Molecular Ecology (2005) **14**, 255–265 doi: 10.1111/j.1365-294X.2004.02404.x

Accounting for coalescent stochasticity in testing phylogeographical hypotheses: modelling Pleistocene population structure in the Idaho giant salamander *Dicamptodon aterrimus*

B. C. CARSTENS,* J. D. DEGENHARDT,* A. L. STEVENSON* and J. SULLIVAN*†
University of Idaho, Department of Biological Sciences, Box 443051, Moscow, ID, 83844–3051, †Initiative in Bioinformatics and Evolutionary Studies (IBEST)

Abstract

Several theoretical studies have demonstrated the importance of accounting for coalescent stochasticity in phylogeographical studies, however, there are few empirical examples that do so in the context of explicit hypothesis testing. Here, we provide an example from the Idaho giant salamander (*Dicamptodon aterrimus*) using 118 mtDNA sequences, nearly 2 kb in length. This species is endemic to mesic forests in northern and central Idaho, and several a priori hypotheses have been erected based both on palaeoclimatic grounds and from phylogeographical studies of codistributed amphibians. Phylogenetic analysis of the *D. aterrimus* data suggests an expansion from a single refugium south of the Salmon River, whereas the inference from nested clade analysis is one of expansion from a single refugium in the Clearwater drainage. Explicit testing of these hypotheses, using geographically structured coalescent simulations to erect null distributions, indicates we can reject expansion from the Clearwater drainage (p_{CLW} = 0.089), but not expansion from the South Fork of the Salmon drainage (p_{SAL} = 0.329). Furthermore, data from codistributed amphibians suggest that there may have been two refugia, and an AMOVA shows that most of the molecular variance partitioned between the Clearwater and the Salmon drainages (54.40%; $P < 0.001$) and within drainages (43.61%; $P < 0.001$). As a result, we also tested three a priori hypotheses which predicted that both the Clearwater and Salmon drainages functioned as refugia during the late Pleistocene; we could reject (P_{CORD} = 0.019) divergence dates during the Cordilleran glacial maxima [*c.* 20 000 years before present (ybp)], during the Sangamon interglacial (*c.* 35 000 ybp; p_{SANG} = 0.032), as well as pre-Pleistocene divergence (*c.* 1.7 Ma; $p_{pP} < 0.001$). Mismatch distributions and Tajima's *D* within the individual drainages provide further support to recent population expansion. This work demonstrates coalescent stochasticity is an important phenomenon to consider in testing phylogeographical hypotheses, and suggests that analytical methods which fail to sufficiently quantify this uncertainty can lead to false confidence in the conclusions drawn from these methods.

Keywords: coalescent stochasticity, *Dicamptodon*, MESQUITE, statistical phylogeography

Received 11 August 2004; revision received 6 October 2004; accepted 6 October 2004

Interpretation of geographically structured genetic variation is complicated by stochastic effects that result from a variety of forces, including demographic and coalescent stochasticity, incomplete lineage sorting, sampling effects, and phylogenetic uncertainty. Although phylogeographers now realize that these stochastic effects can lead to a misinterpretation of genetic data if statistical uncertainty is not accounted for (e.g. Edwards & Beerli 2000; Emerson *et al*. 2001; Knowles & Maddison 2002; Hudson & Turelli 2003), no clear consensus has been reached on how to quantify this uncertainty. One powerful method of exploring

Correspondence: Bryan C. Carstens, E-mail: cars5766@uidaho.edu phone (208) 885-2550, Fax (208) 885-7905

the susceptibility of empirical data to misinterpretation is to conduct coalescent simulations under competing hypotheses of population history (Knowles 2001). By building models of population history based on external evidence (e.g. palaeoclimatic data) and empirical estimates of relevant parameters, the potential for different population histories to produce the geographical patterns of genetic variation seen in an empirical data set can be assessed. However, because coalescent theory treats the genealogy as a random variable, testing a priori hypotheses requires that these simulations be conducted over a range of plausible genealogies. Here, we consider data collected from *Dicamptodon aterrimus*, a salamander endemic to mesic forests in the northern Rocky Mountains and a species from which all data suggests a recent range expansion from Pleistocene refugia. The alternative hypotheses that we consider differ in the number (1 vs. 2) and locality (Clearwater drainage vs. Salmon drainage) of Pleistocene refugia as well as in the depth of the divergence between putative refugia.

The mesic forests of the Pacific Northwest of North America (PNW) occur primarily along the Pacific coast, but an ecologically isolated portion of the PNW mesic forest ecosystem occurs in the northern Rocky Mountains (NRM) west of the continental divide. This region contains populations of > 150 plant, animal, and fungi lineages that also have representatives in the coastal forests (Brunsfeld *et al*. 2001; Nielson *et al*. 2001). One of these mesic forest disjuncts is the Idaho giant salamander (*D. aterrimus*), and the disjunction in this species is apparently the result of the xerification of the Columbia basin caused by the orogeny of the Cascades range some 2–5 Ma (Brunsfeld *et al*. 2001). The deep genetic divergence and reciprocal monophyly between coastal congeners and *D. aterrimus* indicate that the establishment of NRM populations predates the Pleistocene and imply that *D. aterrimus* persisted throughout the Pleistocene glaciation in one or more refugia located somewhere within the NRM.

As it did throughout most of the Northern Hemisphere, Pleistocene glaciation had a tremendous impact on the mesic forests in the NRM. During the most recent glacial maxima, approximately 18 000 years before present (ybp), the Corderillan ice sheet reached well into what is now the panhandle of Idaho (Richmond *et al*. 1965; Waitt & Thorson 1983; Delcourt & Delcourt 1993), and was the latest in a series of glaciers that formed and retreated repeatedly during the last 2 Myr. The advancing glaciers are thought to have forced the mesic forest ecosystem into refugia that were located to the south of glaciation in riverine canyons deep enough to offer climatic insulation (Daubenmire 1975). The Clearwater drainage in north central Idaho has been proposed as one such refugium, based on both geology and phytogeography (Daubenmire 1975). Furthermore, genetic data in the codistributed salamander *Plethodon idahoensis* are consistent with post-Pleistocene expansion from the Clearwater refugium (Carstens *et al*. 2004a). The other possible location for a mesic forest refugium is the South Fork of the Salmon River (SFS). Data from the Rocky Mountain tailed frog (*Ascaphus montanus*) are consistent with expansion from both of these putative refugia (Nielson *et al*. 2001), as well as a relatively ancient (pre-Pleistocene) divergence between populations derived from these refugia. Based on available a priori evidence, we considered five hypotheses regarding Pleistocene population structure in *D. aterrimus*: (1) a single refugium located in the Clearwater drainage; (2) a single refugium located in the Salmon drainage; (3) two refugia, one located in the Clearwater and a second located in the SFS drainage, separated since the middle of the Cordilleran glaciation (i.e. late Pleistocene, *c*. 20 000 ybp); (4) two refugia, located in the Clearwater and the SFS drainage (as above), but separated since the Sangamon interglacial (*c*. 135 000 ybp); and (5) two refugia, again located in the Clearwater and the SFS drainages, but separated by deep divergence (i.e. pre-Pleistocene divergence, *c*. 1.7 Ma). By analysing data simulated under explicit coalescent models based on these hypothesized population histories, we can evaluate the probability that the observed data were generated by each of these histories.

