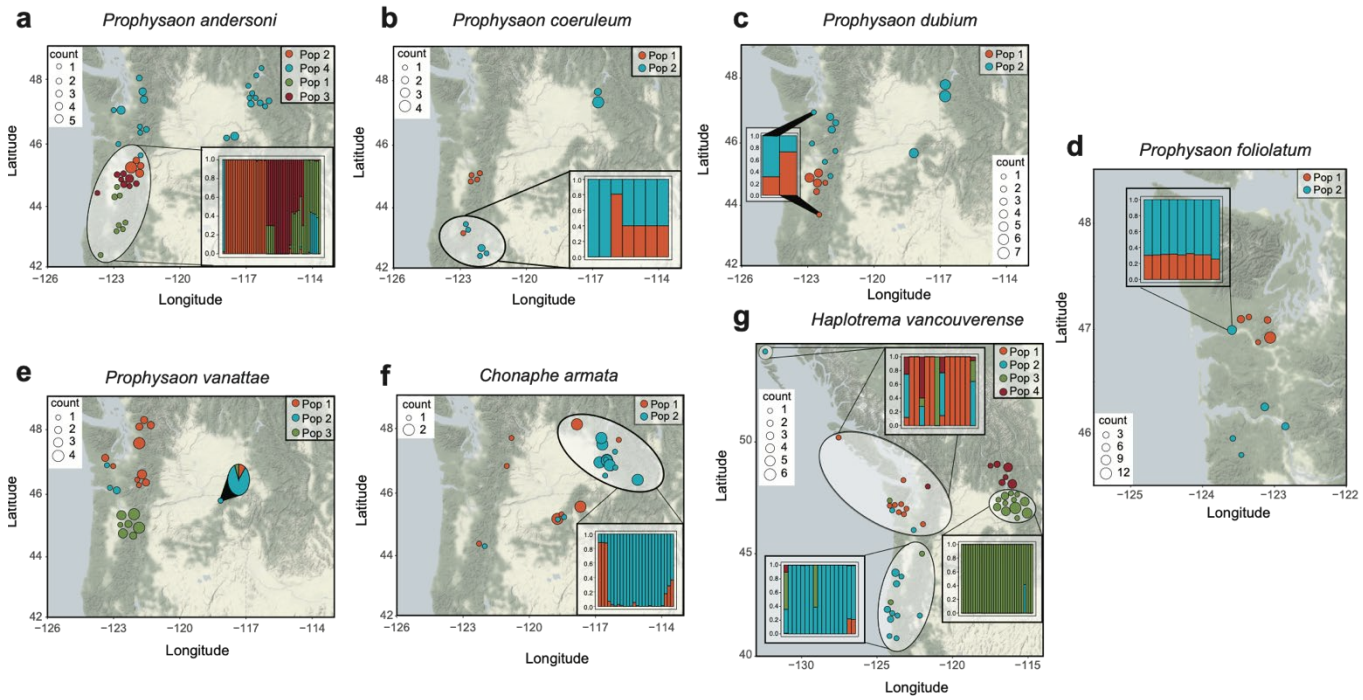

Figure 3: Results of population clustering in STRUCTURE. Collections are plotted on maps depicting portions of the Pacific Northwest that were largely unglaciated (c.f. Fig. 1 for a broader perspective). Colors correspond to the population with the highest membership proportion, and insets show individual membership to each cluster for all localities that contain admixed individuals ($q < 0.90$). The size of points is scaled based on the number of individuals collected from the locality. a) *P. andersoni*: Admixed individuals were found only in the South Cascades, and q-values for all of these individuals are displayed in the inset bar chart. b) *P. coeruleum*. Admixed individuals were found only in southwestern Oregon, and q-values for all of these individuals are displayed in the inset bar chart. c) *P. dubium*: Only two localities showed evidence of mixed ancestry, and the inset bar chart shows q-values for these samples. d) *P. foliolatum*: Only samples from a single locality showed evidence of mixed ancestry, and the inset bar chart displays q-values at this locality. e) *P. vanattae*: A single individual (the only individual at the locality) showed limited evidence of admixture, and the q-values for this individual are displayed in the pie chart. f) *C. armata*: There was only evidence of mixed ancestry in the Northern Rocky Mountains, and the inset bar chart shows q-values of samples from this region. g) *H. vancouverense*: All populations except the northern samples from the Northern Rocky Mountains showed evidence of admixture, and q-values are displayed on inset bar charts. Shading represents topography.

