Classification: Biological Sciences; Evolution

Geography is more important than life history in the recent diversification of the tiger salamander complex

Kathryn M. Everson[a,1], Levi N. Gray[a,1], Angela G. Jones[a], Nicolette M. Lawrence[a], Mary E. Foley[a], Kelly L. Sovacool[a,b], Justin D. Kratovil[a,c], Scott Hotaling[a,d], Paul M. Hime[a,e], Andrew Storfer[d], Gabriela Parra-Olea[f], X. Aguilar-Miguel[g], Eric M. O'Neill[a], Luis Zambrano[h], H. Bradley Shaffer[i], David W. Weisrock[a,2]

[a] Department of Biology, University of Kentucky, Lexington, KY 40506

[b] Department of Computational Medicine & Bioinformatics, University of Michigan, Ann Arbor, MI 48109

[c] Ecology and Evolutionary Biology Department, University of Connecticut, Storrs, CT 06269

[d] School of Biological Sciences, Washington State University, Pullman, WA 99164

[e] Biodiversity Institute, University of Kansas, Lawrence, KS 66045

[f] Departamento de Zoología, Instituto de Biología, Universidad Nacional Autónoma de México, Ciudad de México, México

[g] CIRB, Facultad de Ciencias, Universidad Autónoma del Estado de México, Toluca, Edo. de México, México

[h] Facultad de Medicina Veterinaria y Zootecnia, Universidad Autónoma de Sinaloa, Culiacán Rosales 82260, Sinaloa, Mexico

[i]Department of Ecology and Evolutionary Biology, and La Kretz Center for California Conservation Science, University of California, Los Angeles, CA 90095

[1] K.M.E. and L.N.G. contributed equally to this work

[2] To whom correspondence may be addressed. Email: david.weisrock@uky.edu

ABSTRACT

The North American tiger salamander species complex, including its flagship species the axolotl, has long been a source of biological fascination. The complex exhibits a wide range of variation in developmental life history strategies, including populations and individuals that undergo metamorphosis and those able to forego metamorphosis and retain a larval, aquatic lifestyle (i.e., paedomorphosis). Such disparate life history strategies are assumed to cause populations to become reproductively isolated, but the degree to which they have actually shaped population- and species-level boundaries is poorly understood. Using a large multi-locus dataset from hundreds of samples across North America, we identified genetic clusters with clear signs of admixture across the geographic range of the tiger salamander complex. Population clusters often contain a mixture of paedomorphic and metamorphic taxa, and we conclude that geography has played a large role in driving lineage divergence relative to obligate paedomorphosis in this system. This conclusion is bolstered by model-based analyses demonstrating gene flow between metamorphic and paedomorphic populations. Even the axolotl, a paedomorphic species with an isolated native range, apparently has a history of gene flow with its neighboring populations. This fine-scale genetic perspective on life-history variation establishes a framework for understanding how plasticity, local adaptation, and gene flow contribute to lineage divergence. The axolotl is currently used as the vertebrate model system in regenerative biology, and our findings chart a course for more informed use of these and other tiger salamander species in experimental and field research, including conservation priorities.

Keywords: *Ambystoma*, life history, phylogenetics, population genomics, salamanders

SIGNIFICANCE STATEMENT

Population structure and speciation are shaped by a variety of biotic and abiotic factors. In the tiger salamander complex, one factor that may influence diversification is life history: some taxa are obligately paedomorphic–a condition where adults maintain an aquatic, larval phenotype–while others are facultatively paedomorphic or entirely metamorphic. Using a large multi-locus dataset, we found evidence of gene flow and/or panmixia between obligately and facultatively paedomorphic taxa, suggesting that an obligately paedomorphic life history is not a strong driver of speciation in the tiger salamander complex. We also recovered a history of gene flow between the critically endangered axolotl and its neighboring populations, providing important information for its conservation and captive management.

