MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 5745–5761 doi: 10.1111/mec.12068

Defining evolutionary boundaries across parapatric ecomorphs of Black Salamanders (*Aneides flavipunctatus*) with conservation implications

SEAN B. REILLY,*† SHARYN B. MARKS* and W. BRYAN JENNINGS*‡

Department of Biological Sciences, Humboldt State University, 1 Harpst Street, Arcata, CA 95521, USA, ‡Museu Nacional, Departamento de Vertebrados, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ 20940-040, Brazil

Abstract

The accurate delimitation of evolutionary population units represents an important component in phylogeographic and conservation genetic studies. Here, we used a combined population assignment and historical demographic approach to study a complex of ecomorphologically distinctive populations of Black Salamanders (*Aneides flavipunctatus*) that are parapatrically distributed and meet at a three-way contact zone in north-western California. We used mitochondrial tree-based and multilocus clustering methods to evaluate a priori two- (Northern and Southern) and three (Northern, Coast and Inland) population hypotheses derived from previous studies. Mitochondrial results were consistent with the two- and three-population hypotheses, while the nDNA clustering results supported only the two-population hypothesis. Historical demographic analyses and mtDNA gene divergence estimates revealed that the Northern and Southern populations split during the Pliocene (2–5 Ma). Subdivision of the Southern population into Coast and Inland populations was estimated to be late Pleistocene (0.24 Ma), although our mtDNA results suggested a Pliocene divergence. Effective gene flow estimates ($2N_em$) suggest that either the two- or three-population hypotheses remain valid. However, our results unexpectedly revealed that the Northern population might instead represent two parapatric populations that separated nearly 4 Ma. These results are surprising because the Pliocene divergence between these ecomorphologically conservative forms is similar or older than for the ecomorphologically divergent Coast and Inland sister populations. We conclude that Black Salamanders in north-western California belong to at least three or four populations or species, and these all meet criteria for being Evolutionary Significant Units or 'ESUs' and therefore warrant conservation consideration.

Keywords: anonymous nuclear loci, demography, gene flow, mitochondrial DNA, population delimitation

Received 6 February 2012; revision received 30 August 2012; accepted 11 September 2012

Introduction

The proper delineation of reproductively cohesive populations is of paramount importance to studies in molecular ecology and conservation biology. For example, incorrectly defined populations in phylogeographic studies can be problematic because they can lead to biased, if not spurious, demographic parameter estimates (Beerli 2004; Slatkin 2005). In conservation biology, 'cryptic' populations threatened with extinction may be deprived of legal protection; conversely, ill-defined population units may receive unwarranted protection, thereby potentially diverting limited conservation resources (Avise 2000). Criteria are urgently

†Correspondence: Sean B. Reilly, Museum of Vertebrate Zoology and Department of Integrative Biology, University of California, 3101 Valley Life Sciences Building, Berkeley, CA 94720-3160, USA. Fax: +(510) 643 -8238;
E-mail: sbreilly@berkeley.edu

needed for objectively defining population units, as well as methodologies for their implementation.

One of the major challenges facing phylogeographers and conservation geneticists is that populations often do not occur as discrete entities—rather, they exist on a continuum of population differentiation resulting from varying levels of geographic isolation and gene flow (see Fig. 1 in Waples & Gaggiotti 2006). Although no single best criterion for defining 'evolutionary' populations yet exists, subpopulations exhibiting a lack of genetic cohesiveness with other subpopulations could be justifiably considered as distinct populations (Waples & Gaggiotti 2006). Indeed the demographic approach to elucidating population structure relies on quantitative assays of cohesiveness among subpopulations via estimates of N_em (i.e. product of the effective population size, N_e, and migration rate, m; Wright 1931). Although $N_em < 1$ is often used as the cut-off value for defining populations, this criterion may be too stringent as even $N_em < 25$ can still reflect statistically detectable departures from panmixia (Waples & Gaggiotti 2006). In contrast to the demographic approach, phylogeographers have typically used two nonmutually exclusive approaches for characterizing population structure: (i) using nongenetic data such as geographical area (e.g. island, habitat or sampling locality) or diagnosable morphological characteristics (e.g. color pattern) and (ii) using genetic data (e.g. mtDNA gene tree or multilocus clustering of individuals).

Although the demographic approach represents a more direct method for defining populations, owing to its focus on estimation of demographic parameters to infer population structure, in practice, this method is difficult to implement because decisions about which population units should be compared in the first place must be based upon prior information (e.g. morphological or molecular studies). On the other hand, tree- and cluster-based approaches are relatively easy to implement, but because they are pattern-based they lack key demographic insights into population structure. Nonetheless, the tree and cluster approaches can provide preliminary ideas about population structure, which can then be further evaluated in a demographic study.

Because populations also represent the fundamental units for biological conservation and management (Waples & Gaggiotti 2006), new approaches to defining populations may have implications for preservation of biodiversity. Indeed, one of the primary areas of conservation genetics concerns the identification of population-level genetic diversity that is the result of long-term historical isolation (Moritz 2002). Ryder (1986) introduced the concept of the Evolutionary Significant Unit or 'ESU' to describe a genetically distinctive population (or subspecies) that is worthy of conservation protection. The original criterion for recognizing ESUs simply required a finding of concordance between two independent data sets (e.g. distributional vs. molecular data). Later, Moritz (1994) suggested a more specific criterion: a finding of reciprocal monophyly in a mitochondrial gene tree together with significant divergence among a sample of nuclear loci. Given that some conceptual overlap exists between the evolutionary population concept (*sensu* Waples & Gaggiotti 2006) and the ESU concept, it would be interesting to know whether, in practice, populations defined using explicit historical demographic methods also meet ESU criteria.

The Black Salamander (*Aneides flavipunctatus* [Strauch 1870]; Plethodontidae) is a geographically variable species, which includes a complex of three ecomorphs that are parapatrically distributed at a three-way contact zone in north-western California (Lowe 1950). Although these north-western ecomorphs are considered a single species, previous studies (Lowe 1950; Larson 1980; Lynch 1981; Rissler & Apodaca 2007) suggest that *A. flavipunctatus* near this contact zone could be comprised of at least two or three populations. In this study, our goals are to: (i) assess how well these a priori two- and three-population hypotheses fit our population assignment results obtained from independent data sets (mtDNA vs. multiple nuclear loci) and methods (gene tree vs. allele-clustering); (ii) use historical demographic information from our multilocus coalescent analyses to determine whether these two- or three-population hypotheses are viable in the evolutionary population paradigm and (iii) assess whether or not our empirically supported evolutionary populations are also compatible with the ESU criteria of Moritz (1994).

Materials and methods

Study system

Aneides flavipunctatus inhabits three disjunct regions in north-western California: the Santa Cruz Mountains, the coast ranges from extreme southern Oregon south to Sonoma County, and Shasta County (Lynch 1981). Lowe (1950) used color pattern, microhabitat preference and forest habitat to subdivide *A. flavipunctatus* into three ecomorphs in the main part of the range, which hereafter will be recognized as three tentative 'populations': (i) A 'Coast' population from the southern coastal coniferous forest, with individuals that are black, or black with small off-white flecks, which is commonly found in streamside habitat; (ii) An 'Inland' population that has fewer, larger white spots and occurs in a more open habitat consisting primarily of oak and pine trees; and (iii) a 'Northern' population that possesses a green/grey frosted coloration, which is cryptic against the talus

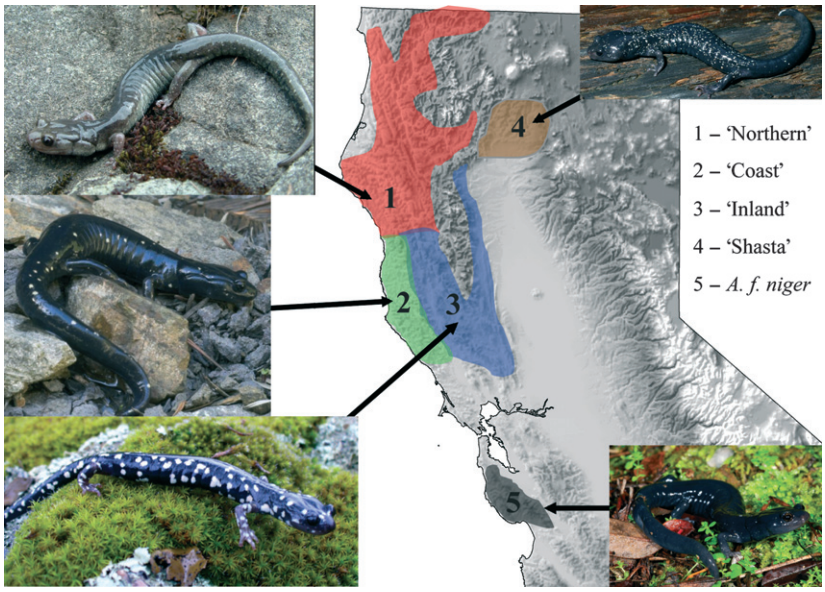

Fig. 1 Ecologically/morphologically distinct populations of *Aneides flavipunctatus* in northwestern California as described by Lowe (1950). Photos clockwise from upper right: G. Nafis, M. Mulks, S. Reilly, A. Gottscho, S. Reilly.

rocks that characterize its preferred microhabitat (Fig. 1). Lowe also identified a three-way contact zone between the parapatrically distributed Coast, Inland and Northern ecomorphs located in the Laytonville–Longvale area of Mendocino County (Fig. 1). Within this contact zone, individual salamanders can exhibit a variety of color patterns (Fig. S1, Supporting information).

