

# INTEGRATING COALESCENT AND ECOLOGICAL NICHE MODELING IN COMPARATIVE PHYLOGEOGRAPHY

Bryan C. Carstens<sup>1</sup> and Corinne L. Richards<sup>2</sup>

*Department of Ecology and Evolutionary Biology, 1109 Geddes Ave., Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109-1079*

<sup>1</sup>E-mail: bcarsten@umich.edu

<sup>2</sup>E-mail: clrichar@umich.edu

Received November 6, 2006

Accepted January 23, 2007

Understanding the factors that contribute to the formation of population genetic structure is a central goal of phylogeographic research, but achieving this goal can be complicated by the stochastic variance inherent to genetic processes. Statistical approaches to testing phylogeographic hypotheses accommodate this stochasticity by evaluating competing models of putative historical population structure, often by simulating null distributions of the expected variance. The effectiveness of these tests depends on the biological realism of the models. Information from the fossil record can aid in reconstructing the historical distributions of some taxa. However, for the majority of taxa, which lack sufficient fossils, paleodistributional modeling can provide valuable spatial-geographic data concerning ancestral distributions. Paleodistributional models are generated by projecting ecological niche models, which predict the current distribution of each species, onto a model of past climatic conditions. Here, we generate paleodistributional models describing the suitable habitat during the last glacial maximum for lineages from the mesic forests of the Pacific Northwest of North America, and use these models to generate alternative phylogeographic hypotheses. Coalescent simulations are then used to test these hypotheses to improve our understanding of the historical events that promoted the formation of population genetic structure in this ecosystem. Results from Pacific Northwest mesic forest organisms demonstrate the utility of these combined approaches. Paleodistribution models and population genetic structure are congruent across three amphibian lineages, suggesting that they have responded in a concerted manner to environmental change. Two other species, a willow and a water vole, despite being currently codistributed and having similar population genetic structure, were predicted by the paleodistributional model to have had markedly different distributions during the last glacial maximum. This suggests that congruent phylogeographic patterns can arise from incongruent ancestral distributions. Paleodistributional models introduce a much-needed spatial-geographic perspective to statistical phylogeography. In conjunction with coalescent models of population genetic structure, they have the potential to improve our understanding of the factors that promote population divergence and ultimately produce regional patterns of biodiversity.

**KEY WORDS:** Coalescent modeling, ecological niche modeling, statistical phylogeography, Pleistocene, comparative phylogeography.

Most phylogeographic research is conducted under the premise that the demographic history of the focal organism can be inferred from a single temporal sample of the genetic and geographic variability. Theoretical explorations pertaining to the genetic ba-

sis of phylogeographic research, such as those related to coalescent theory (Kingman 1982; Hudson 1992), the expected time to monophyly (Pamilo and Nei 1988; Takahata 1989; Hudson and Coyne 2002; Rosenberg 2003), and methods for dating lineage

divergence (Arbogast et al. 2002 and references within) are well developed. However, spatial-geographic aspects of phylogeographic research are less developed (Kidd and Ritchie 2006), possibly because fossil data are not available for many species (Hugall et al. 2002), leaving little direct evidence of the past distribution of the focal organism. Because organismal ranges can change dramatically over time (Lawton 1993; Gaston 1996), incorporating information about these temporal changes is likely to improve inferences about the factors that promote the formation of population genetic structure. For species in which fossil data are not available, paleodistributional modeling allows historical species distributions to be estimated, improving inferences about demographic history (Hugall et al. 2002; Peterson et al. 2004). Paleodistributional modeling may be particularly valuable to comparative phylogeographic studies, in which genealogical concordance among codistributed species is often interpreted as evidence that the population genetic structure in each species has formed in response to similar environmental events. Implicit in this interpretation is the assumption that species that are currently codistributed would have been similarly codistributed in the past.

Paleodistributional models are models of the ecological niche of the organism that have been projected onto models of past climate. This technique uses climatic data from sites in which species are known to exist to determine the set of climatic parameters which best predict the presence of the species (Austin 1985; Peterson 2001). When projected onto a map, the resulting models predict the habitat that a species could potentially inhabit. Ecological niche modeling approaches have been applied to a variety of research questions, including explorations of speciation mechanisms (Graham et al. 2004a), extinction (Martinez-Meyer et al. 2004), and hybrid zones (Swenson 2006). Ecological niche modeling and phylogeographic analyses are complementary because the inferences from one approach can be explored and potentially validated by the other. For example, Hugall et al. (2002) combined these approaches to investigate the phylogeography of a snail endemic to the Australian wet tropics and compared its pattern with that of several vertebrates, Bond et al. (2006) used them to investigate the causes of population extinction in *Apomastus* spiders, whereas Rissler et al. (2006) explored different strategies for conservation by predicting geographic breaks between lineages. Although there are many potential applications within evolutionary biology (e.g., Stigall and Lieberman 2006; Yesson and Culham 2006), the broad taxonomic applicability of this approach makes it particularly suited to comparative phylogeography (Hugall et al. 2002). For example, paleoclimatic reconstructions are useful for identifying putative locations for Pleistocene refugia (Peterson et al. 2004), and may also provide information about potential dispersal corridors.

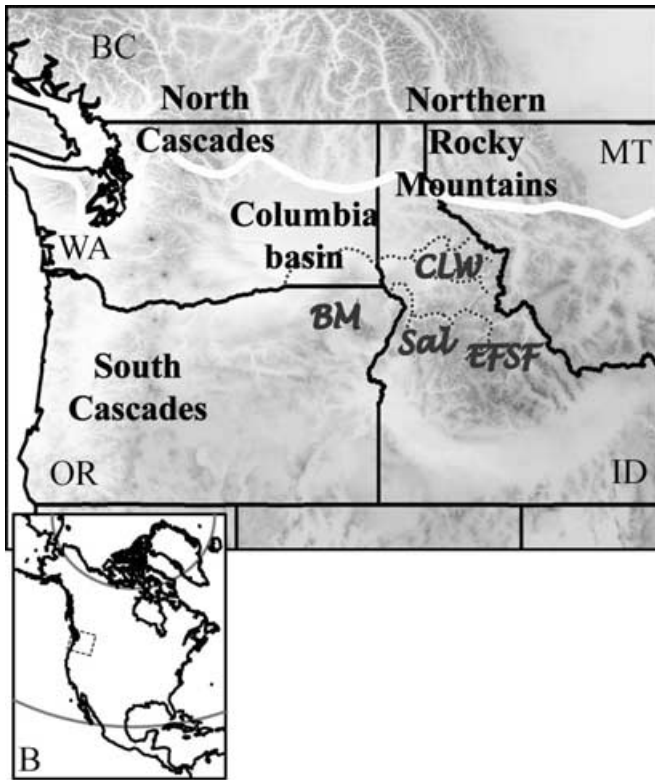
Here we construct models of the distributions of species from the Pacific Northwest mesic forest (PNW-MF) ecosystem

of North America during the last glacial maximum, and compare these putative historical ranges to current ranges. The paleodistributional models are then used as a framework for statistical phylogeographic hypothesis testing, in which parametric simulation is used to construct null distributions of the expected variance of relevant parameters under competing historical scenarios (e.g., Knowles 2004). This approach has proven to be a useful tool for phylogeographic research (e.g., Knowles 2001; Carstens et al. 2005a; DeChaine and Martin 2005; Russell et al. 2005; Steele and Storfer 2006), but in the absence of explicit reconstructions of the past range of the focal organism. By integrating paleodistributional models and statistical phylogeography, we hope to increase the realism of historical models, thereby improving phylogeographic inferences about the relative importance of particular climatic events to the formation of population genetic structure.

## Background

Western hemlock (*Tsuga heterophylla*) and western redcedar (*Thuja plicata*) are the late successional dominants that characterize the temperate mesic forests of the Pacific Northwest (Daubenmire and Daubenmire 1968; Daubenmire 1975). *Tsuga* and *Thuja* forests occupy two disjunct regions in the Pacific Northwest; one in the Cascades range and along the Pacific coast and a second, ecologically isolated portion in the northern Rocky Mountains west of the continental divide (Fig. 1). Fossil data suggest that *Tsuga* and *Thuja* forests have been present in the northern Rocky Mountains since the mid-Eocene (Graham 1993, 1999), or shortly after the formation of the northern Rocky Mountains some 45–36 million years ago (mya) (English and Johnson 2004). The Cascades orogeny during the Pliocene (5–2 mya) created a rain-shadow that dried out the intervening Columbia basin, presumably dividing a formerly continuous coniferous forest (Graham 1999). Today, the 300-km-wide expanse of xeric shrub and steppe habitat in the Columbia basin forms an effective barrier to dispersal between the Cascades and northern Rocky Mountains (Brunsfeld et al. 2001).

The Pleistocene glacial cycles and the accompanying climate change had a large influence on the evolution of the PNW-MF biota, as they did in other regions (e.g., Gates 1993; Hewitt 1996; Schneider et al. 1998). During the Pleistocene, the entire region was subjected to repeated glaciation, which covered a significant portion of the northern Rocky Mountains but only the northernmost portions of the Cascades range (Pielou 1991; Delcourt and Delcourt 1993). Although fossil data suggest that *Thuja* and *Tsuga* forests were widespread in the northern Rockies from the middle Eocene until the Pleistocene (Graham 1993, 1999), it is unclear if the *Thuja* and *Tsuga* forests persisted in the northern Rocky Mountains throughout the Pleistocene glaciations or



**Figure 1.** Map of the Pacific Northwest of North America. The following abbreviations are used: Clearwater River drainage (CLW), Salmon River drainage (Sal), Blue Mountains (BM), East Fork of the South Fork of the Salmon (EFSF), British Columbia (BC), Washington State (WA), Oregon (OR), Idaho (ID), and Montana (MT). The maximum extent of the Cordilleran glacier is shown with the white line (Booth et al. 2004). 1B shows the location of the PNW-MF on a map of the North American continent.