61 **INTRODUCTION**

62 Life history—the complement of traits affecting survival and reproduction over an
63 organism's lifetime—affects ecology and dispersal, and therefore plays a major role in shaping
64 population structure and speciation (1, 2). The axolotl (*Ambystoma mexicanum*) and related
65 salamander species of the North American *Ambystoma tigrinum* complex (Table 1) display
66 enormous variation in life history (3). Some members of this group are considered obligate
67 paedomorphs, where sexually mature adults retain a larval, aquatic body plan that includes
68 external gills and an enlarged tail fin (4). Others are considered obligate metamorphs,
69 transforming from aquatic larvae to terrestrial adults. Most populations are facultative in their
70 developmental life history strategy, transforming under certain genetic and/or environmental
71 conditions (5–8). This lability in life history is believed to have played an important role in the
72 diversification of the *A. tigrinum* complex, particularly in the Trans-Mexican Volcanic Belt
73 (TMVB) region, where several essentially paedomorphic species (including the axolotl) have
74 been described (9–11). Obligate paedomorphosis is estimated to have evolved in this region
75 multiple times (11–13) in association with relatively large, permanent bodies of water.
76 Presumably, restricted gene flow between these isolated, paedomorphic populations led to
77 speciation as well as morphological adaptations to the aquatic lifestyle (14).

78 The term "species complex," while not a formal taxonomic category, is often used to
79 describe groups of closely related lineages early in the divergence process. The tiger salamander
80 species complex has been highlighted as a potentially valuable example of a recent radiation (15)
81 that could provide insight into the early mechanisms initiating and/or maintaining diversity (16–
82 21). However, such species complexes poses a number of challenges: phenotypic differences
83 among lineages may be subtle or absent (i.e., cryptic species), ancestral polymorphisms may be

Table 1. Currently recognized species in the *Ambystoma tigrinum* complex. Life histories, habitat preferences, and conservation status were compiled from AmphibiaWeb (2019) and the IUCN (2019).

Species	Description Reference	Life History	Habitat Preference	Conservation Status
Ambystoma altamirani	Dugès, 1895	Facultative paedomorph	Streams, occasional ponds	Endangered
Ambystoma amblycephalum	Taylor, 1940	Facultative paedomorph	Ponds	Critically endangered
Ambystoma andersoni	Krebs & Brandon, 1984	Paedomorph	Lake Zacapu	Critically endangered
Ambystoma bombypellum	Taylor, 1940	Facultative paedomorph	Lakes / Ponds, occasional streams	Data deficient
Ambystoma californiense	Gray 1853	Metamorph	Vernal pools	Vulnerable
Ambystoma dumerilii	(Dugès, 1870)	Paedomorph	Lake Patzcuaro	Critically endangered
Ambystoma flavipiperatum	Dixon, 1963	Facultative paedomorph	Lakes / Ponds, occasional streams	Endangered
Ambystoma granulosum	Taylor, 1944	Facultative paedomorph	Lakes / Ponds, occasional streams	Critically endangered
Ambystoma leorae	(Taylor, 1943)	Facultative paedomorph	Streams	Critically endangered
Ambystoma lermaense	(Taylor, 1940)	Facultative; often paedomorphic	Lakes / Ponds	Endangered
*Ambystoma mavortium**	Baird, 1850	Facultative paedomorph	Lakes / Ponds	Least concern
Ambystoma mexicanum	(Shaw and Nodder, 1798)	Paedomorph	Lakes Xochimilco & Chalco	Critically endangered
Ambystoma ordinarium	Taylor, 1940	Facultative paedomorph	Streams, ponds	Endangered
Ambystoma rivulare	(Taylor, 1940)	Facultative paedomorph	Streams	Endangered
Ambystoma rosaceum	Taylor, 1941	Facultative paedomorph	Streams, occasional ponds	Least concern
Ambystoma silvense	Webb, 2004	Facultative paedomorph	Lakes / Ponds	Data deficient
Ambystoma taylori	Brandon et al., 1982	Paedomorph	Lake Alchichica (saline)	Critically endangered
Ambystoma tigrinum	Green, 1825	Rare paedomorph	Lakes / Ponds	Least concern
Ambystoma velasci	(Dugès, 1888)	Facultative paedomorph	Lakes / Ponds, Streams	Least concern

**includes 5 recognized subspecies: A. m. mavortium Baird 1850, A. m. melanostictum (Cooper 1859), A. m. diaboli Dunn 1940, A. m. nebulosum Hallowell 1853, and A. m. stebbinsi Lowe 1954*

84 stochastically inherited (i.e., incomplete lineage sorting), and reproductive barriers may be
85 porous (i.e., speciation with gene flow). Furthermore, phenotypic variation may not reflect
86 lineage divergence or speciation processes; for example, trait variation may be the result of
87 plasticity or balancing selection (22). The reliance on phenotype for species diagnosis in the *A.*
88 *tigrinum* complex may require particular scrutiny, as several phenotypic traits, including
89 developmental state and adult color pattern, can be plastic and highly variable (23–25). Given
90 these challenges, fundamental questions remain concerning the evolutionary distinctiveness of
91 component lineages within the tiger salamander complex, calling into question the current
92 taxonomy as an accurate reflection of the underlying population biology.