Several studies provide additional evidence that significant evolutionary divergence exists among *A. flavipunctatus* in Mendocino County. Lynch (1981) found that salamanders north of the Laytonville–Longvale area retain juvenile coloration and body proportions into adulthood, whereas salamanders found just to the south lack these paedomorphic traits; furthermore, individuals north of the contact zone average 17 vertebrae, while those just south average 16 vertebrae. In a molecular study using allozymes, Larson (1980) found that the isolation between coastal and inland populations was greater than that between northern and southern populations in Mendocino County. More recently, Rissler & Apodaca (2007) conducted an mtDNA study of *A. flavipunctatus* and concluded that these salamanders comprise four major lineages (Shasta, Santa Cruz, Northwest and Central) and suggested that each of these lineages should be elevated to species status. The geographic boundaries of these lineages approximately correspond to Lowe's division, except that they grouped salamanders corresponding to Lowe's Coast and Inland ecomorphs together into a 'Central' lineage. The contact zone between the 'Northwest' and 'Central' lineages (*sensu* Rissler & Apodaca 2007) was placed ~30–140 km north of the Laytonville–Longvale zone described by Lowe (1950) and Lynch (1981). Although there is no consensus on the taxonomy of these salamanders or the placement of the Mendocino contact zones, the work to date provides compelling evidence for the existence of a major contact zone where at least two ('Northern vs. Southern') or three ('Northern', 'Coast' and 'Inland') populations have apparently diverged from each other. We formulated this two-population hypothesis based on the findings of Lynch (1981) and Rissler & Apodaca (2007), whereas the three-population hypothesis is based on Lowe's (1950) study.

Molecular data

Genetic sampling. Forty-six salamanders were sampled in the region near the three-way contact zone in Mendocino, southern Humboldt and northern Lake Counties, California. Four additional samples were obtained from the Museum of Vertebrate Zoology located in Berkeley, CA. Our rationale for this sampling approach is that if genetic isolation between populations has been achieved in or near the contact zone, then areas further away from the contact zone would likely be similarly or more isolated. All samples used and their collection locality data are presented in Table S1 (Supporting information). Genomic DNA was extracted from tissues using the DNeasy kit (Qiagen, Valencia, CA, USA).

Mitochondrial loci. We sequenced two mitochondrial loci including *ND4* (and associated tRNAs) and *Cytochrome B (Cytb)* genes. PCR amplification followed standard procedures and was performed using primers listed in Table S2 (Supporting information). PCR products were sequenced in both directions using an ABI3730 sequencer (Applied Biosystems, Foster City, CA, USA). Raw

sequence reads were combined in Codoncode Aligner 3.5.2 (CodonCode Corporation, Dedham, MA, USA). *ND4* sequences from three *A. flavipunctatus* (GenBank accession no's AY274624, AY274642 and AY274665) and one *Aneides lugubris* (GenBank accession no. AF329329) were downloaded from GenBank. *A. flavipunctatus niger* from Santa Cruz Co., CA (MVZ 264073 and MVZ 170983) and *A. lugubris* were used as outgroups to root mtDNA gene trees. Mitochondrial loci were concatenated, and sequence alignments were accomplished using CLUSTAL X (Larkin *et al.* 2007) with manual corrections.

Anonymous nuclear loci. Using genomic DNA from an individual *A. flavipunctatus* (Af_Voucher, Table S1, Supporting information), we followed the methods outlined in Jennings & Edwards (2005) for generating a set of primers that amplify anonymous loci ranging from 200 to 700 base pairs in length. All new loci were subjected to a BLAST analysis to determine whether any locus was homologous with any known genomic entities. The primer sequences and expected PCR product sizes for 12 anonymous nuclear loci are in Table S2 (Supporting information).

Intron locus. We included one intron locus, a 539-bp fragment of the protein-coding proopiomelanocortin gene (*POMC*) (Table S2, Supporting information). PCR and sequencing methods were the same as described earlier.

Nuclear haplotype sequences. We obtained nuclear haplotype sequences for each PCR product using either the TOPO TA cloning and sequencing method (Invitrogen, Carlsbad, CA, USA) or via a bioinformatics approach using the software program PHASE v2.1 (Stephens *et al.* 2001; Stephens & Scheet 2005). For each individual, the most probable pair of allele sequences (probability cutoff 0.50) was used in all subsequent analyses.

Data analyses

Data characteristics. We used DNAsp (Rozas *et al.* 2003) to calculate the following summary statistics for each locus: variable sites, parsimony informative sites, number of haplotypes, haplotype diversity, nucleotide diversity, %GC, and Tajima's D. Tajima's D (Tajima 1989) compares an estimate of θ based on site heterozygosities with another estimate of θ based on segregating sites; loci that evolved under neutral evolution will show a Tajima's D statistic of zero (Gillespie 2004). Nonsignificant results are consistent with a locus being neutral and having stable population size, whereas a positive D reflects a locus that experienced balancing selection or a population reduction and negative D indicates purifying selection or population expansion (Gillespie 2004; Yang 2006).

Intralocus recombination. Our demographic analyses assume no intralocus recombination (Hey & Nielsen 2004); therefore, we analysed our nuclear loci using the program IMgc (Woerner *et al.* 2007). This program produces the largest nonrecombining block of DNA sequence and maximizes the information content of the final data set by equally weighting individual sequences and segregating sites.

MtDNA gene tree estimation. We conducted two different Bayesian analyses using BEAST v1.6.1 (Drummond & Rambaut 2007) and MrBayes v3.1.2 (Ronquist & Huelsenbeck 2003). Our BEAST analysis consisted of two separate runs of 100 000 000 generations, sampling every 1000 generations for a total of 100 000 saved trees per run. These two runs were pooled in LogCombiner after discarding the initial 75 000 saved trees. The remaining 50 000 trees were used to create a maximum clade credibility tree. BEAST also provides estimates of gene divergence times using a relaxed exponential clock model. We used a calibration of 0.8% divergence per million years, which had been used for the salamander *Taricha torosa* (Tan & Wake 1995). We used TRACER v1.5 to assess convergence by examining the ESS values for each parameter and the trace of the posterior values. In the MrBayes analyses, two separate analyses were run with four chains for 20 million generations, sampling every 1000 generations to produce 20 000 trees per run. The first 10 000 trees of each run were discarded as burn-in, and the remaining 20 000 trees were used to produce a 50% majority-rule consensus tree. Clade support was assessed using posterior probabilities.

We also estimated mtDNA gene trees using maximum likelihood (ML) as implemented in RAXML (Stamatakis *et al.* 2005) and GARLI (Zwickl 2006). The GARLI analysis was run only with samples for which sequences for both mtDNA genes were obtained (except outgroups). Clade support was provided by 1000 bootstrap replicates.

Nuclear gene trees. Although the main purpose for acquiring our nDNA data set was for conducting population assignment and historical demography analyses, we also examined the reconstructed gene trees for each nuclear locus, which were inferred using GARLI with mid-point rooting.

Population assignment. Two types of data and analytical methods were used to genetically corroborate or reject the 'two-population hypothesis' ('Northern' and 'South-

ern') or 'three-population hypothesis' ('Northern', 'Coast' and 'Inland'). Note that the Northern population is equivalent in both hypotheses and that the Southern population is equal to the sum of the Coast and Inland populations.

The first method involved using our mtDNA tree to test our population hypotheses by looking for exclusive clusters of haplotypes that correspond with either or both the two- and three-population hypotheses. Note that the criterion of 'exclusivity' of haplotypes for a predefined population can be satisfied either by bifurcation groups in an unrooted tree or clades in a rooted tree (Brower 1999).

The second method used STRUCTURAMA (Huelsenbeck & Andolfatto 2007), which is a multilocus clustering program that assigns nDNA haplotypes to populations. A particularly nice feature of this program is that by running it with the population number specified as unknown, we are in effect letting the program objectively decide how many populations exist. We ran the program with the number of populations fixed at two, three, four and also with the number of populations set as a random variable. Each analysis was run for five million generations, sampling every 100 generations for a total of 50 000 saved steps. We discarded the first 20 000 steps as burn-in, and the remaining steps were summarized to assign individuals to populations.