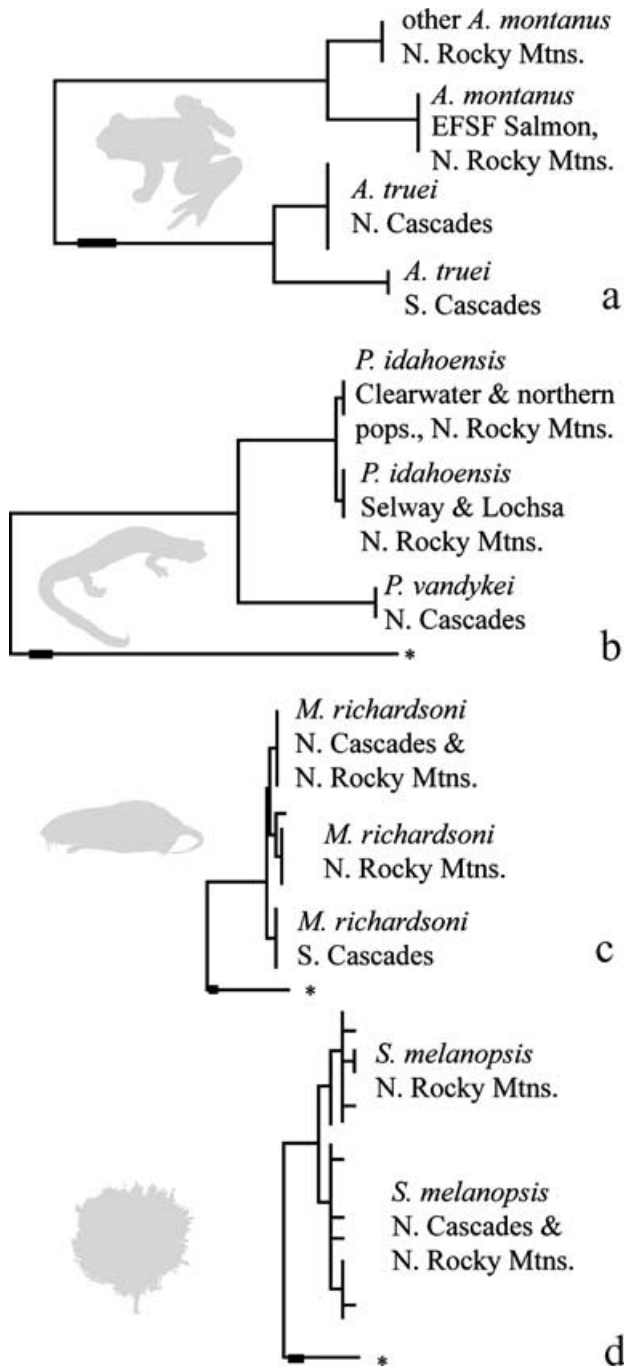
recolonized this region during the Holocene (Mehring 1985). Although large portions of the South Cascades were unglaciated during the Pleistocene, most of the current distribution of *Tsuga* and *Thuja* forests in the northern Rocky Mountains were repeatedly covered by glaciers (Brunsfeld et al. 2001). An exception is the Clearwater river drainage (Fig. 1), which was proposed as a likely Pleistocene refuge (Daubenmire 1952; Detling 1968) because it is currently the southern extent of *Tsuga* and *Thuja* forests in the Rocky Mountains and was the northernmost river system to the south of Pleistocene glaciation.

On the basis of the above, Brunsfeld et al. (2001) proposed two general hypotheses that explain the formation of the disjunct mesic forest distribution. One of these (ancient vicariance) posits that populations in the Cascades and northern Rocky Mountains were isolated following the xerification of the Columbia basin, and subsequently able to persist in refugia throughout the Pleistocene glaciations. The ancient vicariance hypothesis predicts that, within any PNW-MF lineage, populations from the northern Rocky Mountains and Cascades will be reciprocally mono-

phyletic and separated by deep temporal divergence. The second hypothesis (recent dispersal) is based on the lack of *Thuja* and *Tsuga* fossils from the unglaciated portions of the northern Rocky Mountains during the Pleistocene. It posits that populations in northern Rocky Mountains are derived from post-Pleistocene dispersal from either the North or South Cascades, and predicts that haplotypes from the northern Rocky Mountains will be nested within the source populations with extremely recent temporal divergence.

Dispersal and vicariance are classic biogeographic explanations for disjunct distributions, and to date each has been supported for some species by genetic data sampled from PNW-MF species. For example, the predictions of the ancient vicariance hypothesis could not be rejected in three amphibian clades (Nielson et al. 2001; Carstens et al. 2004; Steele et al. 2005) in which Cascadian and northern Rocky Mountains populations are reciprocally monophyletic and separated by up to 10% sequence divergence (Fig. 2a, b). In contrast, the recent dispersal hypothesis could not be rejected in a vole and a willow (Carstens et al. 2005b), in which haplotypes from the northern Rocky Mountains are nested within those from the North Cascades and sequence divergence is low (Fig. 2c, d). It is unclear what factors account for these incongruent evolutionary histories. Codistributed organisms may differ in inherent dispersal abilities, but these dispersal differences do not necessarily produce incongruent phylogeographic patterns. The ecological niches of codistributed organisms are not likely to be identical, and temporal changes in the geographic distribution of these niches may not be congruent across organisms, leading to dissimilar opportunities for dispersal. In combination these differences may result in incongruent patterns even among organisms that currently, and historically, have inhabited the same general region. Nevertheless, within any evolutionary lineage, the paleodistributional models and genetic data should be reciprocally consistent as to the timing and geography of dispersal and vicariance among populations. For example, if the genetic data suggest that there has been recent dispersal between particular regions, one might reasonably expect the paleodistributional model to provide evidence of suitable habitat between these regions. Therefore, each of these approaches offers an independent test of the a priori hypotheses outlined earlier.

Here we construct ecological niche models for four PNW-MF species or species-pairs, and explore the degree to which their distribution has changed since the last glacial maximum of the Pleistocene by projecting the current distribution model from each species onto a reconstruction of the climate from 21,000 years before present (ybp). The last glacial maximum is a key event in the recent history of the PNW-MF because it represents the opposite climatic extreme from the present interglacial climate. As such, if changes in geographic distributions are correlated with climatic changes, the distributions of these organisms would likely have



**Figure 2.** Maximum likelihood estimates of phylogeny, redrawn from Carstens et al. (2005b) for four PNW mesic forest organismal lineages. (a) *A. truei* and *A. montanus*; (b) *P. vandykei* and *P. idahoensis*; (c) *M. richardsoni*; (d) *S. melanopsis*. Scale bars on the branches represent 0.005 substitutions per site. For details of the phylogenies, see Carstens et al. (2005b).

been as different from their present range during the last glacial maximum as during any period in the past. We then test a set of phylogeographic hypotheses, generated from the paleodistributional models, using a coalescent approach. Our hypotheses concern the structure of Pleistocene refugia in two species whose

population genetic structures are consistent with the predictions of the ancient vicariance hypothesis, and the likely routes of dispersal in two other species whose population genetic structures are consistent with the predictions of the recent dispersal hypothesis.

## Materials and Methods

### COLLECTION LOCALITIES

Collection localities were compiled from sources given in Appendix I for the following taxa: sister species of tailed frogs (*Ascaphus truei*, *A. montanus*), sister species of plethodontid salamanders (*Plethodon vandykei*, *P. idahoensis*), a vole (*Microtus richardsoni*), and the dusky willow (*Salix melanopsis*). Paleodistribution models were also generated for other PNW-MF species, including the ecosystem dominants western redcedar (*Thu. plicata*) and western hemlock (*T. heterophylla*), the white-bark pine (*Pinus albicaulis*), and the Pacific Giant salamanders (*Dicamptodon* spp.).

### ENVIRONMENTAL VARIABLES USED IN ECOLOGICAL NICHE MODELING

Mean monthly climate estimates at a spatial resolution of 10 arc-minutes were generated for the last glacial maximum of the Pleistocene (LGM) under the Community Climate System Model (CCSM3: Collins et al. 2006) and statistically downscaled (Ruegg et al. 2006) by Robert J. Hijmans with data provided by the Paleoclimate Modelling Intercomparison Project Phase II (PMIP2). The CCSM3 is composed of four separate models that simultaneously simulate the earth's atmosphere, ocean, land surface and sea-ice, and one central coupler component (National Center for Atmospheric Research: <http://www.cesm.ucar.edu/>).

For current conditions (1950–2000), we used WorldClim, a global climate database (Hijmans et al. 2005) at the same resolution. Using the outputs from the CCSM3 and WorldClim models, which in both cases consisted of maximum and minimum monthly temperature and precipitation estimates, 19 bioclimatic parameters were derived following those of ANUCLIM (Australian National University: <http://cres.anu.edu.au/outputs/anuclim.php>), to capture biologically meaningful aspects of climate variation. These parameters incorporate annual trends (e.g., mean annual temperature, annual precipitation), aspects of seasonality (e.g., annual range in temperature and precipitation) and extreme or potentially limiting environmental factors (e.g., temperature of the coldest and warmest months, and precipitation of the wettest and driest months).

### ECOLOGICAL NICHE MODELING

Climate models were generated using the maximum entropy machine learning algorithm in MAXENT (Phillips et al. 2006). For each species group, a distribution model was generated using the

19 bioclimatic parameters for the current climate and the collection localities. This model was then applied to the set of bioclimatic variables generated from CCSM3 to estimate the geographic distribution and extent of suitable climate for each species during the last glacial maximum. Model performance was evaluated using receiver operating characteristic (ROC) analysis, where model evaluation proceeds by a sensitivity analysis of the presence and absence points predicted correctly by the model. Binomial tests of omission (Phillips et al. 2006) were also conducted by randomly selecting 25% of the occurrence localities as test data and using 10,000 randomly chosen pixels from the study region as random instances. For the binomial test, MAXENT generates *P*-values for 10 different threshold values following different approaches for predicting species presence/absence from continuous model output (reviewed in Liu et al. 2005). Because one of the predictions of the ancient vicariance hypothesis is that suitable habitat in the northern Rocky Mountains will be restricted during glacial maxima, climatic models were based on collection localities from both the inland and coastal species in the amphibian lineages. However, these paleoclimate models did not differ appreciably from the models constructed using only the northern Rocky Mountain species in the inland region.