93 Previous research has produced mixed results regarding levels of genetic differentiation
94 among lineages in the *A. tigrinum* complex, as well as the relative importance of paedomorphosis
95 as a driver of diversification. While population- and species-level structure is evident, an
96 important theme emerging across multiple studies has been that reproductive barriers are porous
97 or incomplete (11, 26, 27). Indeed, Shaffer and McKnight (11) noted "the striking lack of
98 differentiation among the 14 species of the tiger salamander complex" (p. 425). Some studies
99 have found an elevated degree of population structure among paedomorphic populations relative
100 to metamorphic populations (12, 28), but others have demonstrated that paedomorphic taxa and
101 neighboring metamorphic populations are not fully reproductively isolated (29) and can produce
102 viable hybrid offspring under laboratory conditions (30, 31). Furthermore, there is evidence that
103 some populations considered to be obligately paedomorphic are capable of transforming at low
104 rates and may be more appropriately considered facultatively paedomorphic (10, 32, 33),
105 although the absence of transformed individuals in nature is difficult to verify without targeted
106 ecological studies (34).

107 To illuminate the processes underlying diversification in the tiger salamander complex,
108 an important first step must be a range-wide assessment of population structure and the
109 clarification of population- and species-level boundaries. With this groundwork, insights into the
110 role of paedomorphosis and diversification can be addressed. For a study system of this
111 geographic and taxonomic scale, robust inference of population structure requires information
112 derived from throughout the genome combined with thorough range-wide sampling. To meet
113 these criteria, we expand on a large multi-locus dataset containing 95 nuclear loci for 93
114 individuals (35), to produce a data matrix for 347 individuals across the full geographic range of
115 tiger salamanders (Fig. S1). Given the complex and somewhat checkered history of the group's
116 taxonomy, we take a naïve approach, performing population structure analyses without *a priori*
117 identification of taxa to resolve geographic genetic clusters and characterize patterns of
118 admixture. We couple these results with phylogenetic network analyses and model-based tests of
119 migration between metamorphic and paedomorphic populations to provide an additional
120 perspective on the degree to which populations are genetically connected, now or in the recent
121 past. We then infer a species tree for identified genetic clusters in an attempt to provide an
122 updated working hypothesis for the evolution of the group. Finally, we test the overarching
123 hypothesis that life history evolution has driven speciation in the tiger salamander complex,
124 particularly through the assumed isolation and independence of small-range endemics classified
125 as paedomorphic species.
126

127 **RESULTS AND DISCUSSION**

128 *Identification of genetic lineages and gene flow*

129 A principal components analysis (PCA) and a discriminant analysis of principal
130 components (DAPC; 36) both recovered *Ambystoma californiense*, the only confirmed obligate
131 metamorph in the complex, as genetically distinct from all other tiger salamanders (Fig. 1A, S2,
132 S3), which is consistent with previous work (e.g., 11, 37). After removing *A. californiense* from
133 the dataset, DAPC analysis supported the recognition of three primary genetic clusters: U.S.
134 (including samples from southern Canada), northwestern Mexico, and central Mexico (Fig. 1B,
135 S4). All populations in the U.S. and northern Mexico clusters are considered facultatively
136 paedomorphic or primarily metamorphic (38). The central Mexico cluster, however, contains a
137 wealth of life history extremes: four currently recognized species are considered to be obligate
138 paedomorphs and ten others are facultatively paedomorphic to varying degrees (13, 39). A PCA
139 of these data produced similar clustering results, with ordination patterns largely mirroring
140 geographic sampling (Fig. S5). For discussion of the taxonomic implications of our results, see
141 the Supplementary Material.