Estimation of demographic parameters. We used our 13 nuclear loci and the programs IM (Hey & Nielsen 2004) and IMa2 (Hey 2010) to analyse the two- and three-population hypotheses, respectively. These programs apply the Isolation with Migration model to genetic data obtained from closely related populations or species to estimate the marginal posterior probability densities for each of the model parameters. For each divergence event, six population parameters (scaled by mutation rate u) were estimated: sizes of the ancestral (θA) and daughter populations (θ1, θ2); population divergence time (t); migration rate/gene/generation (m1), which is defined as the rate genes come into population 1 from population 2; and (m2) for migration in the opposite direction.

For the IM analyses, several were completed to determine the appropriate prior boundaries for each parameter and to assess convergence before completing one final run for each population comparison. Each run consisted of six chains, a 2.5-million step burn-in, and five-million steps. The program IMa2 requires that the user specify the population (or species) tree before initiating the analysis. Of the three possible population trees, we chose the tree topology having the relationships [(Coast, Inland), Northern], which is supported by Lynch (1981) and by our mtDNA gene tree (see Results). After we determined the priors and assessed convergence as before, a final run consisting of a 10-million step burn-in and 10-million steps sampling every 100 steps for a total of 100 000 saved steps was completed.

Model parameters were converted into demographic and time units following Hey (2005). We assumed a mutation rate of 2.2×10^{-9} substitutions/site/year (Kumar & Subramanian 2002), and multiplied this rate by the number of base pairs in each locus before using the geometric mean of our 13-locus mutation rate (8.16×10^{-7} substitutions/year) to perform parameter conversions. We assumed a generation time of 3 years until sexual maturity, which is based on the estimate for *A. lugubris* (Anderson 1960).

To better understand our population assignment results, we used our IM and IMa2 effective population size and gene flow estimates to calculate $2N_em$ values between each hypothetical population. The $2N_em$ value is defined as the population migration rate, which is the effective rate at which genes come into a population, per generation. Given that the evolutionary consequences of effective gene flow varies between two extreme conditions—genetically isolated populations and a single panmictic one—a criterion for defining separate populations can be based on $2N_em$ values that reflect a departure from panmixia. We therefore chose $2N_em \leq 25$ as our criterion because gene flow, even this large, can be associated with departures from panmixia (Waples & Gaggiotti 2006). For descriptive purposes, we classify $2N_em$ values in terms of their relative strength of genetic isolation: *strong* when $2N_em \leq 1$ (Wright 1931); *moderate* when $1 < 2N_em \leq 5$ and *weak* when $5 < 2N_em \leq 25$ (Waples & Gaggiotti 2006).

Results

Data characteristics

Loci and sequence characteristics. Mitochondrial *Cytb* and *ND4* loci were 387 and 735 bp, respectively, twelve anonymous loci ranged from 188 to 639 bp (average 447 bp), and the POMC intron was 481 bp (Table 1). MtDNA loci exhibited higher levels of nucleotide variation (165 variable sites/1000 bp for *Cytb* and 158 variable sites/1000 bp for *ND4*) than either the anonymous loci (18–250 variable sites/1000 bp, average of 70) or POMC intron (37 variable sites/1000 bp; Table 1). All nuclear loci, except *POMC*, had low (< 50%) GC content, and our BLAST results indicated that none of our twelve anonymous loci matched known genomic regions. Tajima's D tests indicated that both mtDNA loci and six of 12 anonymous loci had D values that were not significantly different than the null expectation of zero; the other six anonymous loci, plus *POMC*, exhibited negative D values that were significantly

Table 1 Summary statistics for the loci used in this study

Locus	No. of bp	No. Sequences	*h*	Haplotype diversity	Variable sites (not including gaps)	Variable Sites/1000 bp	Parsimony informative sites	Tajima's D	%GC
ND4	735	51	38	0.987	116	158	69	−0.973	0.374
Cytb	387	46	35	0.985	64	165	43	−0.301	0.347
SR1	478	92	19	0.799	22	46	6	−1.850*	0.372
SR2	357	62	22	0.921	32	90	22	0.214	0.442
SR4	271	83	6	0.360	5	18	5	−1.239	0.366
SR7	565	89	26	0.817	28	50	13	−1.991*	0.496
SR8	639	72	44	0.964	47	74	18	−1.859*	0.395
SR9	188	71	13	0.654	47	250	44	0.679	0.428
SR12	435	90	32	0.920	26	60	16	−0.662	0.486
SR15	526	89	36	0.949	21	40	19	−0.484	0.373
SR16	273	91	14	0.640	14	51	9	−1.867*	0.482
SR17	455	95	41	0.929	26	57	15	−1.906*	0.446
SR19	621	88	38	0.896	40	64	21	−2.419*	0.376
SR20	550	81	12	0.765	14	25	5	−1.480	0.351
POMC	481	94	18	0.729	18	37	11	−1.870*	0.550

h, number of haplotypes.
*Tajima's D value that is significantly different than the null expectation of zero ($P < 0.05$).

lower than the null expectation of zero (Table 1). Data coverage information can be found in Table S3 (Supporting information).

Population assignment: MtDNA results

MtDNA gene tree. MtDNA analyses, which were based on a total 1122 bp, produced trees having similar topologies. In the BEAST tree (Fig. 2A), three monophyletic groups provide evidence consistent with our three-population hypothesis, whereas a more basal split in this tree yields two sister clades that match our two-population hypothesis. Gene divergences between Northern vs. Southern and Coast vs. Inland clades were estimated to be ~5 and 4.4 Ma, respectively. The uncorrected pairwise sequence divergences ranged from 3 to 5% (Table 2). Interestingly, the Northern clade is subdivided by a gene divergence event that occurred about 4.3 Ma. We shall refer to these two divergent daughter clades as the 'Main Fork Eel River' (= MFER) clade (located near the contact zone area) and 'South Fork Eel River' (= SFER) clade (found to the north). These MFER and SFER clades have 23 fixed base mutations out of 1122 bp between them, which exceeds the other population comparisons (Table 2).

Our MrBayes tree (Fig. 2B) is partially congruent with the BEAST tree. Monophyletic groups reflecting our hypothetical populations were only observed for the Northern and Coast populations. However, note that if the root position is shifted a few nodes, then both the Inland and the more encompassing Southern clade both become monophyletic groups and would therefore match the BEAST tree topology.

The two ML analyses produced similar although partially unresolved topologies (Figs. 2C, D). In both trees, a basal polytomy is observed consisting of three divergent monophyletic lineages: two separate Northern clades (i.e. the MFER and SFER clades) and the Southern clade. Although a sister group relationship between Northern and Southern clades is ambiguous, we did obtain evidence from both trees of a monophyletic Southern clade. Thus, our two ML trees provide results consistent with the existence of Southern, Coast, Inland, SFER and MFER populations.

When the topologies of these four mtDNA trees are compared, we observe the following: a Coast clade was found in all four trees; a Southern clade consisting of nested sister Coast and Inland clades was found in three of four trees; a Northern clade was found in two trees; and all four trees contained MFER and SFER clades. When visualized in an unrooted format (not shown) or as a majority-rule consensus tree of our rooted trees (equivalent to Fig. 2A), all four trees have topologies compatible with our two- and three-population hypotheses. Examination of mitochondrial haplotype distribution around the contact zone not only illustrates concordance with the two- and three-population hypotheses but also reveals fine-scale geographic structuring of the MFER and SFER haplotypes (Fig. 3).

Population assignment: nuclear DNA results

The inferred topologies for the thirteen nuclear gene trees revealed little correspondence to our hypothesized populations (see Figs S2–4, Supporting information).