#### SEQUENCE DATA

Genetic data from four species were used: *P. idahoensis*, 669 bp of the mtDNA Cytochrome *b* gene from 229 individuals (Carstens et al. 2004); *A. montanus*, 1280 bp of the mtDNA genes *cyt b* and ND-4 from 56 individuals (Nielson et al. 2001); *M. richardsoni*, 747 bp of Cytochrome *b* from 58 individuals (Carstens et al. 2005b); *S. melanopsis*, 1159 bp of the *cpDNA* gene maturase *K* and *rp116* intergenic spacer from 155 individuals (Brunsfeld et al. 2006). GenBank accession numbers can be found in Table 1. Two different phylogeographic patterns are represented by these data (Fig. 2). In the *Ascaphus* and *Plethodon* lineages, haplotypes sampled from the northern Rocky Mountains and Cascades are reciprocally monophyletic, and divergence between these populations predates the Pleistocene (Carstens et al. 2005b). In contrast, haplotypes from the northern Rocky Mountains and North Cascades are intermixed in *M. richardsoni* and *S. melanopsis*, and population divergence likely dates to the end of the Pleistocene and Holocene

**Table 1.** GenBank accession numbers for data used in the coalescent modeling.

Species	Accession number
<i>A. montanus</i>	AF277324-277352, AF277370-277353
<i>P. idahoensis</i>	AY572039-572107
<i>M. richardsoni</i>	AY073753-973809
<i>S. melanopsis</i>	DQ875023-875036, DQ975037-875049, DQ060264-060269, DQ060271-060276

(Carstens et al. 2005b). Phylogeographic hypotheses were tested previously in *Dicamptodon* spp. (Carstens et al. 2005a; Steele and Storfer 2006) and were not repeated here. Sequence data with sufficient variability for testing these phylogeographic hypotheses are not currently available in the other species (e.g., *Thu. plicata*, *T. heterophylla*, *Pin. albicaulis*).

#### COALESCENT MODELING IN STATISTICAL PHYLOGEOGRAPHY

Each phylogeographic hypothesis consists of a model of population history that necessarily includes the branching order of population divergence and the temporal divergence between populations (measured in organismal generations). Additional parameters can be incorporated, such as the effective population size or particular demographic events such as population bottlenecks. Although there is no formal procedure for parameter selection, effective models should have the fewest number of parameters that allow them to be differentiated given the data while still capturing the biological essence of the problem (Knowles 2004). Once the model of population history is defined, genealogies consistent with the model of population history are simulated under a neutral coalescent model, and sequence data are simulated on the genealogies. MESQUITE 1.04 (Maddison and Maddison 2004) was used to simulate both the genealogies and the sequence data.

The models of population divergence used to test phylogeographic hypotheses were based on qualitative examinations of the paleodistribution model for each species. For example, in *A. montanus*, the most suitable habitat during the last glacial maximum, is contained within two discrete regions if the threshold for the strength of prediction was above 65, and these divisions were used to form the model of population divergence. Owing to differences in the strength of prediction across taxa, these threshold levels were not equal across species. Because strength of prediction is influenced by factors both extrinsic (e.g., sampling, model performance, influence of various climatic variables) and intrinsic (niche breadth) to the species' ecology, choosing a single value across taxa is difficult (Liu et al. 2005). Nevertheless, in each species, the regions chosen to represent ancestral populations represented the regions with the highest predicted suitability. Data were simulated using models of sequence evolution that match those selected from the empirical data. The mean number of variable sites in the simulated data corresponded to the number of variable sites in the empirical data, and values of  $\theta = 2N_e\mu$  matched those estimated from the empirical data (Table 2). The goal of the simulations was to reproduce the characteristics of the empirical data, regardless of the particular hypothesis being tested, so that comparisons of the test statistic to the null distribution would be meaningful. Because effective population size ( $N_e$ ) is a critical parameter in the evaluation of these models, tests

**Table 2.** Shown for each species: the number of samples with genetic data (N), the gene used (GENE), the length of the sequence (BP), the number of variable sites (S), the model of sequence evolution selected using DT-ModSel (Minin et al. 2003), the generation length used to scale putative models of population history (GEN.),  $\theta$  (with lower and upper 95% CIs) estimated using LAMARC (Kuhner et al. 2005), and the assumed mutation rate (see Carstens et al. 2005b).

Lineage	n	Gene	bp	s	Model	Gen.	$\theta_{\text{low}}$	$\theta_{\text{est}}$	$\theta_{\text{high}}$	$\mu$
<i>A. montanus</i>	56	ND-4, cyt b	1280	41	TrN	6	0.004234	0.006074	0.008974	$1.0 \times 10^{-7}$
<i>P. idahoensis</i>	229	cyt b	669	85	F <sub>81</sub> +G	5	0.034653	0.041251	0.048321	$1.0 \times 10^{-7}$
<i>M. richardsoni</i>	58	cyt b	747	59	GTR+I	1	0.021331	0.027969	0.037073	$1.0 \times 10^{-6}$
<i>S. melanopsis</i>	155	matK+rpl16	1159	7	F <sub>81</sub>	4	0.001226	0.002163	0.003034	$1.0 \times 10^{-6}$

were conducted at values of  $N_e$  that corresponded to empirical estimates of  $\theta$  made using the LAMARC software package (Kuhner et al. 2005). LAMARC generates a maximum-likelihood estimate of  $\theta$  using Markov-chain Monte Carlo approaches to sample the genealogies that contribute the most to the likelihood of the given model. Each test was conducted using an  $N_e$  that corresponds to the point estimate of  $\theta$  (1000 replicates), as well as the 95% upper and lower confidence intervals (100 replicates).

To test the predictions of each phylogeographic hypothesis, null distributions were constructed in the following manner: (1) PAUP\* 4.0b (Swofford 2002) was used to search simulated data for the most parsimonious tree(s) using TBR branch-swapping, 10 random addition replicates, and max trees = 100; (2) a majority-rule consensus of the most parsimonious trees was calculated; (3) a test statistic was computed from each dataset and then used to construct the null distribution; and (4) the same parameter from the empirical data was compared with the simulated null distribution. Two parameters were used to evaluate hypotheses: (1) the probability of monophyly of a particular clade or clades in the genealogy, calculated using a constraint-tree filter in PAUP\*; and (2) the average genetic divergence between the Cascades and northern Rocky Mountains, calculated using PAUP\*. This general approach provides a statistical framework for determining whether the empirical data differed significantly from theoretical expectations under a given model, while taking into account the inherent stochasticity of genetic processes (Knowles and Maddison 2002). Specific details of the models for each species are given below.

#### *Species where the population genetic structure is consistent with the ancient vicariance hypothesis*

In the *A. truei*/*A. montanus* and the *P. vandykei*/*P. idahoensis* species complexes, the predictions of the ancient vicariance hypothesis cannot be rejected (Nielson et al. 2001; Carstens et al. 2004), suggesting that populations in the northern Rocky Mountains persisted throughout the Pleistocene glaciations in one or more refugia. Coalescent simulations under population models that are compatible with the habitat reconstructions at the last glacial maximum provide an explicit test of the putative refugial population structure. In each of the inland species, two models

were tested: (1) The single refugium model (SRM) posits that all populations are derived from a single source population at the end of the Pleistocene. (2) The dual refugia model (DRM) holds that the species were isolated in at least two refugia during the Wisconsinian glaciations, which began some 114,000 ybp (Gibbard and Van Kolfshoten 2004). The location of these putative refugial populations was derived from the paleoclimatic model and corresponded to one of several deep river canyons in northern Idaho. These canyons were previously hypothesized as putative refugia due to the climatic insulation presumably offered by their lower elevation (Daubenmire 1952; Detling 1968). In *A. montanus*, simulations were conducted at three values of  $N_e$  (21170, 30370, 44870), corresponding to the 95% lower confidence interval, point estimate, and 95% upper CI of the parameter  $\theta$ , using a generation length of six years. On the basis of this length, there have been 3500 *A. montanus* generations of since the LGM, and 19,000 generations since the end of the Sangamon interglacial (114,000 ybp). From the Sangamon to the present interglacial represents one complete cycle of Pleistocene climatic fluctuation. In *P. idahoensis*, simulations were conducted across the range of  $N_e$  values (173,265, 206,255, 241,605) using five years per generation. On the basis of this generation length, there have been 4200 generations since the last glacial maximum and 22,800 generations since the Sangamon interglacial. Because there is evidence of dramatic post-Pleistocene population expansion in *P. idahoensis* (Carstens et al. 2004), the SRM and DRM were also tested using population bottlenecks, in which the northern and southern refugial populations were restricted to 0.87 and 0.23, respectively, of the total  $N_e$  during the last glacial maximum. These values are based on  $\theta$  estimated only from these populations using LAMARC. The above simulations used the probability of reciprocal monophyly for clades corresponding to the putative refugia as the test statistic.

#### *Species where the population genetic structure is consistent with the recent dispersal hypothesis*

The ancient vicariance hypothesis was rejected in both *S. melanopsis* and *M. richardsoni* (Carstens et al. 2005b), suggesting that the disjunct distribution in these species resulted from recent

dispersal between the North Cascades and northern Rocky Mountains. These inferences result from both the admixture of haplotypes sampled in these regions, as well as the recent temporal divergence between populations. Coalescent-based paleodistribution models were designed to explore the founding of populations and the formation of population structure. For example, if suitable habitat was restricted to one region during the last glacial maximum, the model of population structure treated this region as the putative source population. In each species, divergence from a single Pleistocene refugium was also modeled. This model predicts that populations were isolated in the Salmon and Clearwater refugia since the beginning of the Wisconsinian glaciation some 114,000 ybp (28,500 generations). For *M. richardsoni*, the probability of reciprocal monophyly and genetic distance were used as test statistics, whereas only genetic distances were used for *S. melanopsis*. In *M. richardsoni*, simulations were conducted across

a range of  $N_e$  (10,666, 13,985, 18,537) using a generation time of one generation per year. In *S. melanopsis*, simulations were conducted across a range of  $N_e$  (1226; 2163; 3034) using a generation length of four years per generation.