Materials and methods

Sample collection

We surveyed the entire known range of *Dicamptodon aterrimus* (Fig. 1) during the summers of 2002 and 2003 and collected tail clips from 108 individuals in 30 localities (Appendix I). We received seven additional specimens from M. Nielson and three specimens from the Museum of Vertebrate Zoology (MVZ) at the University of California-Berkeley (MVZ187984, MVZ187985, MVZ187986). Tissue from a single individual of each of the congeners (*Dicamptodon copei*, *Dicamptodon ensatus*, and *Dicamptodon tenebrosus*) was also obtained from the MVZ (MVZ187978, MVZ192640, and MVZ192671) and used as outgroup taxa. Sample sizes ranged from $n = 6$ to $n = 41$ per drainage, with an average of $n = 16.9$. Therefore, following Saunders *et al*. (1984) [$p = (n - 1)/(n + 1)$, where n is the sample size], we have high probability of sampling the deepest coalescent event in each drainage ($P = 0.71–0.95$). However, our actual probability of sampling the deepest coalescent event within a drainage may be higher, because ancestral haplotypes are expected to be most frequent in the population (Waterson & Guess 1977; Crandall & Templeton 1996).

DNA was extracted from 10 to 20-mg tail clips, which had been stored in 90% EtOH, using the DNeasy Tissue kit (Qiagen, Inc.), following the manufacturer's instructions for rodent tails. To amplify the cytochrome *b* (cyt *b*) gene, we designed primers in the tRNA-Threonine (5'-

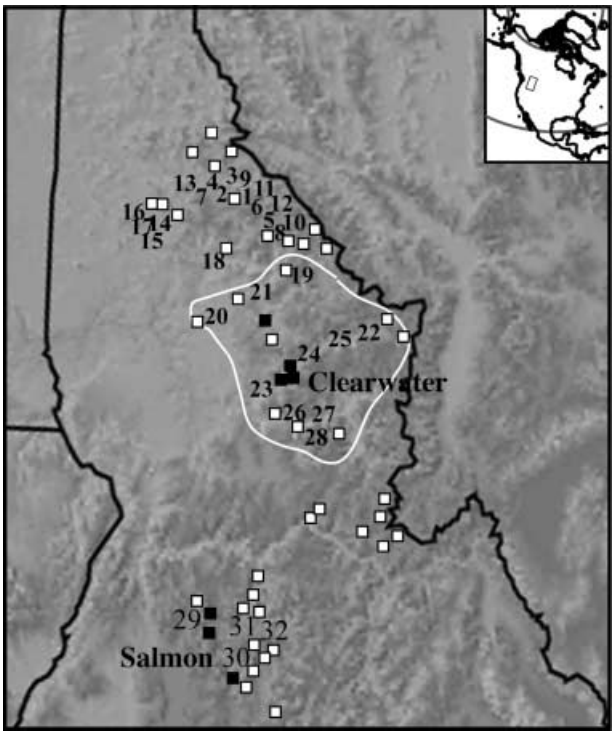

Fig. 1 Map of localities from which *Dicamptodon aterrimus* were collected. Numbers on the map correspond to localities listed in Appendix I. Boxes represent localities that were surveyed for > 2 person-hours without observing any *D. aterrimus*; black boxes are historic localities and white boxes are new localities. The approximate boundaries of the Clearwater drainage are marked by the white line. The inset shows the position of the larger map on the North American continent.

TTCAGCTTACAAGGCTGATGTTTT-3′) and the tRNA-Glucine (5′-TTGTATTCAACTATAAAAAC-3′) using conserved portions of these tRNAs from complete mitochondrial genomes of urodeles deposited in GenBank [*Andrias davidianus* (NC 004926; Zhang *et al*. 2003a), *Ranadon sibiricus* (NC 004021; Zhang *et al*. 2003b), *Mertensiella luschani* (NC 002756; Zardoya *et al*. 2003)]. We initially amplified the first two-thirds of cyt *b* using the primer in the tRNA-Glucine and a primer designed for rodents (H-14963; 5′-GGCAAATA-GAARTATCATT-3′; Sullivan *et al*. 2000). For subsequent PCR (polymerase chair reaction) amplification, we designed lineage-specific internal primers (Forward internal 5′-TCCACCCATACTTTTCTTATAAAGA-3′; Reverse internal 5′-TAATTAGTGGATTTGCTGGTGTAA-3′). We also sequenced a portion of the mitochondrial control region (CR) using the THR (5′-AAAACATCGATCTTGTAAGTC-3′) and 007 (5′-GCACCCAAAGCCAAAATTTTCA-3′) primers from Shaffer & McKnight (1996). Amplicons were purified using polyethylene glycol precipitation, and sequencing reactions were performed with the BigDye Kit (Applied Biosystems) with 20–40 ng of (PCR) product in 15 μl volumes. CentriSep columns (Princeton Separations) were used to filter sequencing reactions, and samples were run on an ABI 377 automated sequencer using 5% Long Ranger polyacrylamide gels. Cyt *b* and the CR were sequenced in both the 5′- and 3′ directions, and edited and aligned with SEQUENCER 3.0 (GeneCodes). Sequences were deposited in GenBank under accession numbers AY728902–AY729015.

Phylogenetic methods

We used maximum likelihood (ML) phylogeny estimation to identify the major clades within *D. aterrimus*. We first pruned redundant haplotypes, and used DT-MODSEL (Minin *et al*. 2003) to select a model of sequence evolution for use in subsequent analyses. This method of model selection incorporates fit, a penalty for over-parameterization, and performance into the selection process. We verified that the model chosen was appropriate by using an absolute goodness-of-fit test (Goldman 1993; Sullivan *et al*. 2000) with 1000 replicates. ML analyses were conducted with PAUP* 4.0 (Swofford 2002) using a heuristic search with the chosen model, TBR branch swapping, 10 random-addition replicates, and a random-addition starting tree. Nodal support was assessed using bootstrap analysis with 200 replicates (Felsenstein 1985). We also conducted 10^7 generations of a Bayesian search using MRBAYES (Huelsenbeck & Ronquist 2001) in order to estimate the Bayesian posterior probability (Bpp) of nodes in the phylogeny estimate, using a flat beta prior for the transition/transversion ratio (e.g. 1, 1) and uniform priors for the topology, shape of gamma distribution of rate heterogeneity, and rate correlations, and using empirical base frequencies as priors.