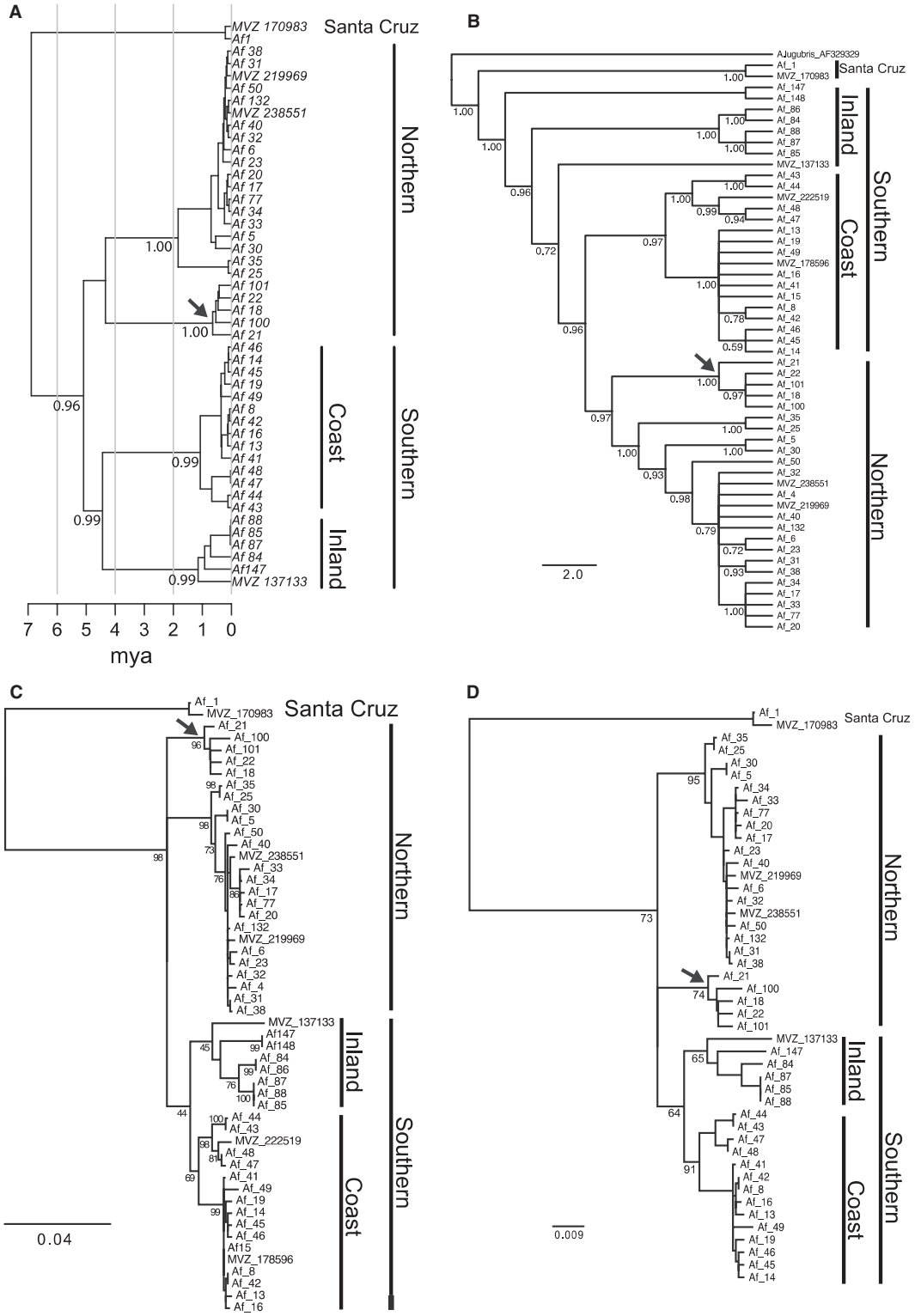

Fig. 2 Mitochondrial DNA phylogenies of the *ND4* and *Cytb* genes. Arrows point to the Main Fork Eel River clade. (A) BEAST analysis. Numbers along bottom scale represent millions of years from present, and numbers below nodes represent Bayesian posterior probabilities. (B) MrBayes 50% majority-rule consensus tree. Numbers below nodes represent Bayesian posterior probabilities. (C) RAXML analysis. Numbers at nodes represent bootstrap support. (D) GARLI analysis. Numbers at nodes represent bootstrap support.

Table 2 F_{st}, fixed nucleotide differences, and uncorrected sequence divergence of mtDNA between populations

	ND4			*Cytb*		
	F_{st}	No. of fixed differences	Average seq. div.	F_{st}	No. of fixed differences	Average seq. div.
Coast vs. SFER	0.716	9	0.033	0.700	8	0.046
Coast vs. Inland	0.510	4	0.030	0.604	4	0.041
Coast vs. MFER	0.662	11	0.035	0.803	8	0.050
SFER vs. MFER	0.757	15	0.036	0.718	8	0.045
SFER vs. Inland	0.748	13	0.044	0.592	5	0.048
Inland vs. MFER	0.660	13	0.041	0.662	8	0.047
Northern vs. Southern	0.499	5	0.036	0.475	1	0.048

MFER, Main Fork Eel River; SFER, South Fork Eel River.

When these nuclear loci were analysed using STRUCTURAMA, the results from the variable population setting supported a three-population model (Table S4, Supporting information). However, further examination shows that a two-population model, comparable to the Northern and Southern lineages from the mtDNA gene trees, is more biologically plausible, because this 'third' population is represented by a single individual (Af_23). The variable setting yielded similar results as when the population number was preset at two. When the number of populations is fixed from 3 to 4, the program failed to identify any geographically structured populations. These analyses also suggested that three MFER salamanders are more closely related to Southern lineage salamanders rather than other Northern lineage salamanders. Figure 4 compares the results from the two assignment analyses.

Historical demography

Because our mtDNA and nDNA assignment analyses yielded results consistent with both the three- and two-population hypotheses, we estimated historical demographic parameters for both evolutionary scenarios.

Populations defined for IMa2 analyses. Prior to conducting the IMa2 analyses, assignment of individual salamanders following a three-population scheme was straightforward because our mtDNA results matched Lowe's (1950) morphology/geography data.

Populations defined for IM analyses. For our IM analyses of a two-population scheme, we constructed three different data sets because of minor variation in our population assignment results (Fig. 4). The first data set,

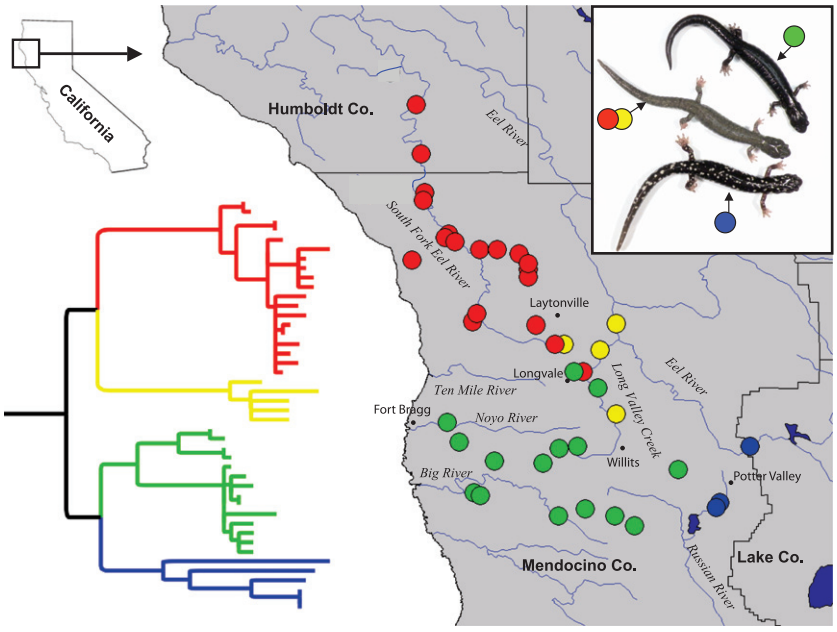

Fig. 3 Geographic placement of mitochondrial haplotypes by clade. Red = South Fork Eel River (SFER) clade, Yellow = Main Fork Eel River (MFER) clade, Green = Coast clade, Blue = Inland clade. Note, SFER + MFER = Northern haplotype group. (photograph: D. Portik).

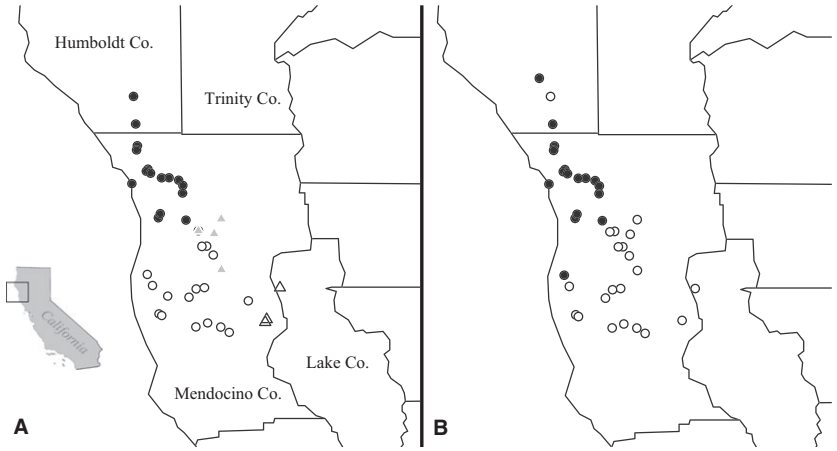

Fig. 4 Comparison of population assignment techniques. (A) Mitochondrial clade groupings: Black circle = South Fork Eel River (SFER), Grey triangle = Main Fork Eel River (MFER), Open circle = Coast, Open triangle = Inland (note, SFER + MFER = Northern haplotype group); (B) STRUCTURAMA two-population model groupings: Black circle = Northern, Open circle = Southern.

informed by our mtDNA results, comprised nuclear haplotypes from the thirteen nuclear loci that corresponded to the Northern and Southern mtDNA haplotype groups. The second data set was based only on our STRUCTURAMA results because eight individuals were assigned differently compared with the mtDNA results, seven of which were located close to the contact zone. One individual from Humboldt County (Af_132) is apparently an outlier, because it is grouped with southern Mendocino County salamanders. Because of the extreme geographic distance between Af_132 and the other individuals with which it was grouped, we assigned Af_132 to a geographically proximate salamander group. Lastly, we constructed a third data set (termed 'two-population pure'), which excluded haplotypes from salamanders assigned to one population using mtDNA and another using nDNA.