142 Within central Mexico, both DAPC and STRUCTURE analyses identified four
143 geographic clusters (hereafter referred to as CM1-CM4) with signatures of admixture within and
144 among them (Figs. 2A, S6-S8). Network analysis (40) also recovered groups that generally
145 correspond to CM1-CM4, but with a large number of reticulate branches (Fig. 3). This starburst-
146 like pattern indicates the presence of conflicting topologies within the dataset (40), which can be
147 caused by biological processes including incomplete lineage sorting (a historical legacy of recent
148 diversification) and gene flow, either current or in the evolutionarily recent past. Phylogenetic

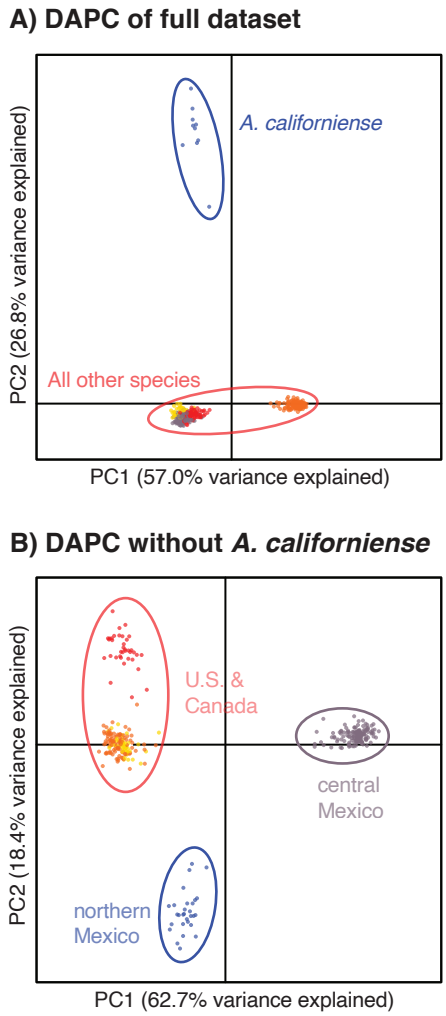

Figure 1. Results of discriminant analysis of principal components (DAPC) on (A) the entire genomic dataset (without *A. opacum* and *A. texanum* outgroups, and (B) the genomic dataset without the highly divergent *A. californiense*. In both plots, points represent individuals and ellipses show the groups identified by DAPC. The first and second principal components from the DAPC are on the x and y axes, respectively. Results from additional values of *K* are provided in Supplementary Figures 2 and 3.