Population genetic methods

The genetic structure of *D. aterrimus* was explored using AMOVA (Cockerham 1969; Excoffier *et al*. 1992) and Nested Clade Analysis (NCA) (Templeton *et al*. 1987). We conducted the AMOVA using ARLEQUIN 2.0 (Excoffier *et al*. 1992; Schneider *et al*. 2000) in two ways. We first defined sample partitions in a pattern designed to correspond to a single refugium. We defined river drainages as populations, and anticipated that most of the molecular variance would occur within populations. Second, we partitioned populations in a manner intended to correspond to two refugia; we defined each river drainage as a population and grouped the six drainages in the Clearwater separately from the Salmon.

For the NCA, we assessed the significance of the correlation between haplotypes and geography by using TCS (Clement *et al*. 2001) to estimate a haplotype network, and nested this network into groupings following Templeton (1998). Unweighted geographical distances were calculated from latitude and longitude, and GEODIS (Posada *et al*. 2000) was used to calculate the average geographical distance of

the haplotypes to the centre of each clade (D_c), the average distance of the haplotypes in each clade to the centre of the clade at the next nesting level (D_n), and the difference between tip clades and interior clades for D_c and D_n. We tested for the significant association of haplotype with geography using a categorical permutation contingency analysis. For nested clades with a significant association of haplotype and geography, we followed the inference key of Templeton (2004).

We also tested for localized population expansion within each of the seven drainages using Tajima's *D* (Tajima 1989) and mismatch distributions (Rogers & Harpending 1992). Nucleotide diversity (μ), the number of pairwise differences (*k*), and the number of polymorphic sites (*s*) were calculated with ARLEQUIN and used in the calculation of Tajima's *D*. Although originally developed as a test of selection, in its absence, significantly negative values of Tajima's *D* are thought to be evidence of expanding populations (Rogers 1995), but can be biased by sampling artefacts and population structure (Hammer *et al*. 2003). In addition, the demographic expansion model of Rogers & Harpending (1992) was used to analyse the pairwise mismatch distributions. Populations that have been historically stable are predicted to have multimodal mismatch distributions, whereas those that have undergone a recent expansion are predicted to be unimodal. We calculated the mismatch distributions with ARLEQUIN (Schneider *et al*. 2000) and used the Raggedness Index of Harpending (1994) to assess the significance of the fit of the distribution to that of an expanding population. We also calculated the maximum corrected sequence divergence within each drainage using PAUP*.

Hypothesis testing

Although some approaches to analysing phylogeographical data, notably Nested Clade Analysis (Templeton 1998), have incorporated aspects of coalescent theory, researchers have had difficulties developing and testing hypotheses that are both explicitly based in coalescent theory and specific to a given taxon and its geographical distribution. This is primarily because most analytical methods lack the flexibility to build taxon-specific null distributions. We used MESQUITE 1.03 (Maddison & Maddison 2004), which provides considerable flexibility, to investigate models of possible population structure during Pleistocene glaciation. We conducted coalescent simulations under five a priori hypotheses of Pleistocene population structure; two single-refugium hypotheses and three dual-refugium hypotheses (Fig. 2). MESQUITE was used to simulate 1000 coalescent genealogies constrained within the population history predicted by the hypothesis being tested. We selected a second model of sequence evolution using DT-MODSEL using only samples from *D. aterrimus* (e.g. without the outgroup), and verified this model with a goodness-of-fit test (Goldman 1993). We then simulated DNA sequences with the same parameters as our actual data on each genealogy on each of the replicate gene trees.

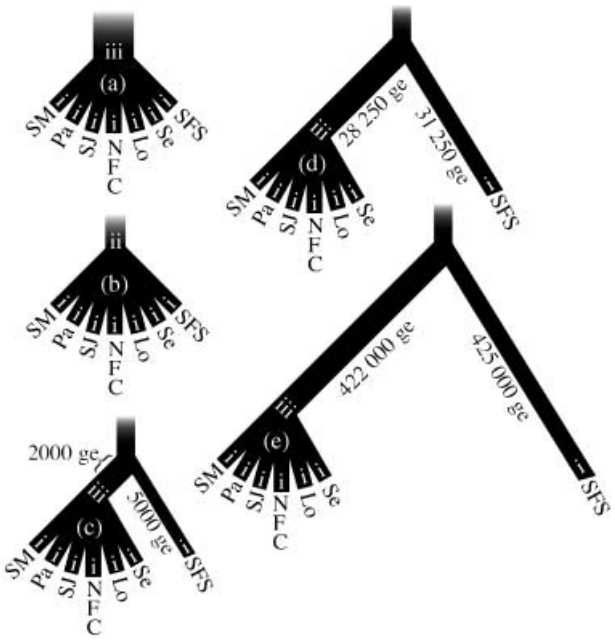

Fig. 2 Illustration of the five hypotheses of Pleistocene population structure: (a) a single Clearwater refugium (b) a single South Fork of the Salmon refugium (c) two refugia, separated from each other divergence dating to the Cordilleran glaciation (d) two refugia, separated by deeper divergence dating to the Sangamon interglacial period, and (e) two refugia, separated by pre-Pleistocene divergence. For each hypothesis, we simulated 1000 data set on coalescent genealogies constrained within these population histories. Each simulation used an N_e = 58 600. Populations with constrained effective population sizes are marked with white 'i' as follows: 51.5% of total N_e (iii), 27.3% of total N_e (ii) 14.3% of total N_e (i). Drainages are abbreviated as follows: South Fork of Salmon (SFS), Selway (Se), Lochsa (Lo), North Fork of the Clearwater (NFC), St. Maries (SM), Palouse (Pa), St. Joe (SJ). All root branches are 5000 generations long. All branches leading from the refugia in 2 A and 2B are 3000 generations long, as are branches leading from the Clearwater refugium in 2C, 2D, and 2E. Other branches are marked with their length in generations (ge).

Tests of a single refugium

Two hypotheses predict that a single drainage functioned as a refugium during the Pleistocene. Hypotheses (1) and (2) assume that extant *D. aterrimus* populations are derived from this refugial population, so we modelled population expansion at the close of the Pleistocene and assumed that population expansion began as the glaciers began to retreat *c*. 12 000 ybp. The single-refugium hypotheses are differentiated only by the expected difference in the effective population size (N_e) of the refugial population (Fig. 2a,b). We used three estimates of N_e; the overall effective population

size, N_e estimated from samples collected in the Clearwater, and N_e estimated from samples collected in the South Fork of the Salmon. Empirical estimates of N_e were computed using MIGRATE-N (v. 1.7.3; Beerli 2002) to estimate theta ($\theta = 2N_e\mu$), which was used to calculate N_e assuming $\mu = 1.0 \times 10^{-7}$. To conduct the coalescent simulations, we set the overall N_e to equal the empirically estimated values and constrained the N_e of the refugial population to a size proportional to the relative N_e of the population sampled from the site of the putative refugium. Thus, if samples from the site of the putative refugium had an effective population size of one-third that of the overall N_e, we would constrain the population size in the simulation prior to population expansion to one-third the total N_e.