## Results

### SPECIES DISTRIBUTION MODELS

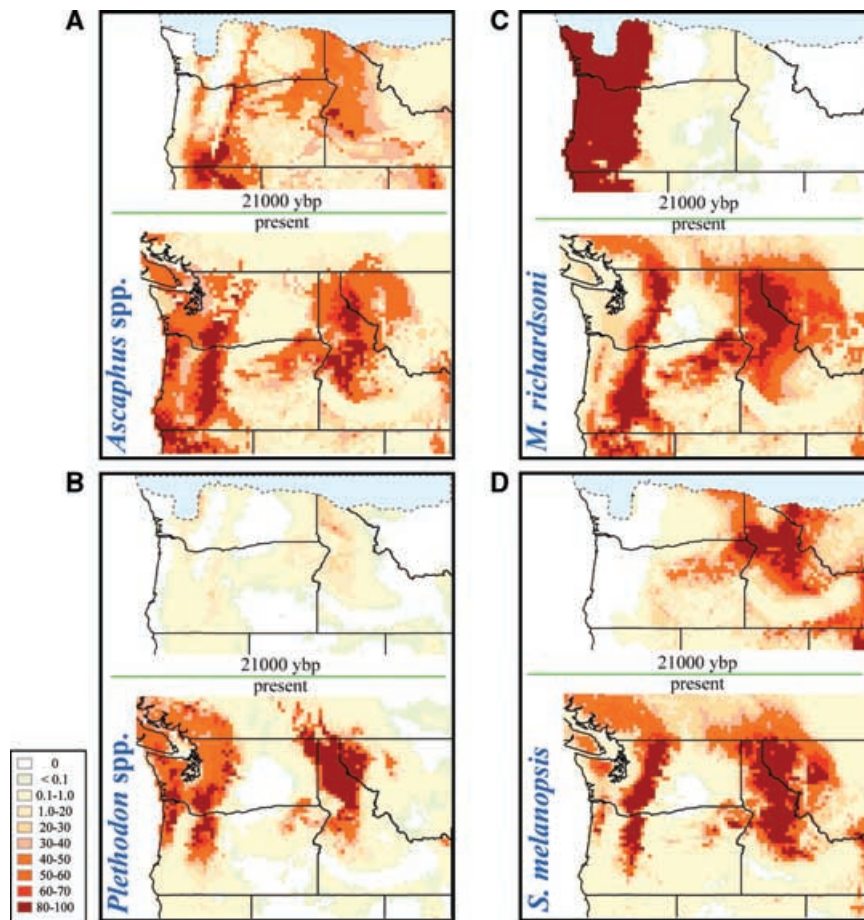
MAXENT was used to generate ecological niche models for each of the focal species or species groups under current climate conditions. For each distribution model, the area under the ROC curve (AUC) was close to one ( $AUC \geq 0.998$ ) indicating better than random predictions ( $0.5 = \text{random}$ ,  $1 = \text{maximum}$ ). Furthermore, all models produced predictions that were significantly better than random ( $P < 0.001$  for all alternative threshold values) based on a binomial test of omission. MAXENT produces a continuous

**Table 3.** Area under curve (AUC) statistic for each of 19 climatic variables, in each of seven PNW mesic forest organismal lineages. For each species or species group, the two climate variables with the highest individual AUC score are marked with an asterisk.

Climate variable	<i>Ascapus</i> spp.	<i>Plethodon</i> spp.	<i>Microtus</i> <i>richardsoni</i>	<i>Salix</i> <i>melanopsis</i>	<i>Dicamptodon</i> spp.	<i>Pinus</i> <i>albicaulis</i>	<i>Tsuga</i> <i>heterophylla</i>	<i>Thuja</i> <i>plicata</i>
Annual mean temperature	0.943	0.966	0.938	0.942	0.913	0.944*	0.900	0.917
Mean diurnal temperature range	0.915	0.941	0.824	0.849	0.856	0.833	0.873	0.826
Isothermality	0.926	0.953	0.965*	0.914	0.899	0.933	0.888	0.889
Temperature seasonality	0.946	0.966	0.951	0.959*	0.882	0.931	0.914	0.929*
Maximum temperature of warmest month	0.923	0.942	0.916	0.841	0.851	0.880	0.853	0.782
Minimum temperature of coldest month	0.945	0.965	0.932	0.942	0.902	0.929	0.908	0.870
Annual temperature range	0.901	0.931	0.884	0.887	0.839	0.835	0.855	0.892
Mean temperature of wettest quarter	0.951	0.974*	0.938	0.941	0.911	0.934	0.765	0.707
Mean temperature of driest quarter	0.927	0.952	0.902	0.850	0.884	0.892	0.928*	0.872
Mean temperature of warmest quarter	0.933	0.964	0.897	0.921	0.880	0.930	0.817	0.811
Mean temperature of coldest quarter	0.954*	0.940	0.937	0.843	0.917*	0.907	0.844	0.828
Annual precipitation	0.901	0.973*	0.811	0.943*	0.879	0.936*	0.918*	0.935*
Precipitation of wettest month	0.940	0.945	0.876	0.819	0.909	0.815	0.882	0.912
Precipitation of driest month	0.915	0.955	0.893	0.894	0.897	0.867	0.887	0.882
Precipitation seasonality	0.913	0.940	0.873	0.914	0.889	0.850	0.866	0.815
Precipitation of wettest quarter	0.938	0.936	0.886	0.860	0.900	0.754	0.898	0.904
Precipitation of driest quarter	0.902	0.963	0.900	0.860	0.880	0.850	0.871	0.760
Precipitation of warmest quarter	0.903	0.941	0.913	0.886	0.883	0.864	0.908	0.923
Precipitation of coldest quarter	0.949*	0.941	0.959*	0.870	0.917*	0.846	0.901	0.912
Test AUC (all 19 variables)	0.999	0.999	0.999	0.998	0.999	0.999	0.999	0.999

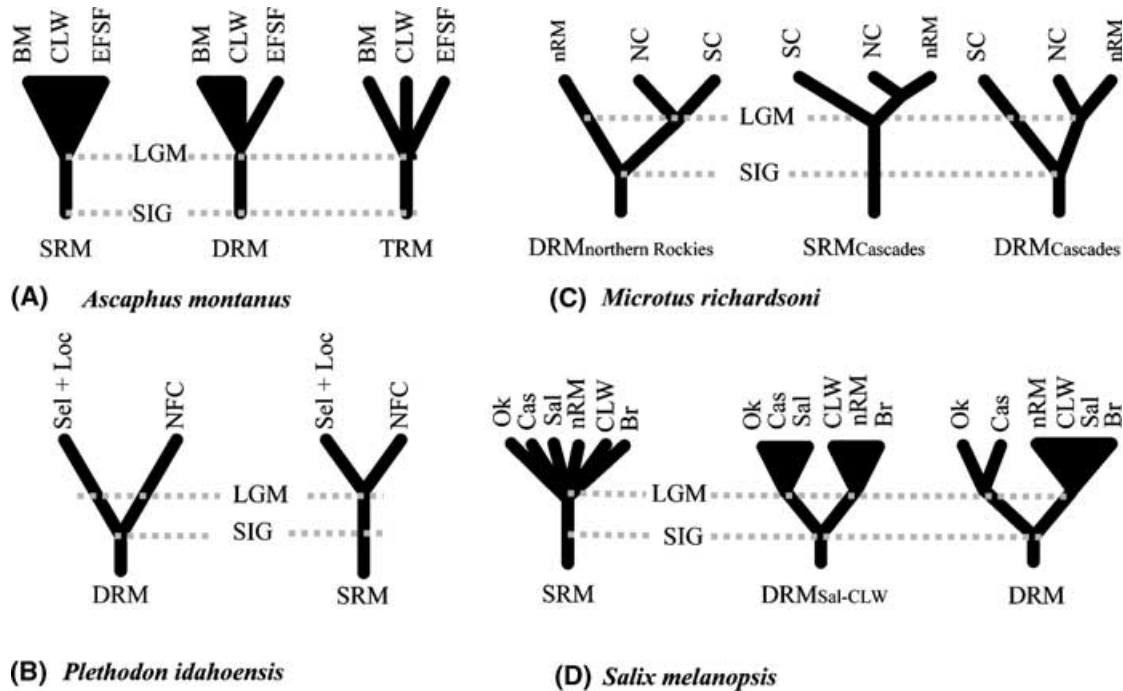
prediction of the climatic suitability of specific geographic locations for species occurrence and reports values ranging from 0 to 100. The climatic variables with the greatest predictive power varied among species (Table 3), as might be expected among organisms that are both ecologically and taxonomically distinct. The ecological niche models differed slightly across species, but each showed regions of highly suitable habitat in both the Cascades and northern Rocky Mountains under current climate conditions. Considering that species were selected for this study on the basis of their current mesic forest disjunct distribution, the similarities in the models for current conditions was expected. In contrast to the current models, the paleodistribution models differed substantially across species (Fig. 3). In some lineages, such as *Ascapthus* and (to a lesser extent) *Plethodon*, the habitat predicted as suitable during the last glacial maximum is similar to the current distribution, showing habitat in both the Cascades and northern Rocky Mountains. In other lineages, the paleodistribution models do not

predict suitable habitat in both regions. In *M. richardsoni*, the paleodistribution model is dominated by coastal and Cascades habitat, which is consistent with expectations derived from the extent of glaciation (see Brunsfeld et al. 2001). However, in *S. melanopsis* the paleodistribution models suggest that there was little suitable habitat along the Pacific coast; these models are dominated by the northern Rocky Mountains and central Oregon highlands regions. In each paleodistribution model, the Columbia basin is a region that is not predicted to contain suitable habitat, supporting notions that it has historically been an effective barrier to dispersal. When all paleodistribution models are considered, suitable habitat corridors for dispersal between the Cascades and northern Rocky Mountains would have existed for *S. melanopsis* and *Ascapthus* spp., but not for *M. richardsoni* or *Plethodon* spp. However, for *S. melanopsis*, the climate model does not clearly indicate the Cascades as a probable refuge, whereas much of the northern and central Rocky Mountains appear to have been suitable.



**Figure 3.** Ecological niche and paleodistribution models for four PNW-MF lineages. Models are shown for *A. truei/A. montanus* (A), *P. idahoensis/P. vandykei* (B), *M. richardsoni* (C), and *S. melanopsis* (D). The strength of prediction for each paleodistribution is indicated according to the key shown. For each species, the top map represents the paleodistribution model (21,000 ybp) and the lower map represents the ecological niche model of the current distribution. The maximum extent of the Cordilleran glacier (e.g., Booth et al. 2004) is shown with the dotted line and colored light blue.