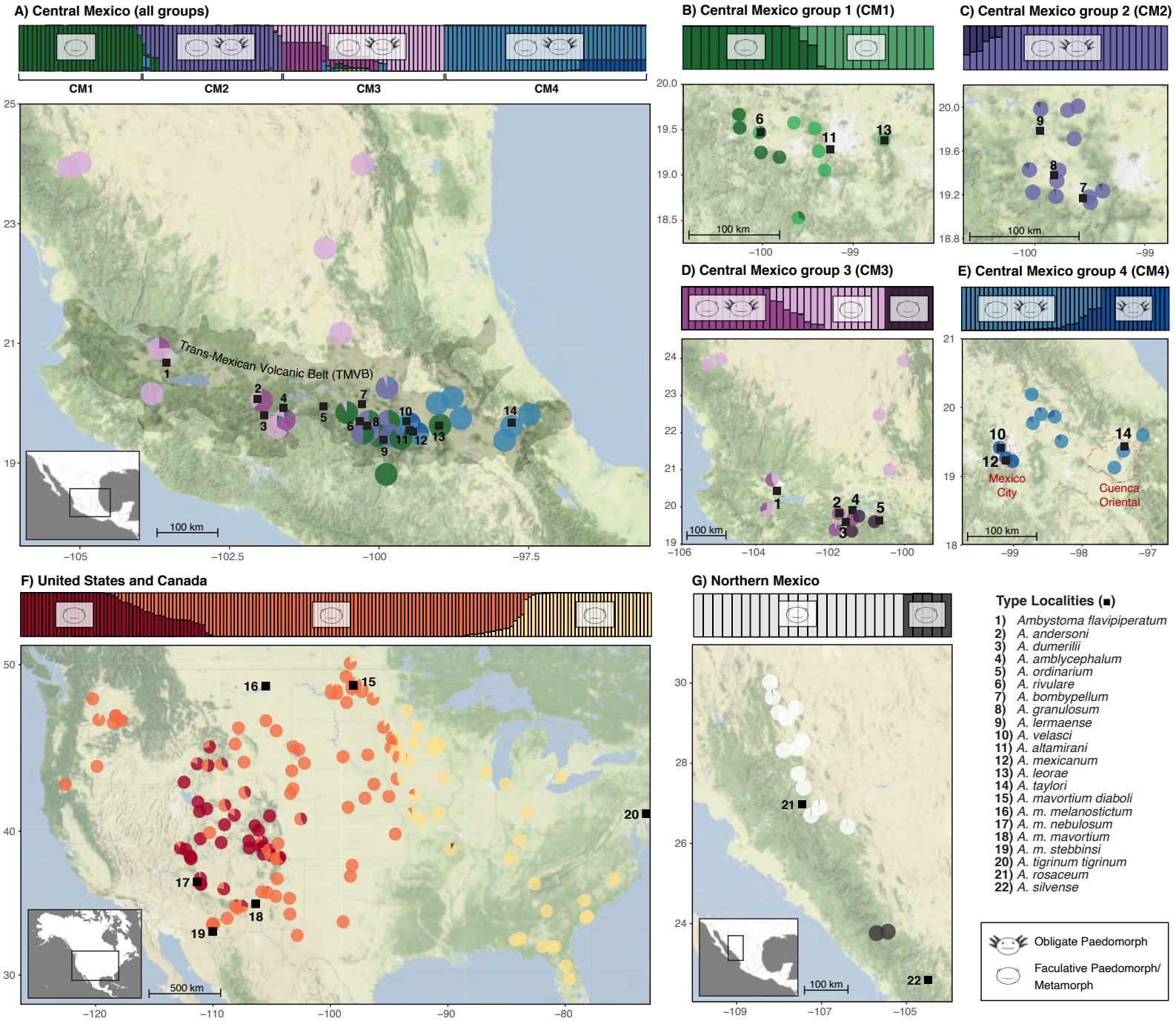

Figure 2. Results of population genomic analyses on the *Ambystoma tigrinum* species complex. (A) STRUCTURE analysis of the central Mexico cluster, which identified four major groups. Each vertical bar represents an individual, while the y-axis shows the membership probability for each group. Pie charts mapped below the assignment plot show the average group membership probability for that locality. Some localities were combined to a single pie chart to reduce visual clutter; see Table S1 for specific coordinates of each sample. The range of the Trans-Mexican Volcanic Belt (TMVB) is identified by shading and is modified from Ferrari (2004). (B-E) Detailed STRUCTURE analyses of each central Mexico subgroup CM1-CM4. A geographic outlier placed in CM2 is thought to represent a geographic range introduction and is not shown here, but is discussed in supplementary text and shown in Fig. S6. The range of the Cuenca Oriental is outlined in red and is modified from Percino-Daniel et al. (2016). (F,G) STRUCTURE results from analyses of the U.S. and northern Mexico groups, respectively. Detailed results are shown in Figs. S8-S9. Black squares on the maps denote the type localities of all species included in this study. On each STRUCTURE membership plot, salamander face symbols denote the presence of obligately or facultatively paedomorphic populations (with or without gills, respectively); groups labeled with both symbols contain a mixture of life history strategies.