Tests of two refugia

The other hypotheses (3–5) were suggested by Nielson *et al*. (2001), who proposed that *Ascaphus montanus*, the Rocky Mountain tailed frog (which is codistributed and frequently syntopic with *D. aterrimus*), persisted throughout the Pleistocene in two refugia, one located in the Clearwater and a second located in the Salmon River drainage. A dual-refugium hypothesis would be supported in the Idaho giant salamander by a deep divergence between modern populations located in the Clearwater and Salmon drainages, as seen in phylogeny of *A. montanus* (Nielson *et al*. 2001). The key difference between the dual-refugium hypotheses is the degree of genetic divergence that separates the Salmon populations from the other populations; we used a date that corresponded to the middle of the Cordilleran glaciation (*c*. 20 000 ybp) for third hypothesis (Fig. 2c), a date that corresponded to the Sangamon interglacial (*c*. 135 000 ybp) for the fourth hypothesis (Fig. 2d), and a pre-Pleistocene date (*c*. 1.7 Ma) for the fifth hypothesis (Fig. 2e). For converting coalescent time (in generations) to absolute time, we assumed a generation length of 4 years (Nussbaum *et al*. 1983). As in the single-refugium hypotheses, we constrained the relative N_e of simulated populations to match values estimated from the data.

Parameters used to test hypotheses

Each hypothesis was tested with the *S* statistic of Slatkin & Maddison (1989). This parameter treats populations as categorical variables and measures the minimum number of sorting (or migration) events implied by the genealogy. This statistic was then used to assess the significance of the discord between the genealogy and the subdivision of our samples into different drainages. We also tested the single-refugium hypotheses with a test statistic based on the difference in genetic distance present in each of the putative refugial population, using corrected ML distances. If extant populations are recently derived from a single refugium, we expect the greatest genetic diversity to be located within the populations occupying the site of this refugium, as these populations would be the oldest. We computed corrected genetic distance matrices for the SFS and Clearwater samples with PAUP*, parsed the greatest genetic distance within each simulated data set, and used these distances to construct null distributions of the difference between the greatest genetic distance present within the SFS and Clearwater drainages ($D_{SFS} - D_{CLW}$).

Results

Samples collected

We sequenced 1951 nucleotides from two regions of the mitochondrial genome, corresponding to portions of three mitochondrial genes in *Mertensiella luschani* (Zardoya *et al*. 2003): the entire cyt *b* gene (positions 1–1140 of our data) corresponding to positions 14228–15266 of the *M. luschani* genome, a portion of the tRNA-Threonine (positions 1141–1176 of our data) corresponding to positions 15267–15302 of the *M. luschani* genome, and a portion of the control region (positions 1177–1951 of our data) corresponding to positions 15772–16546 of the *M. luschani* genome. Fifty-seven haplotypes were sampled in *Dicamptodon aterrimus* (Table 1); 44 haplotypes were unique to single individuals and 13 were found in multiple individuals. The most

Table 1 Number of haplotypes by drainage for *Dicamptodon aterrimus*. Haplotypes sampled in multiple individuals are labelled A–M. The final column, denoted with the asterisk, lists the number of individuals from each drainage with a haplotype not sampled in any other individual. Haplotype letters correspond to those in Figs 2 and 3

Drainage	*n*	A	B	C	D	E	F	G	H	I	J	K	L	M	*
St. Maries	6	3	—	—	—	1	2	—	0	—	—	—	—	—	—
St. Joe	41	14	—	—	—	—	—	3	4	2	—	—	—	—	18
Palouse	16	7	—	2	—	—	—	—	0	—	—	2	—	—	5
NFC*†	9	3	—	—	—	—	—	—	0	2	—	—	—	—	4
Lochsa†	14	2	3	—	3	—	—	—	0	—	—	—	—	—	6
Selway†	13	5	—	—	—	1	—	—	0	1	2	—	—	—	4
EFSF Salmon	19	—	—	—	—	—	—	—	0	—	—	—	10	2	7
Clearwater	36	10	3	—	3	1	—	—	0	3	2	—	—	—	14
Total	118	34	3	2	3	2	2	3	4	5	2	2	10	2	44

*North Fork of the Clearwater; †NFC, Lochsa, and Selway are the major tributaries of the Clearwater drainage.