**Figure 4.** Population models for (A) *A. montanus*, (B) *P. idahoensis*, (C) *M. richardsoni*, and (D) *S. melanopsis*. Dotted lines correspond to the last glacial maximum (21,000 ybp) and the end of the Sangamon interglacial (114,000 ybp). Abbreviations are as follows: North Cascades (NC), South Cascades (SC), Cascades (Cas), northern Rocky Mountains (nRM), Clearwater (CLW), Blue Mountains (BM), East Fork of the South Fork of the Salmon (EFSF), Selway/Lochsa (Sel+Loc), North Fork of the Clearwater (NFC), Okanogan Highlands (Ok), Salmon River (Sal), Bitterroot Mountains (Br).

Taken as a whole, the paleodistributional models suggest that the distributions of these currently codistributed organisms may not have shifted in a congruent manner in response to Pleistocene climatic change. The differences in paleodistribution models among organisms may also be related to the relative importance of different bioclimatic variables across organisms (Table 3).

**COALESCENT MODELS OF POPULATION STRUCTURE**

*A. montanus*

The paleodistribution model for *A. montanus* suggests that there were three regions within the northern Rocky Mountains that contain the most suitable habitat (Fig. 3): the Blue Mountains, the East Fork of the South Fork of the Salmon River, and the Clearwater

**Table 4.** Results of coalescent tests of Pleistocene population structure. Shown (left to right) are the taxon, the hypothesis tested and test statistic used, and the *P*-values at the low, point, and high estimates of the effective population size.

Taxa	<i>H<sub>o</sub></i>	Test stat	<i>N<sub>el</sub>ow</i>	<i>N<sub>e</sub></i>	<i>N<sub>eh</sub>igh</i>
<i>A. montanus</i>	SRM	RECIPROCAL MONO ((CLW+BM)(EFSF))	<0.01	<0.001	<0.01
	DRM	RECIPROCAL MONO ((CLW+BM)(EFSF))	0.26	0.156	0.09
	TRM	RECIPROCAL MONO ((CLW)(BM)(EFSF))	<0.01	<0.001	<0.01
<i>P. idahoensis</i>	SRM	RECIPROCAL MONO ((SEL+LOC)(NFC))	<0.01	<0.001	<0.01
	SRM <sub>Bottleneck</sub>	RECIPROCAL MONO ((SEL+LOC)(NFC))	<0.01	<0.001	<0.01
	DRM	RECIPROCAL MONO ((SEL+LOC)(NFC))	<0.01	<0.001	<0.01
	DRM <sub>Bottleneck</sub>	RECIPROCAL MONO ((SEL+LOC)(NFC))	<0.01	<0.001	<0.01
<i>M. richardsoni</i>	SRM <sub>Cascades</sub>	RECIPROCAL MONO ((SC)(NRM+NC))	0.61	0.390	0.29
		DIST = 0.02856 subs./site	0.08	0.113	0.13
	DRM <sub>Cascades</sub>	RECIPROCAL MONO ((SC)(NRM+NC))	1.0	1.0	1.0
		DIST = 0.02856 subs./site	<0.01	<0.01	<0.01
<i>S. melanopsis</i>	DRM <sub>northernRockies</sub>	RECIPROCAL MONO ((SC)(NRM+NC))	<0.01	<0.01	<0.01
	DRM	DIST = 0.00239 subs./site	<0.01	<0.01	<0.01
	DRM <sub>Sal-CLW</sub>	DIST = 0.00239 subs./site	0.01	0.024	0.01
	SRM	DIST = 0.00239 subs./site	0.33	0.651	0.52

River system. The later two of these regions have been proposed as sites of possible Pleistocene refugia in *A. montanus* (Nielson et al. 2001), and a DRM (Fig. 4A) in which populations in the East Fork of the South Fork of the Salmon River were isolated from those in the Clearwater River–Blue Mountains during the Wisconsinian glaciation could not be rejected (Table 4). However, a model of a single refugium, as well as a model with three refugia (e.g., a refugium located in each of the regions with predicted habitat suitability > 60), could be rejected across a range of  $N_e$ . This suggests that populations of *A. montanus* were isolated in two glacial refugia during the Wisconsinian glaciation. Another interesting aspect of the paleodistribution model is the suggestion that there were regions of suitable habitat in the central Oregon highlands. Although gene flow during the last glacial maximum has been previously suggested (Pauken and Metter 1971), there is no genetic evidence in support of gene flow between the Cascades and northern Rocky Mountains (Daugherty 1979; Nielson et al. 2001; Carstens et al. 2005b).

#### *P. idahoensis*

Only a small area of habitat is projected as suitable for the plethodontid salamanders during the last glacial maximum, and there is little indication of dispersal corridors between the Cascades and northern Rocky Mountains (Fig. 3). This finding is consistent with the predictions of the ancient vicariance hypothesis and with the deep genetic divergence between the coastal and inland species. In *P. idahoensis*, two discrete regions within the Clearwater River drainage may have contained the best habitat, and a dual refugium model was designed based on this prediction (Fig. 4B). Because there may not have been a dramatic difference in habitat quality throughout the Clearwater region, a SRM was also tested. Each of these models was rejected across the range of  $N_e$  values, both with and without population bottlenecks (Table 4). In *P. idahoensis*, where the  $N_e$  is large (~200,000) in comparison to the modeled temporal divergence between populations (e.g., < 30,000 generations, or 0.15N), these results are consistent with theoretical expectations (Rosenberg 2003; Hudson and Turelli 2003), which predict that it takes on the order of 4N generations for reciprocal monophyly to form. These results indicate that northern Rocky Mountain populations of *P. idahoensis* have been isolated from each other for a significant period of time.

#### *M. richardsoni*

The paleodistribution model for *M. richardsoni* (Fig. 3) suggests that the vast majority of habitat was located in the Cascades during the last glacial maximum, with a far smaller and less suitable region in the northern Rocky Mountains. A model that posits the persistence of a northern Rocky Mountain refugium throughout the last glaciation, in addition to a refugium in the Cascades (Fig. 4C),

can be rejected (Table 4). Two other models were tested which propose that extant populations in the northern Rocky Mountains are the result of recent inland dispersal from the Cascades as the Cordilleran glacier retreated (Fig. 4C). The SRM in the Cascades posits that populations in the North Cascades, South Cascades, and northern Rocky Mountains were derived from a single population that began to expand at the end of the Pleistocene. However, because haplotypes from the South Cascades and North Cascades + northern Rocky Mountains are reciprocally monophyletic, an additional DRM was tested. It posits that populations from the South Cascades and North Cascades were isolated during the late Pleistocene, and that northern Rocky Mountains populations were derived from the North Cascades. The DRM can be rejected on the basis of the amount of genetic divergence between the South Cascades and North Cascades + northern Rocky Mountain populations (Table 4). This suggests that *M. richardsoni* persisted as a single population located in the Cascades during the Wisconsinian glaciation, and began to diverge as the glaciers retreated.

#### *S. melanopsis*

The paleodistribution model suggests that a large area of contiguous habitat suitable for *S. melanopsis* existed at the last glacial maximum from the central Oregon highlands through central Idaho and to the east (Fig. 3). Given this large region, we tested and could not reject a single-refugium model (Fig. 4) proposing that all extant populations are recently derived from the same source population (Table 4). The paleodistribution model suggested that a region of habitat in southern Oregon was largely isolated from other suitable habitat, so a dual refugium model (Fig. 4) was also tested in *S. melanopsis*. Furthermore, Brunfeld et al. (2006) proposed a vicariant divergence within the NRM, suggesting that populations of *S. melanopsis* persisted throughout the late Pleistocene in two refugia, located in the Clearwater and Salmon river canyons. Both of these hypotheses were rejected (Table 4).

## Discussion

Determining if codistributed species responded to climate change in an independent or concerted manner (e.g., Sullivan et al. 2000) is a basic objective of comparative phylogeography. By quantifying habitat suitability during the last glacial maximum, paleodistributional modeling provides an estimate of the historical distribution that can be used to formulate explicit predictions about the locations of dispersal corridors and unsuitable habitat. In this study, such models serve as the basis for a predictive, hypothesis-testing framework by suggesting alternative population histories (e.g., Fig. 4). This predictive framework improves the rigor of phylogeographic investigations by providing an

inference approach that considers the stochastic variance associated with genetic processes (Knowles and Maddison 2002).

Paleoclimatic modeling augments previous inferences about the evolutionary history of the organisms that inhabit the PNW-MF ecosystem. Phylogeographic hypothesis testing suggested that three amphibian lineages (*A. truei*/*A. montanus*, *Dicamptodon* spp., and *P. vandykei*/*P. idahoensis*) have similar evolutionary histories (Carstens et al. 2005b). In each lineage, populations were isolated in the Cascades and northern Rocky Mountains during the Pleistocene. The paleodistribution models for *Ascaphus* and *Plethodon* show regions of suitable habitat in both the Cascades and northern Rocky Mountains (Fig. 3), as does the paleodistribution model for *Dicamptodon* (see online Supplementary Figure S1a). Taken in total, the amphibian lineages form an evolutionary cohort; exhibiting broad similarities in their population genetic structure likely due to their concerted response to environmental change. Similarly, in the ecosystem dominants *T. heterophylla* and *Thu. plicata*, paleodistributional models (see online Supplementary Figure S1b, c) suggest that two disjunct regions contained the most suitable habitat, one in the Cascades and one in the northern Rocky Mountains. For these lineages, the models of past habitat suitability are largely congruent with the inference that populations persisted throughout the Pleistocene in multiple glacial refugia (Brunsfeld and Sullivan 2005), with only *D. aterrimus* apparently confined to a single refugium (see Carstens et al. 2005a). Models of two refugia cannot be rejected in *A. montanus* and *P. idahoensis*, and genetic data for these species as well as the codistributed plant *Cardamine constancei* (Brunsfeld and Sullivan 2005) suggest structured refugia. However, all species did not respond uniformly to the geological and climatic forces that shaped the ecosystem (Carstens et al. 2005b). The integration of paleodistributional models with coalescent tests of phylogeographic hypotheses provides evidence that may explain these independent responses.