Results from IMa2 and IM analyses. Multiple independent IMa2 runs produced comparable parameter estimates. In the three-population analysis, the Inland population has a lower population size (59 000 individuals) than the Coast (580 000) and Northern (483 000) populations (Table 3; Fig. S5, Supporting information). In the two-population (IM) analyses, the estimates for each of the three data sets were similar to each other, as estimates for the Southern population size ranged from 236 000 to 269 000 individuals, whereas the Northern population estimates ranged from 248 000 to 349 000 individuals. The size of the population ancestral to the Coast and Inland populations was estimated to be 88 000 individuals; however, the size of the ancestral population for the Northern and Southern populations could not be reliably estimated in either the Ima2 or IM analyses (Table 3; Fig. S6, Supporting information).

Our results suggest the Coast and Inland populations diverged from each other ~240 000 years ago (with a 95% C.I. of 130 000–430 000 years ago; Table 3; Fig. S7, Supporting information). Population divergence between the Northern vs. Coast/Inland ('Southern') populations occurred ~2.24 Ma (with a 95% C.I. of 1.15–6.02 Ma). For the two-population scheme, our estimates were even older, ranging from 4.7 to 5.2 Ma, with a 95% C.I. of about 3.1–9.4 Ma.

The $2N_em$ values for the Coast vs. Inland and Northern vs. Inland population comparisons were all < 1.0, suggesting a strong level of genetic isolation between these populations in the time since initial divergence (Fig. 5), whereas a moderate level of isolation apparently exists between the Northern vs. Coast populations ($2N_em \sim 2$ in each direction). The posterior probability densities for the three contemporary population size parameters (Fig. S5, Supporting information) and six migration parameters (Figs S8 and S9, Supporting information), upon which the $2N_em$ values are based, indicate that these are all robust $2N_em$ estimates. The $2N_em$ values for the ancestral Coast–Inland population vs. the Northern population suggest asymmetric gene flow in the past resulting in a strong amount of genetic isolation for the southern populations, whereas the Northern population is weakly isolated from the southern ones.

Our two-population analyses detected gene flow moving northwards across the contact zone at a greater rate than southwards gene flow, a result that holds regardless how we conducted our population assignments (Fig. 5). Moreover, we observed negligible variation among $2N_em$ estimates describing southwards gene flow but observed more substantial variation among values for northward gene flow (Fig. 5). Our estimates for the effective size of the Northern population were consistent in our two-population analyses; therefore, variation in $2N_em$ values is due to the estimates of the migration parameter alone. This indicates that our results for this parameter were sensitive to how we assigned salamanders to the Northern and Southern populations.

Table 3 Converted demographic parameters of the IMa2 and IM analyses

Pop. assign scheme	Value	Divergence time (Ma)	Effective population size (millions of individuals)						Population migration rate ($2N_em$)							
		*t*0	*t*1	*q*0	*q*1	*q*2	*q*3	*q*4	*m*0 > 1	*m*1 > 0	*m*0 > 2	*m*2 > 0	*m*1 > 2	*m*2 > 1	*m*2 > 3	*m*3 > 2
3 Pop. mtDNA	HiPt	0.24	2.24	0.580	0.059	0.483	0.088	1.531*	0.27	<0.01	2.06	2.02	0.12	0.01	22.18	0.03
	HPD95Lo	0.13	1.15	0.310	0.019	0.329	0.019	0.604	0.00	0.00	0.00	0.00	0.00	0.00	8.60	0.00
	HPD95Hi	0.43	6.02	1.371	0.125	0.690	0.151	1.531*	7.21	3.03	17.48	9.17	2.63	1.84	–*	–*
		t		*q*0	*q*1	*q*A			*m*0 > 1	*m*1 > 0						
2 Pop. mtDNA	HiPt	4.71	–	0.269	0.334	–*	–	–	<0.01	7.26	–	–	–	–	–	–
	HPD90Lo	3.09	–	0.216	0.247	–*	–	–	<0.01	4.08	–	–	–	–	–	–
	HPD90Hi	8.01	–	0.332	0.449	–*	–	–	0.36	12.83	–	–	–	–	–	–
2 Pop. nDNA	HiPt	4.95	–	0.264	0.349	–*	–	–	0.06	8.21†	–	–	–	–	–	–
	HPD90Lo	3.07	–	0.207	0.256	–*	–	–	<0.01	<0.01	–	–	–	–	–	–
	HPD90Hi	9.09	–	0.332	0.461	–*	–	–	8.13	14.74	–	–	–	–	–	–
2 Pop. Pure	HiPt	5.23	–	0.235	0.248	–*	–	–	<0.01	3.65	–	–	–	–	–	–
	HPD90Lo	3.12	–	0.179	0.175	–*	–	–	<0.01	<0.01	–	–	–	–	–	–
	HPD90Hi	9.39	–	0.299	0.325	–*	–	–	4.36	6.98	–	–	–	–	–	–

For the three-population scheme, 0 = Coast, 1 = Inland, 2 = Northern, 3 = Coast/Inland ancestor, 4 = Northern/Southern ancestor, *t*0 = Coast/Inland divergence time, *t*1 = Northern/Southern divergence time, *q* = effective population size, *m* = population migration rate (example: *m*0 > 1 is the effective rate at which genes come into the Coast population from the Inland population, per generation). For the two-population scheme 0 = Southern, 1 = Northern, A = Southern/Northern ancestor.

*Parameter unable to be estimated or is influenced by the prior distribution.

†Value for a second peak.

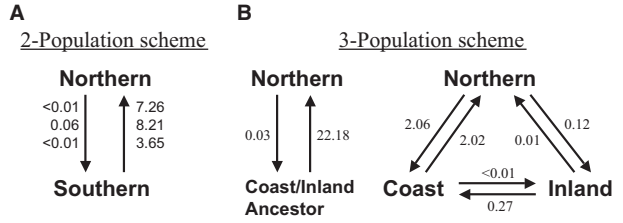

Fig. 5 Summary of estimated $2N_em$ values: (A) for the two-population scenario, using three different approaches to population assignment (from top to bottom: mtDNA, nDNA, and 'two-population pure'), (B) for the three-population mtDNA scenario.

Our mtDNA results revealed a deep divergence within the Northern population, which prompted us to conduct a preliminary IM analysis of our nDNA data set comparing only the MFER and SFER mtDNA haplotype clade groups (see Table S5, Supporting information). Indeed, the results suggest a considerable population divergence time of 3.7 Ma, negligible southwards gene flow ($2N_em$ = 0.17) and some northwards gene flow ($2N_em$ = 6.78). Sizes of the MFER and SFER populations were estimated to be ~136 000 and 302 000 individuals, respectively.

Discussion

Population assignment: perspectives from mtDNA and nDNA

Concordance between morphological and genetic variation is compelling because it reinforces the notion that morphological and/or molecular data can shed light on population structure (Zink 1994; Avise 2000; Blondel *et al.* 2006; Rojas-Soto *et al.* 2010; Dudaniec *et al.* 2011). Indeed, the sensitivity of mtDNA as an indicator of population structure became evident in this study, as our trees showed concordance with Lowe's (1950) three variably colored ecomorphs and thereby supported the three-population scheme. However, due to the nested nature of our a priori population schemes, this same mtDNA tree could also be presented as evidence for a two-population scheme, in agreement with data from another phenotypic study (Lynch 1981). Our clustering analysis of 13 nDNA loci also supported the two-population scheme. Unexpectedly, our mtDNA analysis detected a deep divergence among haplotype clades in the Northern population, which was not reflected in any of our nDNA trees or from the nDNA clustering analysis. We discuss the significance of this latter observation below.

Another discrepancy between the mtDNA vs. nDNA assignment results concerns the precise position of the contact zone. Along the coastline, the nDNA contact zone lies at the Noyo River, ~25 km (due to a sampling gap) south of the mtDNA contact zone. Further inland near Longvale, the nDNA contact zone is ~5 km north of the mtDNA break (Fig. 4). The Northern and Southern mtDNA clades come into contact with each other in the area between Laytonville and Willits (just south of Longvale), as Lowe (1950) and Lynch (1981) had postulated. Our north/south mtDNA break is located at least 30 km south of the region identified as such by Rissler & Apodaca (2007). The Coast and Inland clades contact each other just west of Potter Valley, and the Inland and Northern clades apparently contact each other along the MFER between Lake Pillsbury and Long Valley Creek. Finding the precise location of contact zones is important for at least two reasons. First, an improved understanding of the speciation process could be gained from study of accurately mapped contact zones. Second, contact zones may be important for conservation purposes because gene flow across these population boundaries can maintain or enhance genetic diversity, thereby providing populations with genetic substrate for adapting to changing environmental conditions (Moritz 2002).