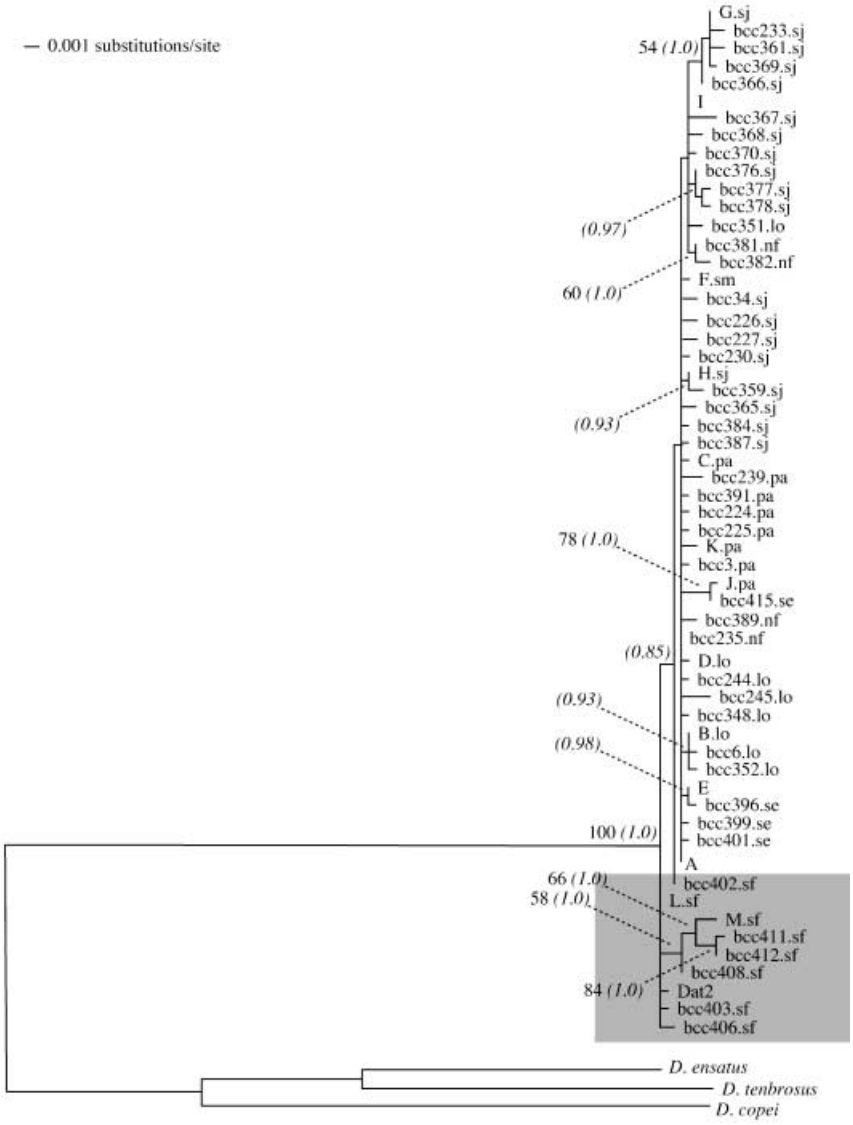

Fig. 3 ML phylogeny estimate for the *Dicamptodon aterrimus* complex using a HKY + Γ model of sequence evolution (ln $L = -4364.94360$). ML bootstrap values from 200 replicates are shown for nodes recovered in > 50% of the bootstrap data sets. Bayesian posterior probabilities are shown in italics. Taxon labels correspond to haplotypes listed in Table 1, with unique haplotypes labelled according to their collection number. All haplotypes are also labelled by drainage as follows: St. Maries (.sm), St. Joe (.sj), Palouse (.pa), North Fork of the Clearwater (.nf), Lochsa (.lo), Selway (.se), or South Fork of the Salmon (.sf). Haplotypes from the South Fork of the Salmon are shown enclosed in a grey box.

frequently sampled haplotype, labelled 'A', was found in 29% of all individuals and was present in 18 of the 32 localities, none of which were located in SFS.

Phylogenetic analyses

For the phylogenetic analyses (including the outgroups), we selected the HKY + Γ model of sequence evolution using DT-MODSEL with the following parameters: $\pi_A = 0.31943$, $\pi_C = 0.18268$, $\pi_G = 0.15652$, $\pi_T = 0.34137$, t-ratio = 1.704, $\alpha = 0.208$, ncat = 8. This model could not be rejected with an absolute goodness-of-fit test ($P = 0.688$) and was used in subsequent phylogenetic analyses. We found a single best ML tree, with a likelihood score ln $L = -4364.94360$ (Fig. 3). *D. aterrimus* monophyly is well supported (100% bs; 1.0 Bpp). Within *D. aterrimus*, haplotypes from the South Fork of the Salmon River are paraphyletic and basal to a large clade containing haplotypes from the Clearwater drainage and other drainages north of the Clearwater. Most of the relationships within *D. aterrimus* are poorly supported by the bootstrap analysis, and only two clades were found within a single drainage: the clade composed of five haplotypes from the St. Joe (G + bcc233 + bcc361 + bcc366 + bcc369; 54 bs; 1.0 Bpp), and the clade composed of four haplotypes from SFS (M + bcc408 + bcc411 + bcc412; 58% bs; 1.0 Bpp). The overall phylogenetic pattern suggests expansion from a single refugium in the SFS.

Population genetic analyses

The AMOVA conducted with populations partitioned according to a single refugium suggested that most of the molecular

Table 2 Diversity statistics by river drainage for *Dicamptodon aterrimus*. The name of the drainage, the number of specimens collected in that drainage (*n*), the probability of capturing the deepest coalescent event (*Prob*), the number of haplotypes (# *hap*) the number of polymorphic sites (*s*), the nucleotide diversity (π) ± one standard deviation, Tajima's *D* (with significant values in bold), the *P*-value for the mismatch distributions, and the greatest corrected sequence divergence within the drainage (*Seq. Division*, in units of substitutions per site), are shown

Drainage	*n*	*Prob*	# *haps*	*s*	π	*D*	*mismatch*	*Seq. Division*
St. Maries	6	0.71	3	2	0.0004 ± 0.0004	1.52	0.303	0.00103
St. Joe	41	0.95	22	29	0.0013 ± 0.0008	**−1.75**	0.849	0.00368
Palouse	16	0.89	9	11	0.0008 ± 0.0006	**−1.43**	0.394	0.00155
Clearwater	36	0.95	20	26	0.0016 ± 0.001	*n/a*	*n/a*	0.00476
NFC*†	9	0.80	6	6	0.0009 ± 0.0007	0.18	0.913	0.00208
Lochsa†	14	0.88	9	12	0.0011 ± 0.0007	**−1.11**	0.400	0.00367
Selway†	13	0.75	7	9	0.0009 ± 0.0006	**−1.11**	0.508	0.00368
SF Salmon	19	0.90	9	17	0.002 ± 0.0011	−0.23	0.67	0.00583
Total	118	0.98	57	76	0.0018 ± 0.001	*n/a*	*n/a*	0.00801

*North Fork of the Clearwater; †NFC, Lochsa, and Selway are the major tributaries of the Clearwater drainage.

variance occurred within drainages (68.24%; $P < 0.001$), but with significant differentiation (31.76% ($P < 0.001$) among all drainages. When we partitioned the samples in a matter consistent with two refugia, most of the molecular variance was partitioned between the Clearwater and the Salmon drainages (54.40%; $P < 0.001$) and within drainages (43.61%; $P < 0.001$), with only a small percentage of the variance partitioned among the drainages in the Clearwater or to the north of the Clearwater (1.99%; $P = 0.14$). There is also evidence for population expansion in each of the drainages from Tajima's *D* and/or the mismatch distributions (Table 2).

The permutation test of NCA indicated a significant association between haplotypes and geography in two subclades (2–9, 3–1; Fig. 4). For both of these nested subclades, the inference reached following the key of Templeton (2004) was one of range expansion. Like the phylogeny estimate, the NCA appears to support a single refugium, but unlike the phylogeny estimate, the inference from NCA is that the Clearwater was the site of this refugium. Both the placement of haplotype 'A' at the centre of the network, which suggests that it is ancestral (Crandall & Templeton 1996), and the significant range expansion in nested clades that contain samples from the Clearwater drainage support this inference.