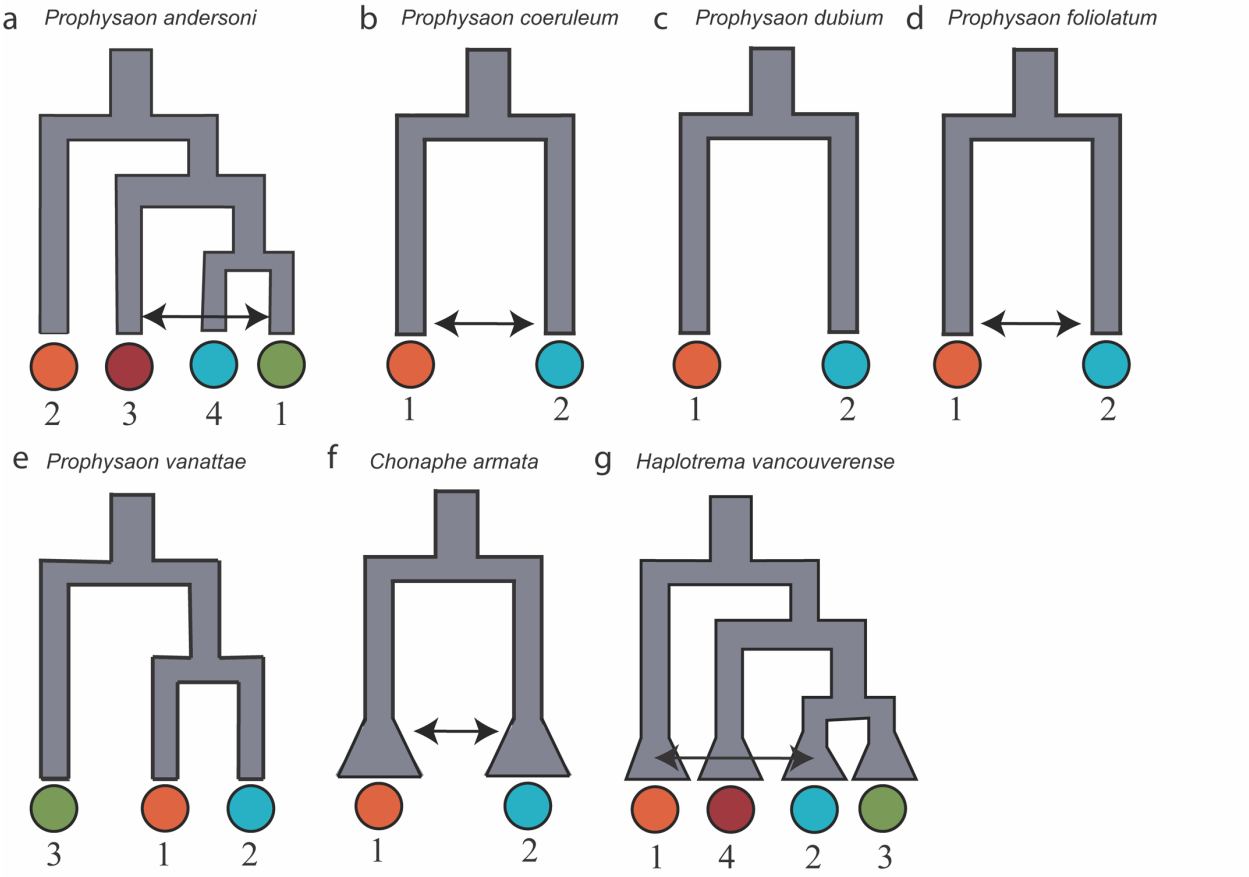

Figure 4: Results of demographic model selection in delimitR. Population colors correspond to colors on the maps in Figure 3. Arrows indicate the presence of gene flow upon secondary contact. Expanding branch tips (triangles) indicate population expansion.