Interestingly, five salamanders sampled near the contact zone had Northern mtDNA haplotypes, yet these individuals were assigned to the Southern population in the nDNA clustering analysis. This pattern would be expected if the mtDNA lineage boundary remained static while nuclear gene flow moved from south to north, a signal our migration estimates recovered (see below). One mechanism that would promote this pattern would be male-biased northward dispersal, where males disperse farther than females and thus bring new nuclear genes into an area but not new mtDNA haplotypes (Galbreath *et al.* 2011; Toews & Brelsford 2012). Such a dispersal bias was shown in *Ensatina* salamanders, in which males disperse nearly twice as far as females (Staub *et al.* 1995). The pattern of unidirectional gene flow could also be influenced by asymmetric hybridization, which was documented in another study on *Ensatina* (Devitt *et al.* 2011).

Historical demography

The large effective population size for the Coast population is consistent with observations of high local densities of salamanders in the field (S. B. Reilly, personal observation). This population occupies extensive and continuously distributed coastal forests, much like the situation for the comparably sized (in terms of N_e) Northern population. The lower effective population size for the Inland population is in accord with field observations of relatively low population densities (S. B. Reilly, personal observation). Inland salamanders inha-

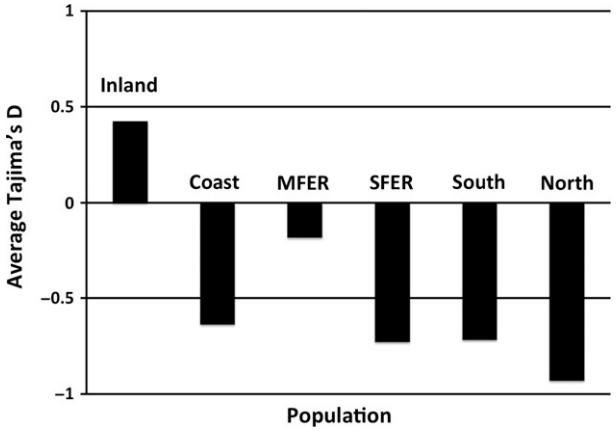

Fig. 6 Average values of Tajima's D for the population units considered in this study (see Table S6, Supporting information for actual *D* values).

bit a region governed by a hot and dry climate, where they occur in scattered rock piles that are often surrounded by areas of less favourable open grassland and oak/pine woodland. The size of the ancestral effective population size suggests that the Inland population has remained similar in effective size since the time of divergence, whereas the Coast population underwent a tenfold expansion. This agrees with results from the Tajima's D tests, which yielded negative D values for nearly all nuclear loci. Assuming these loci are neutral, we interpret our findings as evidence for recent population expansion. Indeed, Tajima's D calculated for all candidate populations (Fig. 6; Table S6, Supporting information) reveals trends consistent with the ecology and size of each population: Inland salamanders, which live in relatively unfavourable environments and exist in low population densities, may be declining (positive D); Coast and Northern populations, which live in favourable areas and occur in higher densities, might be growing (negative D); and lastly, the SFER population (found in mesic coastal areas) shows a more negative D value than the MFER population (found in the drier inland areas).

Both population scenarios predict a genetic break between a Northern population and either a single Southern population or two parapatric southern populations. This hypothetical north–south break is further corroborated by our divergence time estimates, which ranged between 2.2 and 5.2 Ma, suggesting that the ancestral population was subdivided into northern and southern populations during the Pliocene epoch (2.0 –5.3 Ma). This result holds regardless of how we conducted our assignment analyses and is consistent with our mtDNA gene divergence times (~5 Ma). Our three-population analysis yielded a population divergence time between the Coast vs. Inland populations of 240 000 years ago (middle-to-late Pleistocene), whereas our mtDNA estimate suggests an older divergence time (~4.4 Ma). Such a discrepancy between divergence estimates could be explained by at least two causes, which are not mutually exclusive: (i) an overabundance of ancestral polymorphisms in our mtDNA relative to the time since population divergence (Edwards & Beerli 2000) and (ii) improper calibration for mtDNA and/or nDNA (Lee & Edwards 2008; Carling *et al.* 2010). Despite these substantial divergence times, the gene flow history is needed to assess the degree of reproductive cohesion among these tentative populations (Waples & Gaggiotti 2006).

Although we observed some variation in our $2N_em$ estimates (Fig. 5), which may be attributable to how we conducted our population assignments and demographic analyses, the general picture supported by these results does not refute the two- or three-population hypotheses. However, a growing body of evidence suggests that the latter hypothesis may better represent the contemporary populations of *A. flavipunctatus* in our study. First, the lowest $2N_em$ values in the three-population analysis suggest that the Inland salamanders are genetically isolated from their relatives to the west and north, a finding that is supported by the allozyme findings of Larson (1980). Secondly, our mtDNA gene and multilocus population divergence estimates propose that the ancestral Southern population diverged into two daughter populations either during the Pliocene or in the middle-to-late Pleistocene, respectively. Lastly, the morphological and ecological distinctiveness of the Inland salamanders further corroborates their uniqueness at the population, if not species, level.

Assumptions of the isolation with migration model

Our demographic parameter estimates may depend on the assumptions of the IM model, which include loci independently assort in meiosis, loci are selectively neutral, recombination occurs only among loci and not within them, populations are unstructured and no genetic contributions from unsampled populations (Hey & Nielsen 2004; Strasburg & Rieseberg 2010). In view of this, we elected to use a modest sample of anonymous loci because these markers have properties that likely enable them to meet the IM assumptions.

First, because anonymous loci are presumably scattered randomly throughout the genome, they are probably unlinked from each other. Moreover, owing to their random genomic locations, such loci are also likely to be selectively neutral because recent whole-genome studies have revealed that ~90% of vertebrate genomes consist of nonfunctional DNA (Meader *et al.* 2010; Ponting & Hardison 2011). Although introns likely experi-

ence little or no natural selection, these loci can lose some genetic diversity via genetic hitchhiking (Lee & Edwards 2008; Thomson *et al.* 2010). Although nuclear loci can recombine via crossing over, the potentially confounding effects of this evolutionary force can be ameliorated by using methods (e.g. the 'four gamete test'; Hudson & Kaplan 1985) to detect historical recombination and then truncating any putatively recombined locus into the longest unrecombined sequence block (e.g. Jennings & Edwards 2005; Woerner *et al.* 2007). Lastly, a recent simulation study by Strasburg & Rieseberg (2010) suggests that the IM model may be robust to violations of the nonstructured population assumption. Moreover, they also found that the divergence time and gene flow parameters in the IM model would experience only minor biases when unsampled populations occupy a small extent of the periphery of focal populations, which is the case in our study (i.e. at the extreme northern and southern range limits of our study populations).

Population assignment revisited: MtDNA vs. nuclear DNA

The recent emergence of studies using large multilocus data sets has spawned a debate over the relative merits of mtDNA vs. nDNA markers for revealing phylogeographic patterns and processes (Edwards *et al.* 2005; Zink & Barrowclough 2008; Edwards & Bensch 2009). Although theoretical reasons exist for believing that large numbers of independent nuclear markers will outperform mtDNA in estimating historical demographic parameters, disagreement persists regarding whether mtDNA or nDNA should be used to define populations (Zink & Barrowclough 2008; Edwards & Bensch 2009).

Our mtDNA results are consistent with both of our *a priori* population schemes. Thus, in this situation the choice between schemes is subjective, illustrating the arbitrary nature of tree topology-based population delimitation practices (Avise 2000). Without demographic information, especially gene flow history, the mtDNA tree topology cannot, by itself, specify the contemporary population boundaries. Consequently, the more objective multilocus clustering approaches would seem to be the ideal solution to our population delimitation problem. However, in our study the clustering analysis apparently lacked the resolving power necessary to delimit at least three populations of *A. flavipunctatus*. This result is not surprising given a simulation study suggesting that use of a small sample (e.g. 10 or fewer) of 'low mutation' nuclear loci presents challenges to identifying the true number of populations (Waples & Gaggiotti 2006). Thus, our use of an insufficient number of anonymous loci could have limited our ability to recognize three populations. Accordingly, we predict that using a larger number of low mutation nuclear loci (dozens or more) or a modest number of high mutation loci (e.g. 8–20 microsatellite loci) would discriminate at least three populations.

Our clustering analysis not only apparently failed to recognize three populations, but it may have also missed a deep and hitherto unknown genetic division within the Northern population. Indeed the 4% mtDNA divergence between these clades is comparable to species-level divergences reported for other plethodontid salamanders (Rissler & Taylor 2003; Rovito 2010). Furthermore, results from our IM analysis suggest that the mtDNA-defined 'populations' (i.e. MFER and SFER individuals) diverged nearly 4 Ma, and that there has been negligible to low northwards-only gene flow between them since divergence. This result is particularly surprising because salamanders in the Northern clade exhibit little ecomorphological variation (although SFER salamanders are darker than the more intensely frosted MFER salamanders; S. B. Reilly, personal observation), whereas individuals from the Coast and Inland lineages show strikingly divergent color patterns and ecologies. Thus, ecomorphological divergence was a strong predictor of genetic divergence between the Coast and Inland populations, but not between the MFER and SFER candidate populations. That the MFER and SFER entities represent actual populations or species needs to be corroborated in future studies.