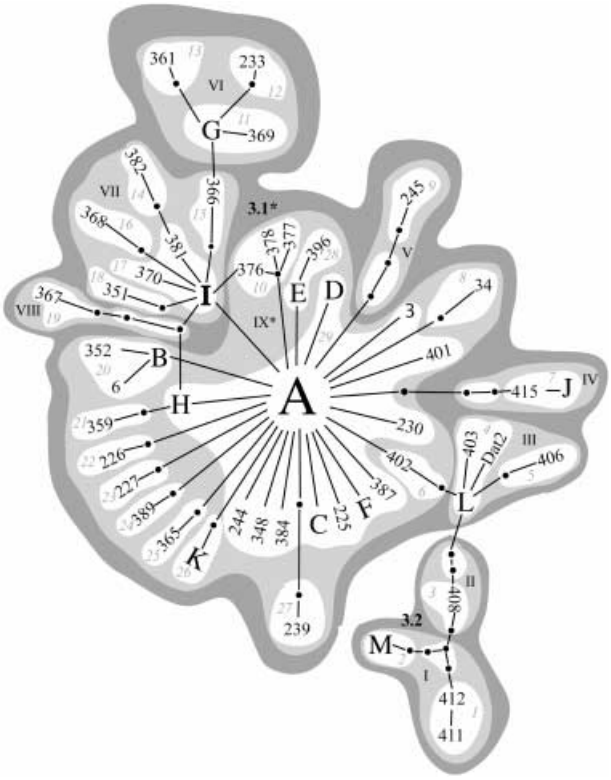

Fig. 4 Haplotype network among all *Dicamptodon aterrimus* haplotypes. Haplotypes are connected by lines, and missing steps are labelled with a black dot. Nesting levels for the NCA are shown, with the one-step clades in white, the two-step clades in light grey, and the three-step clades in dark grey. The clade numbers are also shown, with clades that have a significant association of haplotype and geography denoted with an asterisk.

Tests of phylogeographical hypotheses with MESQUITE

We constructed our Pleistocene population hypotheses with several empirical estimates (using MIGRATE-N) of θ?: $\theta_{TOTAL} = 0.01172$; $\theta_{CLW} = 0.00604$; $\theta_{SFS} = 0.0032$. We were able to calculate $N_{eTOTAL} = 58\,600$, $N_{eCLW} = 30\,200$, and $N_{eSFS} = 16\,000$ with these θ-values, assuming $\mu = 1.0 \times 10^{7}$. From our data, we computed Slatkin and Maddison's $S = 50$ and $D_{SFS} - D_{CLW} = 0.00107$. We selected the F_{81} + I model of sequence evolution for the *D. aterrimus* data, with the following parameters: $\pi_A = 0.31951$, $\pi_C = 0.18263$, $\pi_G = 0.1564$, $\pi_T = 0.34146$, *pinv* = 0.882271. This model could not be rejected with an absolute goodness-of-fit test ($P = 0.58$) and was used in subsequent coalescent simulations.

We tested five hypotheses of putative population structure. The Clearwater refugium hypothesis could be rejected at the $P = 0.1$ significance level using Slatkin and Maddison's S ($p_{CLW} = 0.089$), as well as $D_{SFS} - D_{CLW}$ ($P = 0.010$). We could not reject the SFS hypothesis with either S ($p_{SFS} = 0.329$) or $D_{SFS} - D_{CLW}$ ($P = 0.117$). However, we were able to reject all of the two-refugia models with the S statistic ($p_{CORD} = 0.019$; $p_{SANG} = 0.032$; $p_{pP} < 0.001$). The coalescent simulations suggest that the hypothesis of a single refugium in the SFS has the strongest support in our data.

Discussion

Stochasticity and hypothesis testing

Phylogeographical research is conducted at the interface of phylogenetics and population genetics, and borrows analytical methods from each. While systematists have developed approaches for dealing with phylogenetic uncertainty (see Huelsenbeck & Rannala 1997), comparable approaches have yet to be widely adopted by phylogeographers to deal with coalescent stochasticity. While several theoretical studies (Edwards & Beerli 2000; Emerson *et al*. 2001; Knowles & Maddison 2002; Hudson & Turrelli 2003) have demonstrated the importance of quantifying coalescent uncertainty, few empirical studies have done so (but see Knowles 2001; Carstens *et al*. 2004b). Two types of approaches can be used to quantify coalescent uncertainty; methods that estimate a confidence interval around a parameter of interest and methods that use simulation to construct a null distribution of the parameter under a particular hypothesized population history. The growing awareness of coalescent stochasticity and the need to account for it has led to the introduction of a variety of new analytical methods. MESQUITE is perhaps the most flexible among these; it allows taxon-specific models to be built and tested, and can be integrated into other analytical methods.

Traditional approaches suggest that *Dicamptodon aterrimus* was present in a single refugium during the Pleistocene, but were unable to identify the site of this refugium. The NCA suggested that the Clearwater drainage was the site of this refugium on the basis of two results. The haplotype at the centre of the network, which is inferred as ancestral and not found in the South Fork of the Salmon drainage, and the inference of range expansion in nested clades is located within the Clearwater drainage. Conversely, the phylogeny estimate is inconsistent with this scenario because the basal haplotypes are from the SFS. Furthermore, there is more genetic diversity within the SFS population than within all of the populations that comprise the Clearwater drainage, and we expect the source population to contain the greatest genetic diversity. Using the coalescent simulations, we were able to reject each of the two-refugia hypotheses (strongly) as well as the Clearwater refugium hypothesis (weakly). These findings illustrate that it is critical both to build rigorous a priori hypotheses that are specific to particular taxa of interest and to model the expectations of those hypotheses using coalescent simulations.

The quality of the coalescent simulations depends on several factors. First, models of population structure (e.g. Fig. 2) are of critical importance. We based the depth of these population histories on geological dates in order to test a set of hypotheses that spanned an array range of plausible Pleistocene population structures. Effective population size was used to differentiate between the two single-refugium hypotheses, and we scaled the size of the population at critical times in the past in a manner consistent with the geological record of the NRM during glaciation. Our goal was to build simple, yet differentiable, models that captured the important differences among alternative hypotheses in order to test them. Once we built these models, 1000 coalescent genealogies were simulated within the population trees, and sequence data was simulated on these genealogies (i.e. gene trees) using the model of nucleotide substitution chosen from our empirical data. The accuracy of simulation-based tests is also dependent on the appropriateness of this substitution model. We first conducted these simulations using the model of sequence evolution selected from the data including the outgroup (HKY + Γ); we were unable to differentiate among the two single-refugium hypotheses and the dual refugia hypothesis with shallow divergence (Fig. 2a–c). However removal of the outgroup allowed us simplify the model; we could replace Γ-distributed rates with a invariable sites and use a single substitution type. The F_{81} + I model fits the ingroup acceptable well (see goodness of fit test) and should have a much reduced variance relative to the HKY + Γ. Thus simulation tests under this ingroup-specific model apparently provide more resolution for hypothesis testing. However theoretical work and simulation studies are needed to assess the possible effects of misspecification of both population models and models of nucleotide substitution.