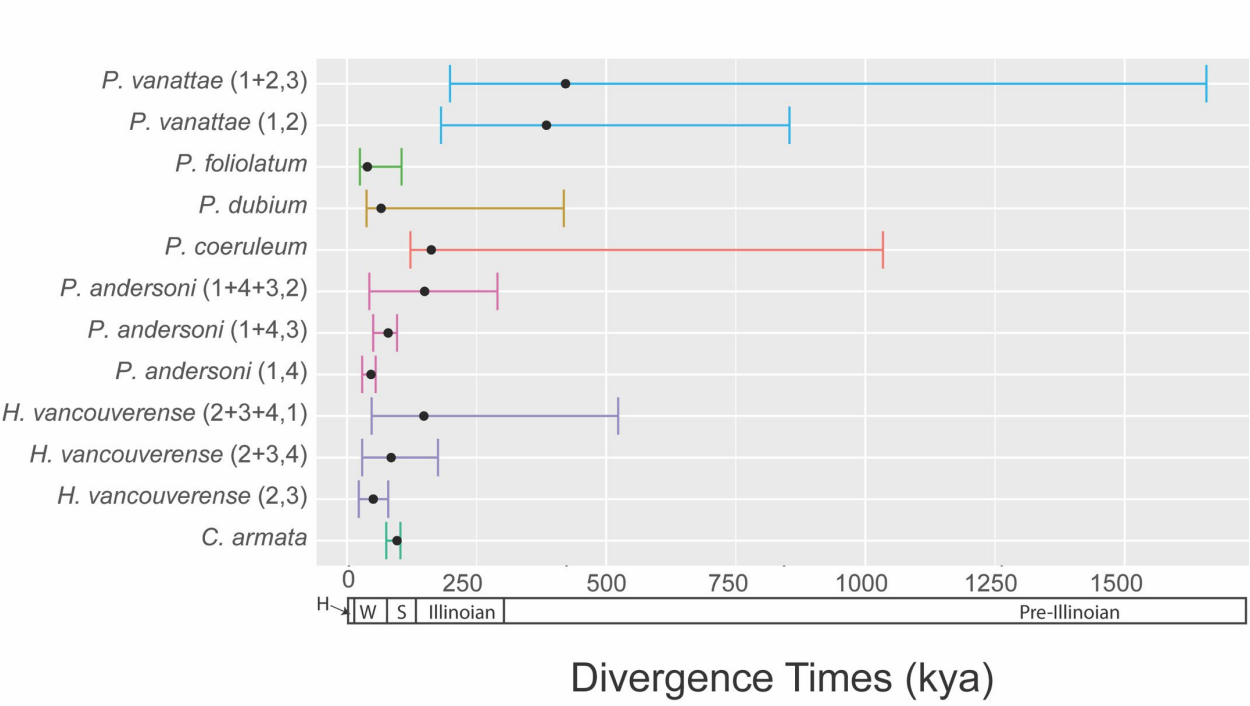

Figure 5: Divergence times estimated using the composite likelihood approach implemented in fastsimcoal2. Median divergence times estimated under the best model for each species, plus 95% HDI from bootstrap replicates. Scale is compared to the dates of glacial-interglacial cycles: Pre-Illinoian, Illinoian, Sangamonian Interstadial (S), Wisconsin (W), and Holocene (H) (Batchelor et al., 2014).