Although mtDNA and nDNA data may exhibit discordance (see Toews & Brelsford 2012 for review), these independent genomic assays can agree in other ways. For example, despite incongruence between the locations of mtDNA and nDNA contact zones subdividing Black Salamanders into northern and southern groups, our mtDNA gene tree still exhibited reciprocal monophyly suggesting that past mtDNA gene flow, if it occurred at all, did not spread far beyond the contact zone region. Reciprocal monophyly was also observed for Coast vs. Inland mtDNA haplotypes suggesting that reproductive cohesion has largely been lost between them as well. These findings were corroborated by our multilocus population divergence and effective gene flow estimates.

It is noteworthy that our empirically supported evolutionary populations (Coast, Inland, Northern) as well as the two newly described candidate populations (MFER, SFER) fit criteria for ESUs, because this suggests that ESUs might be useful indicators of evolutionary populations (*sensu* Waples & Gaggiotti 2006). However, the detection of ESUs in this study may have been facilitated by the highly structured populations of *A. flavipunctatus*, which are typical of low vagility organisms such as terrestrial salamanders (Larson *et al.* 1984; Mo-

ritz 1994). In contrast, ESUs may be difficult to detect in cases of rapid speciation or hybridization because mtDNA will likely not show reciprocal monophyly in those cases (Moritz 1994). Given these potential limitations of current ESU criteria, future efforts directed at developing specific criteria for different levels of reproductive cohesion (i.e. specific values of $2N_em$) seem to offer the best hope for determining whether the estimated evolutionary boundaries for a given 'species' are representative of a single panmictic population, separate evolutionary populations, or even separate species.

A major phylogeographic break in north-western California

What caused these population divergences? Lowe (1950) hypothesized that a retraction of the humid coastal redwood forest from the higher elevation coastal plateau in northern Mendocino may have separated *A. flavipunctatus* populations into northern and southern isolates by creating a barrier of less favourable habitat between them, and a subsequent expansion of the forests forced them into secondary contact. Our genetic data corroborate this hypothesis. However, while secondary contact following allopatric divergence remains a viable hypothesis to explain the divergence of these salamanders, we note that the topology of our mtDNA tree is also consistent with a parapatric divergence scenario; that is, we see possible instances of primary contact as evident in the observed sister group relationships for the Northern vs. the two southern populations, as well as between the two southern populations. The boundary between the Northern vs. the two southern populations is defined by a sharp change in total rainfall (Oregon Climate Service) and talus rock availability, whereas the contact zone separating Coast and Inland populations is characterized by an abrupt transition in forest habitat, temperature, and moisture. Such contact zones, when present in the midst of steep environmental gradients, may cause parapatric divergence and speciation (Endler 1977). Other observations consistent with parapatric divergence include: (i) there is an apparent lack of physical barriers to dispersal, (ii) the color patterns of individuals in each population are cryptic against their background (i.e. possibly adaptive), (iii) the three contemporary populations are parapatrically distributed, and (iv) these sister-population groups meet where transitions in environmental conditions are most extreme. Landform changes could have also played a role in subdividing *A. flavipunctatus* into populations, because this region is a geological contact zone comprised of three tectonic plates (Norris & Webb 1990).

In addition to *A. flavipunctatus*, other terrestrial vertebrates also have a contact zone in Mendocino Co., California. This zone has been called the 'North Coast Divide' (Nussbaum 1976; Good 1989) because many amphibian and reptile species either have north vs. south phylogeographic breaks, or their northernmost or southernmost range limit in the same area (Zweifel 1952; Sessions & Kezer 1987; Good 1989; Smith 1995; Stebbins 2003; Shaffer *et al.* 2004; Pereira & Wake 2009). The presence of these multiple taxa sharing a similar zone of contact or range limit suggests that most if not all of them were subject to common environmental forces. Accordingly, this biogeographical pattern may indicate multiple speciation events in progress or having already been completed. Alhough our results suggest that the ancestral *A. flavipunctatus* population in this region became subdivided into at least three populations during the Pliocene-Pleistocene time interval, comparative phylogeographic studies are needed to determine whether or not these codistributed taxa underwent divergence simultaneously or at different times.

Conclusions

This study shows that mtDNA tree topology-based assignment can be profitably merged with clustering approaches using multiple nuclear loci to provide preliminary estimates for the number of evolutionary populations. These tentative population schemes can then be more rigorously analysed using historical demographic analyses, resulting in the corroboration or reformulation of these hypotheses. In this study, examination of our mtDNA and nDNA assignment results in light of our historical demographic analyses enabled us to conclude that Black Salamanders (*Aneides flavipunctatus*) in northwestern California are likely comprised of at least three populations or species. We also provide an example illustrating the ability of mtDNA to reveal apparently cryptic populations not detected by slower evolving nuclear loci, although we also recognize that very large genomic data sets may also perform well in this regard (e.g. Wang *et al.* 2003; Novembre *et al.* 2008). From a conservation perspective, all of our empirically supported populations and candidate populations qualified as ESUs. Although ESUs may be useful indicators of evolutionary populations for some study systems such as in Black Salamanders, we suggest that more important advances will be made when reproductive cohesion at the intrapopulational, interpopulational and interspecific levels can be more accurately and precisely estimated from genetic data.

Acknowledgements

We thank D. Wake for his help in the development of this project and his inspirational work on the evolutionary biology of

plethodontid salamanders. M. Mulks and A. Gottscho assisted with fieldwork and photography. A. Baker, M. Hart and L. Smith provided assistance in the laboratory. We also thank J. White, E. Metz, M. Weksler, and the McGuire laboratory group for helpful comments on this research and for support. E. Taylor and two anonymous reviewers provided many constructive criticisms that improved the quality of this paper. Collecting permits were granted by the CDFG (no. 9761) and handling of animals was governed by Humboldt State University IACUC Protocol no. 07/08.B.34.A. This research was supported by a grant from the Save-the-Redwoods League and from new faculty laboratory start-up funds provided to WBJ at Humboldt State University. This research was in partial fulfilment for an M.A. in Biology degree for SBR at Humboldt State University.

References

Anderson PK (1960) Ecology and evolution in island populations of salamanders in the San Francisco Bay region. *Ecological Monographs*, **30**, 359–386.

Avise JC (2000) Phylogeography: The History and Formation of Species. Harvard University Press, Cambridge, Massachusetts.

Beerli P (2004) Effects of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Molecular Ecology*, **13**, 827–836.

Blondel J, Thomas DW, Charmantier A, Perret P, Bourgault P, Lambrechts MM (2006) A thirty-year study of phenotypic and genetic variation of blue tits in Mediterranean habitat mosaics. *BioScience*, **56**, 661–673.

Brower AVZ (1999) Delimitation of phylogenetic species with DNA sequences: a critique of Davis and Nixon's population aggregation analysis. *Systematic Biology*, **48**, 199–213.

Carling MD, Lovette IJ, Brumfield RT (2010) Historical divergence and gene flow: coalescent analyses of mitochondrial autosomal and sex-linked loci in *Passerina* buntings. *Evolution*, **64**, 1762–1772.

Devitt TJ, Baird SJE, Moritz C (2011) Asymmetric reproductive isolation between terminal forms of the salamander ring species *Ensatina eschscholtzii* revealed by fine-scale genetic analysis of a hybrid zone. *BMC Evolutionary Biology*, **11**, 245.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Dudaniec RY, Schlotfeldt BE, Bertozzi T, Donnellan SC, Kleindorferb S (2011) Genetic and morphological divergence in island and mainland birds: informing conservation priorities. *Biological Conservation*, **144**, 2902–2912.

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

Edwards SV, Bensch S (2009) Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008. *Molecular Ecology*, **18**, 2930–2933.

Edwards SV, Kingan SB, Calkins JD *et al.* (2005) Speciation in birds: genes, geography, and sexual selection. *Proceedings of the National Academy of Sciences*, **102**, 6550–6557.

Endler JA (1977) Geographic Variation, Speciation, and Clines. Princeton University Press, Princeton, New Jersey.

Galbreath KE, Cook JA, Eddingsaas AA, DeChaine EG (2011) Diversity and demography in Beringia: multilocus tests of paleodistribution models reveal the complex history of arctic ground squirrels. *Evolution*, **65**, 1879–1896.

Gillespie JH (2004) Population Genetics, A Concise Guide, 2nd edn. The Johns Hopkins University Press, Baltimore, Maryland.

Good DA (1989) Hybridization and cryptic species in *Dicamptodon* (Caudata: Dicamptodontidae). *Evolution*, **43**, 728–744.

Hey J (2005) On the number of New World founders: a population genetic portrait of the peopling of the Americas. *PLoS Biology*, **3**, 965–975.

Hey J (2010) Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, **27**, 905–920.

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, **167**, 747–760.

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111**, 147–164.

Huelsenbeck JP, Andolfatto P (2007) Inference of population structure under a Dirichlet process model. *Genetics*, **175**, 1787–1802.