Unlike the amphibian lineages, there is genetic evidence of gene flow between populations from the northern Rocky Mountains and North Cascades in both *S. melanopsis* and *M. richardsoni* (Carstens et al. 2005b). Gene flow is suggested by the reticulation of haplotypes from North Cascades and northern Rocky Mountain populations, along with the shallow temporal divergence and high rates of gene flow estimated using the program MDIV (Nielsen and Wakeley 2001). However, without the spatial-geographic information provided by the paleodistributional modeling, it is difficult to understand the historical processes that produced these patterns. Before the genetic data were collected, it was expected that the Cascades were the source population because there was a much greater extent of unglaciated habitat in this region during the Pleistocene (Brunsfeld et al. 2001). The basal position of haplotypes from portions of the northern Rocky Mountains in

both *S. melanopsis* and *M. richardsoni* contradicted this expectation, suggesting that populations in the North Cascades were instead derived from populations in the northern Rockies. Furthermore, portions of the Pacific Northwest were proposed as a potential suture zone by Remington (1968), a proposal which was supported by recent analysis (Swenson and Howard 2005), and in *S. melanopsis* and *M. richardsoni* the possibility that populations in these regions were isolated during interglacials and came into secondary contact during full glacials is not incompatible with the genetic data. Absent the paleodistribution models, it is difficult to determine which of these scenarios is most probable.

In both *S. melanopsis* and *M. richardsoni*, population divergence between the northern Rocky Mountains and the Cascades dates to the late Pleistocene (Carstens et al. 2005b). The paleodistributional models for these species are markedly different (Fig. 3), with the northern Rocky Mountains containing most of the suitable habitat for *S. melanopsis* and the Cascades containing most of the suitable habitat for *M. richardsoni*. Although these models only represent a single point in time, they cast doubt on any notion that these species are members of the same evolutionary cohort. For each species, coalescent models demonstrate that the sampled empirical data could have been produced by post Pleistocene expansion from a single refugium, but the location of these refugia may have differed. The *M. richardsoni* paleodistributional model implies that its refuge would have been in the Cascades, whereas the *S. melanopsis* refugium was likely located in the northern Rocky Mountains. Although the genetic pattern of these species is similar, the paleodistributional models suggest that this similarity was produced by very different ancestral distributions. Nevertheless, it is clear that the last glacial maximum was a period of population expansion and gene flow. Similarly, the paleodistribution model for *Pin. albicaulis* (see online Supplementary Figure S1d) suggests abundant opportunity for gene flow between the northern Rocky Mountains and Cascades at this time, which is congruent with genetic evidence (Richardson et al. 2002).

The dramatic difference between the current and paleodistribution models for some lineages support suggestions that species' ranges can change dramatically in size during times of climate change (Lessa et al. 2003; Lyons 2003; Myers et al. 2005). For example, *M. richardsoni* has apparently expanded its range over 1000 km (from the North Cascades through British Columbia down the Rocky Mountains into Utah and Wyoming) since the last glacial maximum. During the same period, *S. melanopsis* colonized new habitat in the Cascades and recently deglaciated regions in British Columbia, probably from several source populations (Brunsfeld et al. 2006). Coalescent modeling demonstrates that the range expansions implied by the projected past habitat are

consistent with the genetic data, and that species distributions in the PNW-MF have shifted dramatically, and to a large degree independently, since the last glacial maximum.

Of the paleodistribution models presented here, for only one case (*Ascaphus*) is the paleodistribution seemingly incompatible with the genetic data. This model (Fig. 3) indicates the potential for gene flow between the Cascades and northern Rocky Mountains during the last glacial maximum. Several factors may explain this incongruence: (1) paleodistributional models assume that species physiological limits are constant over time, which may not be the case (Hadly et al. 1998; Davis and Shaw 2001). Violating this assumption would produce an inaccurate estimate of the past distribution. (2) There may be biases in terms of where the samples are collected, for example, if certain types of habitat are more accessible to researchers (Graham et al. 2004b). (3) The numbers of sampled localities are also important in generating an accurate projection. However, the models presented here use between 400 (*Ascaphus*) and 40 (*Salix*) collection localities, and ecological niche modeling has been shown to be accurate so long as at least 20 localities are used to construct the model (Stockwell and Peterson 2002). (4) The ecological niche model may be overpredicted due to inadequate or biased sampling (Peterson and Vieglais 2001; Anderson et al. 2003). Alternatively, the overprediction may indicate regions of suitable habitat that are not occupied due to extrinsic factors such as barriers to dispersal or interspecific competition (Anderson et al. 2002; Graham et al. 2004a,b; Araújo and Guisan 2006). Overprediction in ecological niche models may lead to overprediction in paleodistribution models as well. Apparent inconsistencies (such as between the *Ascaphus* genetic data and past habitat model) merit further investigation and have the potential to direct researchers towards new questions. In the case of *Ascaphus*, the deep divergence between inland and coastal species could indicate dispersal limitations, or if range expansion throughout the Pleistocene glacial maxima brought coastal and inland *Ascaphus* into secondary contact, mechanisms may have evolved to reinforce reproductive isolation.

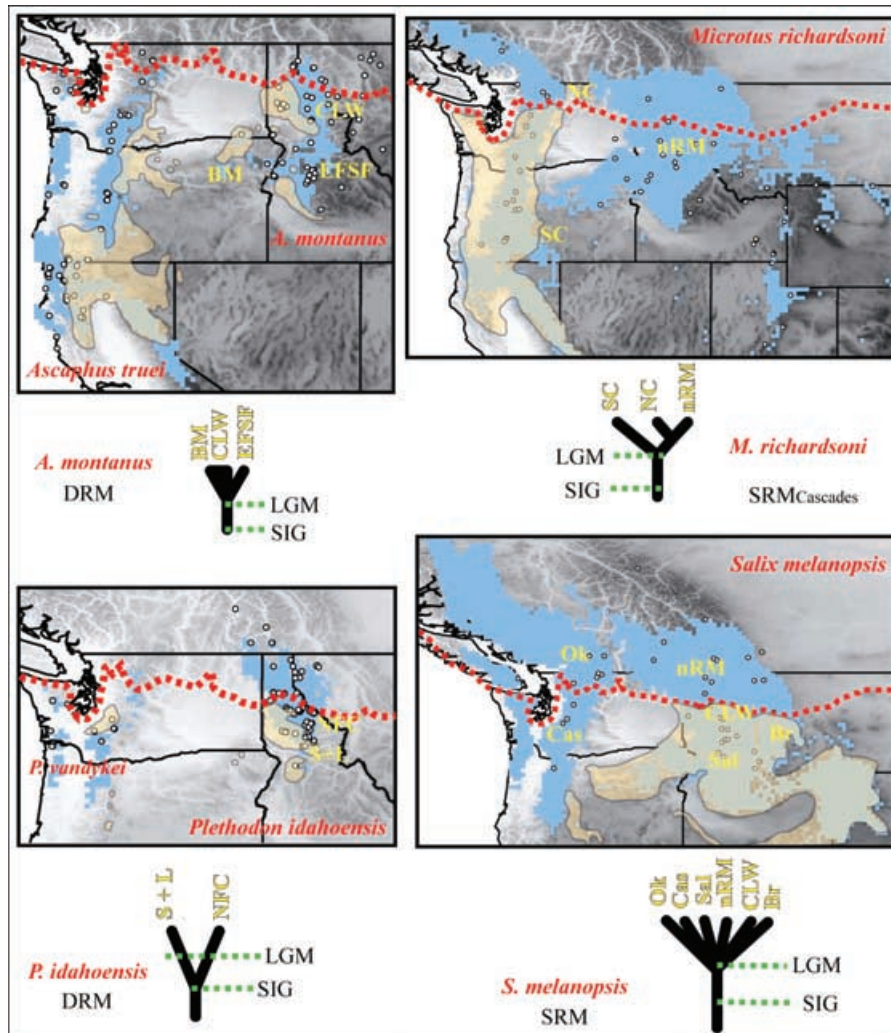
The development of paleodistribution models from additional points in the past would improve the integrative approach taken here. For example, one key prediction of the ancient vicariance hypothesis is that habitat would have been severely restricted during glacial maxima. Although the paleodistribution model from the last glacial maximum provides one test of this hypothesis, the temporal divergence between taxa such as *A. truei* and *A. montanus* indicates that these species have persisted throughout many glacial maxima, and conditions may well have been different for each. Paleoclimate models from additional points during the Pleistocene, when they become available, will allow a comparison of the relative impact of the Wisconsinian glaciation versus previous glacials on population genetic structure.

## ECOLOGICAL NICHE MODELING AND COMPARATIVE PHYLOGEOGRAPHY

Phylogeographic research can be characterized on the basis of how its hypotheses are treated. Traditionally, phylogeography has been used to formulate or describe hypotheses (e.g., Avise 2000); a tradition perhaps derived from the cladistic interpretation of phylogeny as a hypothesis of relationships. Statistical phylogeography is influenced by coalescent-based population genetics and seeks to test a priori hypotheses using methods that account for the stochastic variance inherent to genetic data (Knowles and Maddison 2002). Much of the stimulus for statistical phylogeography results from the realization that phylogeographic inferences based on qualitative interpretations of the relationship between genealogy and geographic distribution have the potential for over interpretation (Edwards and Beerli 2000; Hudson and Turelli 2003; Wakeley 2003). The testing of a priori hypotheses accounts for the stochastic variance inherent to the process of gene coalescence. However, the effectiveness of this approach depends on the ability of the models to capture the important aspects of biological history (Knowles 2004).

Knowles (2004) described three key components of experimental design in statistical phylogeographic studies: (1) defining a set of hypotheses; (2) deciding on the model complexity of the hypotheses; and (3) integrating information from external data, such as bioclimatic or paleoecological data. Our study demonstrates how paleoclimatic modeling can contribute to the formulation of alternative hypotheses, by identifying putative sites for refugia or routes for dispersal. We also show how the amount of suitable habitat predicted by paleodistribution models can guide decisions about model complexity. For example, a model of population history that incorporates a bottleneck during the last glacial maximum is more appropriate in *P. idahoensis* than in *S. melanopsis* (Fig. 5) because of the differences in the size of the projected habitat. By providing an empirical framework for the generation and testing of population divergence models' statistical phylogeography, models of the distribution of past habitat allow phylogeographers to better understand the factors that promote population divergence. When conducted on a regional scale, this may allow a better understanding of how population processes contribute to regional patterns of biodiversity.