Conservation implications

Our results have significant implications for the conservation of *D. aterrimus*. They imply that much of the diversity within *D. aterrimus* was restricted to a small population during the Pleistocene, and that current populations are the result of a post-Pleistocene expansion, most probably from the SFS. When current population densities are considered, these results are troubling. During fieldwork in the summer of 2003, we were unable to locate *D. aterrimus* from three historic collection localities in the SFS and four historic localities in the Clearwater, suggesting that populations in these localities are extremely small or may have undergone recent extirpations. We expected *D. aterrimus*

to occur at highest densities at localities within the likely Pleistocene refugium, but *D. aterrimus* appears to occur at lower densities within the SFS and Clearwater drainages (Fig. 1) when compared to the qualitatively higher densities observed on the Palouse and St. Joe rivers. Populations within the Clearwater appear to have been harmed by two events, the use of DDT (dichlorodiphenyltrichloroethane) in the Lochsa drainage (a major tributary of the Clearwater) during the 1960s (Moore 1996) and widespread logging in the Lochsa River and North Fork of the Clearwater River drainages. Idaho giant salamanders may be vulnerable to logging practices because their density is strongly correlated with the availability of cover items (e.g. large rocks and logs) in the streams that they inhabit as larvae (Parker 1991), and deforestation increases the sedimentation in the streams, thereby decreasing the availability of cover items. This is particularly harmful to reproduction because *Dicamptodon* oviposits in undercut banks (Nussbaum 1969), which are particularly susceptible to increased sedimentation. We were only able to locate the salamanders in minimally logged tributaries that had an abundance of cover items. Given the scarcity of *D. aterrimus* in drainages that contain the highest genetic diversity and which may be the sites of the oldest populations, the Idaho giant salamander merits further investigation and protection.

Conclusions

We have demonstrated using an empirical data set that coalescent stochasticity can complicate the testing of phylogeographical hypotheses. Stochastic forces can produce patterns in empirical data that have evolved on a complex geography that appear to be internally inconsistent. Coalescent simulations that model hypothesized population histories offer one method of quantifying the range of coalescent uncertainty, and as such are critically important additions to the set of methods commonly used by phylogeographical researchers.

Acknowledgements

We received critical advice from C. Peterson concerning the best locations to search for *Dicamptodon aterrimus*. We thank the Museum of Vertebrate Zoology at the University of California-Berkeley and M. Neilson for providing samples, and J. Carstens, M. Wirtz, and C. Steele for assistance in collecting samples. We thank members of the Sullivan lab for reviewing this manuscript prior to submission. Funding for B. Carstens was provided by the University of Idaho Presidential Fellowship. Funding for computer hardware used for much of the computational aspect of this work was provided with funding from NIH National Center for Research Resources grants P20 RR16454 for the BRIN Program, and P20 RR16448 for the COBRE Program. Funding for J. Degenhardt and A. Stevenson was provided by the IDFG State Wildlife Grant F-09–03. Funding used for laboratory reagents and supplies was provided by NSF EPSCoR 809935. This manuscript was significantly improved as a result of reviews provided by G. Wallis and four anonymous reviewers, and we are grateful for their efforts.

References

Beerli P (2002) MIGRATE: *Documentation and Program, Part of* LAMARCK *Version 1.5*. http://evolution.genetics.washington.edu/lamarc.html.

Brunsfeld SJ, Sullivan J, Soltis DE, Soltis PS (2001) Comparative phylogeography of northwestern North America: A synthesis. In: *Integrating Ecology and Evolution in a Spatial Context* (eds Silvertown J, Antonovics J), pp. 319–339. Blackwell Publishing, Williston, VT.

Carstens BC, Stevenson AL, Degenhardt JD, Sullivan J (2004a) Testing nested phylogenetic and phylogeographic hypotheses in the *Plethodon vandykei* species group. *Systematic Biology*, **53**, 781–792.

Carstens BC, Sullivan J, Davalos LM, Larsen PA, Pedersen SC (2004b) Exploring population genetic structure in three species of Lesser Antillean bats. *Molecular Ecology*, **13**, 2557–2566.

Clement M, Posada D, Crandall K (2001) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.

Cockerham CC (1969) Variance of gene frequencies. *Evolution*, **23**, 72–83.

Crandall KA, Templeton AR (1996) Applications of intraspecific phylogenetics. In: *New Uses for New Phylogenies* (eds Harvey PH, Brown AJL, Maynard Smith J, Nee S), pp. 81–102. Oxford University Press, New York, NY.

Daubenmire R (1975) *Floristic Plant Geography of Eastern Washington and Northern Idaho*. Brigham Young University Press, Provo, UT.

Delcourt PA, Delcourt HR (1993) Paleoclimates, paleovegetation, and paleofloras during the late Quaternary. In: *Flora of North America*, Vol. 1, 1st edn. Editorial Committee, pp. 71–94. Oxford University Press, New York.

Edwards SV, Beerli P (2000) Gene divergence, population divergence, and the variance in coalescent time in phylogeographic studies. *Evolution*, **54**, 1839–1854.

Emerson BC, Paradis E, Thebaud C (2001) Revealing the demographic histories of species using DNA sequences. *Trends in Ecology and Evolution*, **16**, 707–716.

Excoffier L, Smouse P, Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, **136**, 343–359.

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

Goldman NJ (1993) Statistical tests of models of DNA substitution. *Journal of Molecular Evolution*, **36**, 182–198.

Hammer MF, Blackmer F, Garrigan F, Nachman MW, Wilder JA (2003) Human population structure and its effects on sampling Y chromosome sequence variation. *Genetics*, **164**, 1495–1509.

Harpending RC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591–600.

Hudson RR, Turelli M (2003) Stochasticity overrules the 'three-times rule': genetic drift, genetic draft, and coalescent times for nuclear loci versus mitochondrial data. *Evolution*, **57**, 182–190.

Huelsenbeck JP, Rannala B (1997) Phylogenetic methods come of age: testing hypotheses in a phylogenetic context. *Science*, **276**, 227–232.

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754–755.

Knowles LL (2001) Did the Pleistocene glaciations promote divergence? Test of explicit refugial models in Montane grasshoppers. *Molecular Ecology*, **10**, 691–701.

Knowles LL, Maddison WP (2002) Statistical phylogeography. *Molecular Ecology*, **10**, 2623–2635.

Maddison WP, Maddison DR (2004) MESQUITE: a Modular System for Evolutionary Analysis, version 1.01. available at http://mesquiteproject.org.

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

Moore B (1996) *The Lochsa Story: Land Ethics in the Bitterroot Mountains*. Mountain Press Publications, Missoula, MT.

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

Nussbaum RA (1969) Nests and eggs of the Pacific giant salamander, *Dicamptodon ensatus* (Eschscholtz). *Herpetologica*, **25**, 257–262.

Nussbaum RA, Brodie Jr ED, Storm RM (1983) *Amphibians and Reptiles of the Pacific Northwest*. University of Idaho Press, Moscow, ID.

Parker MS (1991) Relationship between cover availability and larval Pacific giant salamander density. *Journal of Herpetology*, **25**, 355–357.

Posada D, Crandall KA, Templeton AR (2000) GEODIS: a program for the nested cladistic analysis of the geographical distribution of genetic haplotypes. *Molecular Ecology*, **9**, 487–488.