Jennings WB, Edwards SV (2005) Speciational history of Australian Grass Finches inferred from thirty gene trees. *Evolution*, **59**, 2033–2047.

Kumar S, Subramanian S (2002) Mutation rates in mammalian genomes. *Proceedings of the National Academy of Sciences*, **99**, 803–808.

Larkin MA, Blackshields G, Brown NP *et al.* (2007) ClustalX. *Bioinformatics*, **23**, 2947–2948.

Larson A (1980) Paedomorphosis in relation to rates of morphological and molecular evolution in the salamander *Aneides flavipunctatus*. *Evolution*, **34**, 1–17.

Larson A, Wake DB, Yanev KP (1984) Measuring gene flow among populations having high levels of genetic fragmentation. *Genetics*, **106**, 293–308.

Lee JY, Edwards SV (2008) Divergence across Australia's Carpentarian barrier: statistical phylogeography of the red-backed fairy wren (*Malurus melanocephalus*). *Evolution*, **62**, 3117–3134.

Lowe CH (1950) Speciation and ecology in salamanders of the genus *Aneides*. PhD Dissertation, University of California, Los Angeles, California.

Lynch JF (1981) Patterns of ontogenetic and geographic variation in the black salamander, *Aneides flavipunctatus*. *Smithsonian Contributions to Zoology*, **324**, 1–53.

Meader S, Ponting CP, Lunter G (2010) Massive turnover of functional sequence in human and other mammalian genomes. *Genome Research*, **20**, 1335–1343.

Moritz C (1994) Defining 'Evolutionary Significant Units' for conservation. *Trends in Ecology and Evolution*, **9**, 373–375.

Moritz C (2002) Strategies to protect biological diversity and the evolutionary processes that sustain it. *Systematic Biology*, **51**, 238–254.

Norris RM, Webb RW (1990) Geology of California. Wiley and Sons, New York, NY.

Novembre J, Johnson T, Bryc K *et al.* (2008) Genes mirror geography within Europe. *Nature*, **456**, 98–101.

Nussbaum RA (1976) Geographic variation and systematics of salamanders of the genus *Dicamptodon* Strauch (Ambystomatidae). *Miscellaneous Publications of the Museum of Zoology, University of Michigan*, **149**, 1–94.

Oregon Climate Service (OCS). Available from http://www.ocs.orst.edu. Accessed June 20, 2011.

Pereira RJ, Wake DB (2009) Genetic leakage after adaptive and nonadaptive divergence in the *Ensatina eschscholtzii* ring species. *Evolution*, **63**, 2288–2301.

Ponting CP, Hardison RC (2011) What fraction of the human genome is functional? *Genome Research*, **21**, 1769–1776.

Rissler LJ, Apodaca JJ (2007) Adding more ecology into species delimitation: ecological niche models and phylogeography help define cryptic species in the Black Salamander (*Aneides flavipunctatus*). *Systematic Biology*, **56**, 924–942.

Rissler LJ, Taylor DR (2003) The phylogenetics of Desmognathine salamander populations across the southern Appalachians. *Molecular Phylogenetics and Evolution*, **27**, 197–211.

Rojas-Soto OR, Westberg M, Navarro-Siguenza AG, Zink RM (2010) Genetic and ecological differentiation in the endemic avifauna of Tiburon Island. *Journal of Avian Biology*, **41**, 398–406.

Ronquist F, Huelsenbeck JP (2003) MrBayes3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Rovito SM (2010) Lineage divergence and speciation in the Web-toed Salamanders (Plethodontidae: Hydromantes) of the Sierra Nevada, California. *Molecular Ecology*, **19**, 4554–4571.

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends in Ecology and Evolution*, **1**, 9–10.

Sessions SK, Kezer J (1987) Cytogenetic evolution in the plethodontid salamander genus *Aneides*. *Chromosoma*, **95**, 17–30.

Shaffer HB, Fellers GM, Voss SR, Oliver JC, Pauly GB (2004) Species boundaries, phylogeography and conservation genetics of the red-legged frog (*Rana aurora/draytonii*) complex. *Molecular Ecology*, **13**, 2667–2677.

Slatkin M (2005) Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. *Molecular Ecology*, **14**, 67–73.

Smith HM (1995) Handbook of Lizards: Lizards of the United States and Canada. Cornell University Press, Ithaca, New York.

Stamatakis A, Ludwig T, Meier H (2005) RAxML-III: a fast program for maximum likelihood–based inference of large phylogenetic trees. *Bioinformatics*, **21**, 456–463.

Staub NL, Brown CW, Wake DB (1995) Patterns of growth and movements in a population of *Ensatina eschscholtzii platensis* (Caudata: Plethodontidae) in the Sierra Nevada, California. *Journal of Herpetology*, **29**, 593–599.

Stebbins RC (2003) Western Reptiles and Amphibians, 3rd edn. Peterson Field Guide Series, New York, NY.

Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *American Journal of Human Genetics*, **76**, 449–462.

Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics*, **68**, 978–989.

Strasburg JL, Rieseberg LH (2010) How robust are "Isolation with Migration" analyses to violations of the IM model? A simulation study. *Molecular Biology and Evolution*, **27**, 297–310.

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

Tan AM, Wake DB (1995) MtDNA phylogeography of the California newt, *Taricha torosa* (Caudata, Salamandridae). *Molecular Phylogenetics and Evolution*, **4**, 383–394.

Thomson RC, Wang IJ, Johnson JR (2010) Genome-enabled development of DNA markers for ecology, evolution, and conservation. *Molecular Ecology*, **19**, 2184–2195.

Toews DPL, Brelsford A (2012) The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, **21**, 3907–3930.

Wang Z, Hill GE, Baker AJ, Edwards SV (2003) Reconciling actual and inferred population histories in the House Finch (*Carpodacus mexicanus*) by AFLP analysis. *Evolution*, **57**, 2852–2864.

Waples RS, Gaggiotti O (2006) What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, **15**, 1419–1439.

Woerner AE, Cox MP, Hammer MF (2007) Recombination-filtered genomic datasets by information maximization. *Bioinformatics*, **23**, 1851–1853.

Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 98–160.

Yang Z (2006) Computational Molecular Evolution. Oxford University Press, Oxford, UK.

Zink RM (1994) The geography of mitochondrial DNA variation, population structure, hybridization, and species limits in the Fox Sparrow (*Passerella iliaca*). *Evolution*, **48**, 96–111.

Zink RM, Barrowclough GF (2008) Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology*, **17**, 2107–2121.

Zweifel RG (1952) Pattern variation and evolution of the mountain kingsnake, *Lampropeltis zonata*. *Copeia*, **1952**, 152–168.

Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. PhD Dissertation, The University of Texas, Austin, Texas.

This study is based on S.B.R.'s Master's Thesis research project supervised by S.B.M. and W.B.J. All authors designed the study; S.B.R. collected samples from the field, performed laboratory work, analysed data, and wrote the manuscript; S.B.M. and W.B.J. participated in the manuscript writing. All authors share similar research interests about the ecology and evolution of amphibians and reptiles..

Data accessability

DNA sequences: GenBank accessions JX544070–JX544733.

Sampling locations, primer sequences, data coverage, STRUCTURAMA assignments, MFER/SFER converted IM parameters, and locus-specific Tajima's D and nucleo-

tide diversity values uploaded as online supplemental material tables.

Photos of ecomorphs, anonymous loci gene trees, and IM/IMa2 posterior distribution plots uploaded as supplemental material figures.

IM and IMa2 input files, STRUCTURAMA input file, tree files, and final DNA sequence assemblies: DRYAD entry doi:10.5061/dryad.kh095.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Sample localities and associated museum voucher numbers.

Table S2 Primer sequences and approximate product size.

Table S3 Data coverage.

Table S4 STRUCTURAMA population assignments.

Table S5 Converted IM parameters for the SFER and MFER mtDNA population comparison.

Table S6 Tajima's D and nucleotide diversity values for each locus by population.

Fig. S1 Color patterns of *Aneides flavipunctatus* observed near the Mendocino Co. contact zone.

Fig. S2 Midpoint rooted maximum likelihood gene trees for loci *SR1*, *SR2*, *SR4*, and *SR7*.

Fig. S3 Midpoint rooted maximum likelihood gene trees for loci *SR8*, *SR9*, *SR12*, and *SR15*.

Fig. S4 Midpoint rooted maximum likelihood gene trees for loci *SR16*, *SR17*, *SR19*, and *SR20*.

Fig. S5 Estimated posterior probability distributions for contemporary effective population size (θ) from IMa2 and IM.

Fig. S6 Estimated posterior probability distributions for ancestral effective population size (θ) from IMa2 and IM.

Fig. S7 Estimated posterior probability distributions for population divergence time (*t*) from IMa2 and IM.

Fig. S8 Estimated posterior probability distributions for population migration rates (*m*) from IMa2.

Fig. S9 Estimated posterior probability distributions for population migration rates (*m*) from IM.

Appendix S1 Development of anonymous nuclear loci.

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