Because species histories are often incredibly complex (Knowles 2004), paleodistribution models may be particularly useful to comparative studies, in which they provide a common framework for the formulation of hypotheses, and supply qualitative evidence as to the degree of codistribution among organisms, either presently or at some point in the past. The advantages of statistical phylogeography are also most pronounced in comparative studies, in which general regional hypotheses provide a metric for the comparison of organisms with differing life-history traits (Carstens et al. 2005b). The best application of these models may



**Figure 5.** Paleodistributional models and current distributions are shown for *A. truei* and *A. montanus*, *P. vandykei*/*P. idahoensis*, *S. melanopsis*, and *M. richardsoni*. Open circles represent actual sampling localities. Predicted current ecological niches for each species are shown shaded in blue. Regions of predicted habitat during the LGM are shown in transparent yellow. Regions with predicted suitability > 65 are shown for *Ascaphus*, > 30 are shown for *Plethodon*, > 45 are shown for *S. melanopsis*, and > 99 are shown for *M. richardsoni*. The maximum extent of the Cordilleran glacier is shown with a dotted red line. Under each map is a representation of the population model that is most consistent with the genetic data. These models, and the abbreviations of the terminal branches, correspond to those in Figure 4.

be to the recent past, and as such they complement approaches that integrate phylogenetics and geography and are concerned with more ancient events (e.g., Losos and Glor 2003; Ree et al. 2005; Yesson and Culham 2006).

Comparative phylogeography is a synthetic discipline (Arbogast and Kenagy 2001) that aims to provide an understanding of the processes that generate regional patterns of biodiversity (Bermingham and Moritz 1998). Integrating phylogeographic and ecological niche modeling may lead to new insights into the processes that promote speciation, particularly by helping to elucidate the role of local adaptation in allopatric divergence (Weins 2004). Phylogeography is entering an era in which studies use multilo-

cus nuclear data (Zhang and Hewitt 2003). These studies necessarily employ statistical phylogeographic approaches, because inferences derived from a visual interpretation of the genealogies cannot be made when there is conflict among the genealogies sampled from multiple loci (e.g., Dolman and Moritz 2006; Carstens and Knowles 2007). When multilocus nuclear data are used, models of the historical divergence of populations are vital to phylogeographic investigations because they provide a framework for hypothesis testing and structured-coalescent models (Knowles and Carstens 2007). Estimates of the paleodistribution add a valuable spatial-geographic component to these models, improving phylogeographic inferences.

## ACKNOWLEDGMENTS

We thank D. Alvarado-Serrano, J. Degenhardt, P. Fine, and A. Stevenson for commenting on various drafts of this manuscript, the Laboratoire des Sciences du Climat et de l'Environnement (LSCE) and the PMIP2 modeling (<http://www-lsce.cea.fr/pmip2/>) group for providing paleoclimatic data, and Robert Hijmans for processing these data. Training in ecological niche modeling was made possible by funding from the University of Michigan's Rackham Graduate School. We thank D. Dyer at MONTU for providing collection localities for *Thu. plicata*. We thank K. Crandall, M. Hedin, and one anonymous reviewer for comments that improved the quality of this manuscript.

## LITERATURE CITED

- Anderson, R. P., A. T. Peterson, and M. Gómez-Laverde. 2002. Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. *Oikos* 98: 3–16.
- Anderson, R. P., D. Lew, and A. T. Peterson. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecol. Model.* 162: 211–232.
- Arbogast, B. S., S. V. Edwards, J. Wakeley, P. Beerli, and J. B. Slowinski. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annu. Rev. Ecol. Syst.* 33:707–740.
- Arbogast, B. S. and G. J. Kenagy. 2001. Comparative phylogeography as an integrative approach to historical biogeography. *J. Biogeogr.* 28:819–825.
- Araújo, M. B., and A. Guisan. 2006. Five (or so) challenges for species distribution modeling. *J. Biogeogr.* 33: 1677–1688.
- Austin, M. P. 1985. Continuum concept, ordination methods and niche theory. *Annu. Rev. Ecol. Syst.* 16:39–61.
- Avise, J. 2000. *Phylogeography: the history and formation of species*. Harvard Univ. Press, Cambridge, MA.
- Bermingham, E. and C. Moritz. 1998. Comparative phylogeography: concepts and applications. *Mol. Ecol.* 7:367–369.
- Bond, J. E., D. A. Beamer, T. Lamb, and M. Hedin. 2006. Combining genetic and geospatial analyses to infer population extinction in mygalomorph spiders endemic to the Los Angeles region. *Anim. Conser.* 9:145–157.
- Booth, D. B., K. G. Troost, J. J. Clague, and R. B. Waitt. 2004. The Cordilleran Ice Sheet. Pp. 17–43 in: A. R. Gillespie, S. C. Porter, and B. F. Atwater, eds. *The Quaternary period in the United States*. Elsevier Press, New York.
- Brunsfeld, S. J., J. Sullivan, D. E. Soltis, and P. S. Soltis. 2001. Comparative phylogeography of northwestern North America: a synthesis. Pp. 319–339 in J. Silvertown and J. Antonovics, eds. *Integrating ecology and evolution in a spatial context*. Blackwell Publishing, Williston, VT.
- Brunsfeld, S. J. and J. Sullivan. 2005. A multi-compartmented glacial refugium in the northern Rocky Mountains: evidence from the phylogeography of *Cardamine constancei* (Brassicaceae). *Conser. Genet.* 6:895–904.
- Brunsfeld, S. J., T. R. Miller, and B. C. Carstens. 2006. Insights into the biogeography of the Pacific Northwest of North America: evidence from the phylogeography of *Salix melanopsis* (Salicaceae). *Syst. Bot.* 32:129–139.
- Carstens, B. C., and L. L. Knowles. 2007. Shifting distributions and speciation: genomic resolution of species divergence during rapid climate change. *Mol. Ecol.* 16:619–627.
- 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. *Syst. Biol.* 53:781–792.
- Carstens, B. C., J. D. Degenhardt, A. L. Stevenson, and J. Sullivan. 2005a. Accounting for coalescent stochasticity in testing phylogeographic hypotheses: modelling Pleistocene population structure in the Idaho Giant Salamander *Dicamptodon aterrimus*. *Mol. Ecol.* 14:255–265.
- Carstens, B. C., S. J. Brunsfeld, J. R. Demboski, J. D. Good, and J. Sullivan. 2005b. Investigating the evolutionary history of the Pacific Northwest mesic forest ecosystem: hypothesis testing within a comparative phylogeographic framework. *Evolution* 59:1639–1652.
- Collins, W. D., et al. 2006. The community climate system model version 3 (CCSM3). *J. Clim.* 19:2122–2143.
- Daugherty, C. H. 1979. Population ecology and genetics of *Ascaphus truei*: an examination of gene flow and natural selection. Ph. D. diss., University of Montana, Missoula, MT.
- Daubenmire, R. 1952. Plant geography of Idaho. Pp. 1–17 in: R. J. Davis, ed. *Flora of Idaho*. Brigham Young University Press, Provo, UT.
- Daubenmire, R. and J. B. Daubenmire. 1968. *Forest vegetation of Eastern Washington and Northern Idaho*. Tech. Bull. 60, Washington Agricultural Experiment Station, Washington State University, Pullman, WA.
- Daubenmire, R. 1975. *Floristic plant geography of eastern Washington and northern Idaho*. Brigham Young University Press, Provo, UT.
- Davis, M. B. and R. G. Shaw. 2001. Range shifts and adaptive responses to quaternary climate change. *Science* 292:673–679.
- DeChaine, E. G., and A. P. Martin. 2005. Historical biogeography of two alpine butterflies in the Rocky Mountains: broad-scale concordance and local-scale discordance. *J. Biogeogr.* 32:1943–1956.
- Delcourt, P. A. and H. R. Delcourt. 1993. Paleoclimates, paleovegetation, and paleofloras during the late Quaternary. Pp. 71–94 in Editorial Committee. *Flora of North America*, Vol. 1, 1st edn. Oxford Univ. Press, New York.
- Detling, L. E. 1968. Historical background of the flora of the Pacific Northwest. Bull. No. 13, Mus. Nat. Hist., Univ. Oregon, Eugene, OR.
- Dolman, G., and C. Moritz. 2006. A multilocus perspective on refugial isolation and divergence in rainforest skinks (*Carlia*). *Evolution* 60:573–582.
- Edwards, S. V., and P. Beerli. 2000. Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* 54:1839–1854.
- English, J. M., and S. T. Johnston. 2004. The Laramide Orogeny: what were the driving forces? *Int. Geol. Rev.* 46: 833–883.
- Gaston, K. J. 1996. Species-range-size distributions: patterns, mechanisms, and implications. *Trends Ecol. Evol.* 11:197–201.
- Gates, D. M. 1993. *Climate change and its biological consequences*. Sinauer and Associates Press, Sunderland, MA.
- Gibbard, P. and T. Van Kolfshoten. 2004. The Pleistocene and Holocene epochs. Pp. 441–452 in F. M. Gradstein, J. G. Ogg, and A. G. Smith, eds. *A geologic time scale*. Cambridge Univ. Press, London.
- Graham, A. 1993. History of the vegetation: cretaceous-tertiary. Pp. 57–70 in *Flora of North America*, Vol. I. Oxford Univ. Press, New York.
- Graham, A. 1999. *Late Cretaceous and Cenozoic History of North American Vegetation*. Oxford Univ. Press, New York.
- Graham, C. H., S. Ron, J. C. Santos, C. J. Schneider, and C. Moritz. 2004a. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in Dendrobatic frogs. *Evolution* 58:1781–1793.
- Graham, C. H., S. Ferrier, F. Huettman, C. Moritz, and A. T. Peterson. 2004b. New developments in museum-based informatics and applications in biodiversity analysis. *Trends Ecol. Evol.* 19:497–503.
- Hadly, E. A., M. H. Kohn, J. A. Leonard, and R. K. Wayne. 1998. A genetic record of population isolation in pocket gophers during Holocene climatic change. *Proc. Natl. Acad. Sci. USA* 95:6893–6896.
- Hewitt, G. M. 1996. Some genetic consequences of the ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* 58:247–276.