Richmond GM, Fryxell R, Neff GE, Weis PL (1965) The Cordilleran ice sheet of the northern Rocky Mountains and related Quaternary history. In: *The Quarternary of the United States* (eds Wright HE Jr, Frey DG), pp. 231–242. Princeton University Press, Princeton, NJ.

Rogers A (1995) Genetic evidence for a Pleistocene population explosion. *Evolution*, **49**, 608–615.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Saunders IW, Tavare S, Watterson GA (1984) On the genealogy of nested subsamples from a haploid population. *Advances in Applied Probability*, **16**, 471–491.

Schneider S, Roesli D, Excoffier L (2000) *ARLEQUIN v2.0: Documentation and Program*. URL: http://anthro.unige.ch/arlequin.

Shaffer HB, McKnight ML (1996) The polytypic species concept revisited: genetic differentiation and molecular phylogenetics of the tiger salamander *Ambystoma tigrinum* (Amphibia: Caudata) complex. *Evolution*, **50**, 417–433.

Slatkin M, Maddison WP (1989) A cladistic measure of gene flow inferred from phylogenies of alleles. *Genetics*, **123**, 603–613.

Sullivan J, Arellano E, Rogers DS (2000) Comparative phylogeography of Mesoamerican highland rodents: concerted versus independent responses to past climatic fluctuations. *American Naturalist*, **155**, 755–768.

Swofford DL (2002) *PAUP*. Phylogenetic Analysis Using Parsimony (and Other Methods), Version 4*. Sinauer Associates, Sunderland, Massachusetts.

Tajima F (1989) The effect of change in population size on DNA polymorphism. *Genetics*, **105**, 437–460.

Templeton AR (1998) Nested clade analysis of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology*, **7**, 381–397.

Templeton AR (2004) Statistical phylogeography: methods of evaluating and minimizing inference errors. *Molecular Ecology*, **4**, 789–809.

Templeton AR, Boerwinkle E, Sing CF (1987) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping. I. Basic theory and analysis of alcohol dehydrogenase activity in *Drosophila*. *Genetics*, **117**, 343–351.

Waitt Jr RB, Thorson RM (1983) The Cordilleran ice sheet in Washington, Idaho, and Montana. In: *Late Quaternary Environments of the United States, the Late Pleistocene* (eds Wright HE Jr, Porter SC), pp. 53–70. University of Minnesota Press, Minneapolis, MN.

Waterson GA, Guess HA (1977) Is the most frequent allele the oldest? *Theoretical Population Biology*, **11**, 141–160.

Zardoya R, Malaga-Trillo E, Veith M, Meyer A (2003) Complete nucleotide sequence of the mitochondrial genome of a salamander, *Mertensiella luschani*. *Gene*, **317**, 17–27.

Zhang P, Chen YQ, Liu YF, Zhou H, Qu LH (2003a) The complete mitochondrial genome of the Chinese giant salamander, *Andrias davidianus* (Amphibia: Caudata). *Gene*, **311**, 93–98.

Zhang P, Chen YQ, Zhou H, Wang XL, Qu LH (2003b) The complete mitochondrial genome of a relic salamander, *Ranodon sibiricus* (Amphibia: Caudata) and implications for amphibian phylogeny. *Molecular Phylogenetics and Evolution*, **28**, 620–626.

This paper represents a portion of Bryan Carstens' Ph.D. dissertation conducted in the laboratory of Jack Sullivan. Bryan is interested in methodological approaches to analysing phylogeographic data and comparative phylogeography. Jeremiah Degenhardt is finishing his B.S. in Zoology in May of 2005, and hopes to pursue graduate work in bioinformatics. Angela Stevenson eamed her B.S. in Zoology in December of 2004. She is interested in ontering a graduate program and investigating mechanisms of prczygotic isolation. Jack Sullivan conducts empirical research on the regional phylogeography of western North America and theoretical research on methods of selecting models of sequence evolution for use in phylogenetic and phylogeographic analyses.

Appendix I

Collection localities for specimens used in this analysis. Numbers in the left-most column correspond to localities shown in Fig. 1. The name of the locality, the drainage it is from, and the number of specimens collected are shown in the centre of the table, with the latitude and longitude shown on the right side of the table. Localities 18–28 occur in the Clearwater drainage, the proposed location of one Pleistocene refugium, while localities 29–32 occur in the East Fork of the South Fork of the Salmon drainage, a proposed location for the second refugium.

	Population	Drainage	*n*	Latitude	Longitude
1	Bluff Creek	St. Joe River	1	41 11 N	115 29 W
2	Boulder Creek	St. Joe River	3	47 13 N	116 01 W
3	Burton Creek	St. Joe River	1	47 15 N	115 56 W
4	Fishhook Creek	St. Joe River	4	47 14 N	115 51 W
5	Fly Creek	St. Joe River	7	47 07 N	115 23 W
6	Lick Creek	St. Joe River	5	47 10 N	115 51 W
7	Marble Creek	St. Joe River	4	47 13 N	116 02 W
8	Mosquito Creek	St. Joe River	1	47 09 N	115 25 W
9	Nugget Creek	St. Joe River	1	47 12 N	115 35 W
10	Prospector Creek	St. Joe River	1	47 13 N	115 56 W
11	Quartz Creek	St. Joe River	5	47 12 N	115 31 W
12	Simmons Creek	St. Joe River	4	47 08 N	115 24 W
13	South River Road	St. Joe River	6	47 14 N	115 59 W
14	Charlie Creek	St. Maries	6	47 03 N	116 31 W
15	Eldorado Gulch	Palouse River	10	47 01 N	116 31 W
16	Mannering Creek	Palouse River	4	47 03 N	116 40 W
17	Moscow Gulch	Palouse River	4	47 01 N	116 32 W
18	Breakfast Creek	North Fork Clearwater	1	46 53 N	115 57 W
19	Quartz Creek	North Fork Clearwater	4	46 48 N	115 27 W
20	Reeds Creek	North Fork Clearwater	2	46 37 N	115 59 W
21	Silver Creek	North Fork Clearwater	2	46 44 N	115 56 W
22	Grave Creek	Lochsa River	2	46 26 N	115 05 W
23	Macaroni Creek	Lochsa River	1	46 17 N	115 24 W
24	Squaw Creek	Lochsa River	2	46 30 N	114 52 W
25	Weir Creek	Lochsa River	9	46 28 N	115 02 W
26	Gedney Creek	Selway River	1	46 03 N	115 19 W
27	Glover Creek	Selway River	9	46 04 N	115 22 W
28	Meadow Creek	Selway River	5	46 02 N	115 17 W
29	Pearl Creek	S. Fork Salmon River	1	45 06 N	116 02 W
30	Roaring Creek	S. Fork Salmon River	10	44 45 N	115 41 W
31	Four Mile Creek	S. Fork Salmon River	2	44 59 N	115 40 W
32	Reegan Creek	S. Fork Salmon River	3	44 57 N	115 35 W