- 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. *Int. J. Climatol.* 25:1965–1978.
- Hudson, R. R. 1992. Gene trees, species trees and the segregation of ancestral alleles. *Genetics* 131:509–512.
- Hudson, R. R., and J. A. Coyne. 2002. Mathematical consequences of the genealogical species concept. *Evolution* 56:1557–1565.
- Hudson, R. R., and M. Turelli. 2003. Stochasticity overrules the three-times rule: genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution* 57:182–190.
- Hugall, A., C. Moritz, A. Moussalli, and J. Stanisic. 2002. Reconciling paleodistribution models and comparative phylogeography in the Wet Tropics rainforest land snail *Arosophia bellendenkerensis* (Brazier 1875). *Proc. Natl. Acad. Sci. USA* 99:6112–6117.
- Kidd, D. M., and M. G. Ritchie. 2006. Phylogeographic information systems: putting the geography into phylogeography. *J. Biogeogr.* 33:1851–1865.
- Kingman, J. F. C. 1982. The Coalescent. *Stochastic process. Appl.* 13:235–248.
- Knowles, L. L. 2001. Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Mol. Ecol.* 10:691–701.
- Knowles, L. L., and W. P. Maddison. 2002. Statistical phylogeography. *Mol. Ecol.* 11:2623–2635.
- Knowles, L. L. 2004. The burgeoning field of statistical phylogeography. *J. Evol. Biol.* 17:1–10.
- Knowles, L. L., and B. C. Carstens. 2007. Inferring a population-divergence model for statistical phylogeographic tests in montane grasshoppers. *Evolution*, 61:477–493.
- Kuhner, M. K., J. Yamato, P. Beerli, L. P. Smith, E. Rynes, E. Walkup, C. Li, J. Sloan, P. Colacurcio, and J. Felsenstein. 2005. LAMARC v. 2.0. at <http://evolution.gs.washington.edu/lamarc.html>
- Lawton, J. H. 1993. Range, population abundance and conservation. *Trends Ecol. Evol.* 8:409–413.
- Lessa, E. P., J. A. Cook, and J. L. Patton. 2003. Genetic footprints of demographic expansion in North America, but not Amazonia, during the Late Quaternary. *Proc. Natl. Acad. Sci. USA* 100:10331–10334.
- Liu, C., P. M. Berry, P. D. Dawson, and R. G. Pearson. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28:385–393.
- Losos, J. B., and R. E. Glor. 2003. Phylogenetic comparative methods and the geography of speciation. *Trends Ecol. Evol.* 18:220–227.
- Lyons, S. K. 2003. A quantitative assessment of the range shifts of Pleistocene mammals. *J. Mammal.* 84:385–402.
- Maddison, W. P., and D. R. Maddison. 2004. MESQUITE: a modular system for evolutionary analysis. Version 1.01. Available at <http://mesquiteproject.org>
- Martinez-Meyer, E., A. T. Peterson, and W. W. Hargrove. 2004. Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity. *Glob. Ecol. Biogeogr.* 13:305–314.
- Mehring, Jr., P. E. 1985. Late-Quaternary pollen record from the interior Pacific Northwest and northern Great Basin of the United States. Pp. 167–189 in V. M. Bryant and R. G. Holloway Jr., eds. *Pollen records of Late-Quaternary North America Sediments*. Amer. Assoc. Strat. Palyn. Found., Dallas, TX.
- Minin, V., Z. Abdo, P. Joyce, and J. Sullivan. 2003. Performance-based selection of likelihood models for phylogeny estimation. *Syst. Biol.* 52:674–683.
- Myers, P., B. L. Lundrigan, and R. Van de Kopple. 2005. Climate change and the distribution of *Peromyscus* in Michigan: is global warming already having an impact? Pp. 101–125 in E. A. Lacey and P. Myers, eds. *Mammalian diversification: from chromosomes to phylogeography (A celebration of the career of James L. Patton)*. University of California Press, Berkeley, CA.
- Nielsen, R., and J. Wakeley. 2001. Distinguishing migration from isolation: a Markov-chain Monte Carlo approach. *Genetics* 158:885–896.
- Nielsen, M. K., 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.
- Pauken, R. J., and D. E. Metter. 1971. Geographic representation of morphological variation among populations of *Ascaphus truei* Stejneger. *Syst. Zool.* 20:431–441.
- Pamilo, P., and M. Nei. 1988. Relationships between gene trees and species trees. *Mol. Biol. Evol.* 5:568–583.
- Peterson, A. T. 2001. Predicting species geographic distributions based on ecological niche modeling. *Condor* 103:599–605.
- Peterson, A. T., and D. A. Vieglais. 2001. Predicting species invasions using ecological niche modeling: new approaches from bioinformatics attach a pressing problem. *BioScience* 51: 363–371.
- Peterson, A. T., E. Martinez-Meyer, and C. Gonzalez-Salazar. 2004. Reconstructing the Pleistocene geography of the *Aphelocoma* jays (Corvidae). *Divers. Distr.* 10:237–246.
- Pielou, E. C. 1991. *After the ice age: the return of life to glaciated North America*. Univ. of Chicago Press, Chicago.
- Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* 190:231–259.
- Ree, R. H., B. R. Moore, C. O. Webb, and M. J. Donoghue. 2005. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59:2299–2311.
- Remington, C. L. 1968. Suture-zones of hybrid interactions between recently-joined biotas. Pp. 321–428 in T. Dobzhansky, M. K. Hecht, and W. C. Steere, eds. *Evolutionary biology*. Appleton-Crofts, New York.
- Richardson, B. A., S. J. Brunfeldt, and N. B. Klopfenstein. 2002. DNA from bird-dispersed seed and wind-disseminated pollen provides insights into postglacial colonization and population genetic structure of whitebark pine (*Pinus albicaulis*). *Mol. Ecol.* 11:215–227.
- Rissler, L. J., R. J. Hijmans, C. H. Graham, C. Moritz, and D. B. Wake. 2006. Phylogeographic lineages and species comparisons in conservation analyses: a case study of the California herpetofauna. *Am. Nat.* 167:655–666.
- Rosenberg, N. A. 2003. The shapes of neutral gene genealogies in two species: probabilities of monophyly, paraphyly, and polyphyly in a coalescent model. *Evolution* 57:1465–1477.
- Ruegg, K. C., R. J. Hijmans, and C. Moritz. 2006. Climate change and the origin of migratory pathways in the Swainson's thrush, *Catharus ustulatus*. *J. Biogeogr.* 33:1172–1182.
- Russell, A. L., R. A. Medellín, and G. F. McCracken. 2005. Genetic variation in the Mexican free-tail bat (*Tadarida brasiliensis mexicana*). *Mol. Ecol.* 14:2207–2222.
- Schneider, C. J., M. Cunningham, and C. Moritz. 1998. Comparative phylogeography and the history of endemic vertebrates in the Wet Tropics rainforests of Australia. *Mol. Ecol.* 7:487–498.
- Steele, C. A., B. C. Carstens, A. Storfer, and J. Sullivan. 2005. Testing hypotheses of speciation timing in *Dicamptodon copei* and *Dicamptodon aterrimus* (Caudata: Dicamptodontidae). *Mol. Phylog. Evol.* 36:90–100.
- Steele, C. A., and A. Storfer. 2006. Coalescent-based hypothesis testing supports multiple Pleistocene refugia in the Pacific Northwest for the Pacific giant salamander (*Dicamptodon tenebrosus*). *Mol. Ecol.* 15:2477–2487.
- Stigall, A. L., and B. S. Lieberman. 2006. Quantitative paleobiogeography: GIS, phylogenetic biogeographical analysis, and conservation insights. *J. Biogeogr.* *In press*.

- Stockwell, D. R. B. and A. T. Peterson. 2002. Effects of sample size on accuracy of species distribution models. *Ecol. Model.* 148:1–13.
- Sullivan, J., E. Arellano, and D. S. Rogers. 2000. Comparative phylogeography of Mesoamerican highland rodents: concerted versus independent response to past climatic fluctuations. *Am. Nat.* 155:755–768.
- Swofford, D. L. 2002. PAUP\*. Phylogenetic Analysis Using Parsimony (and other methods). Version 4. Sinauer Assoc., Sunderland, MA.
- Swenson, N. G. 2006. GIS-based niche models reveal unifying climatic mechanisms that maintain the location of avian hybrid zones in a North American suture zone. *J. Evol. Biol.* 19:717–725.
- Swenson, N. G., and D. J. Howard. 2005. Clustering of contact zone, hybrid zones, and phylogeographic breaks in North America. *Am. Nat.* 166:581–591.
- Takahata, N. 1989. Gene genealogy in three related populations: consistency probability between gene and population trees. *Genetics* 122:957–966.
- Wakeley, J. 2003. Inferences about the structure and history of populations: coalescences and intraspecific phylogeography. Pp. 193–213 in R. Singh and M. Uyenoyama, eds. *The Evolution of Population Biology*. Cambridge Univ. Press, Cambridge, UK.
- Weins, J. J. 2004. What is speciation and how should we study it? *Am. Nat.* 163:914–923.
- Yesson, C., and A. Culham. 2006. Phyloclimatic modeling: combining phylogenetics and bioclimatic modeling. *Syst. Biol.* 55:785–802.
- Zhang, D.-X. and G. M. Hewitt. 2003. Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Mol. Ecol.* 12:563–584.

Associate Editor: K. Crandall

## Supplementary Material

The following supplementary material is available for this article:

**Figure S1.** Ecological niche and paleodistribution models for four PNW-MF lineages. Models are shown for *Dicamptodon* spp. (1a), *T. heterophylla* (1b), *Thu. plicata* (1c), and *Pin. albicaulis* (1d). The strength of prediction for each paleodistribution is indicated according to the key shown. For each species, the top map represents the paleodistribution model (21,000 ybp) and the lower map represents the ecological niche model of the current distribution. The maximum extent of the Cordilleran glacier (e.g., Booth et al. 2004) is shown with the dotted line and colored light blue.

This material is available as part of the online article from:

<http://www.blackwell-synergy.com/doi/abs/10.1111/j.1558-5646.2007.00117.x>

(This link will take you to the article abstract.)

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.