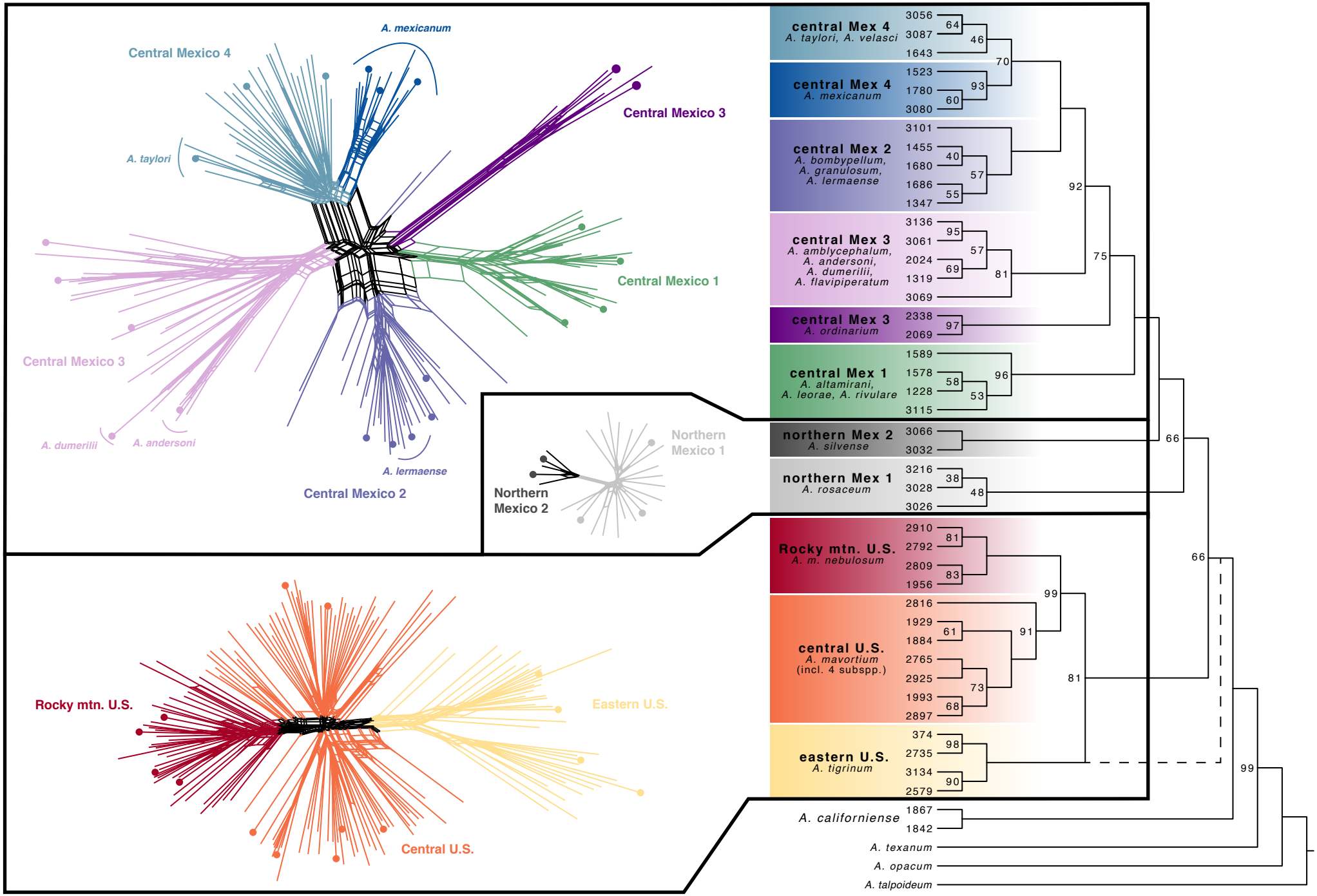

Figure 3. Results of phylogenetic network analyses (left) and the quartets-based phylogenetic analysis (right) of the *Ambystoma tigrinum* species complex. Clades are colored according to their cluster assignments in population genetic analyses, as shown in detail in Fig. 2. In phylogenetic networks, species names are only shown for taxa that are obligately or strongly paedomorphic; all other taxa (not shown) are considered to be facultatively paedomorphic or metamorphic. Note that the quartet tree analysis used a subset of individuals, indicated by solid circles on the phylogenetic networks. Phylogenetic results from Bayesian and maximum-likelihood concatenated analyses were largely concordant with the quartets tree (Fig. S10), although one important difference (the relationship of the eastern U.S. clade) is indicated by a dashed line. Bootstrap support (BS) values are indicated on each node of the quartets tree (BS values > 95 are not shown), while tips are labeled with individual IDs (Table S1). Branch lengths are not scaled to time or substitution rate.

149 analyses (41–43) recovered monophyletic CM1, CM2, and CM4 clusters (see below for
150 discussion of CM3), but support values were often low within the central Mexico clade (Fig. 3).
151 Based on type locality and range information, the CM1 cluster includes three stream-
152 breeding and facultatively paedomorphic taxa: *Ambystoma altamirani* Duges 1895 (44), *A.*
153 *leorae* Taylor 1943 (45), and *A. rivulare* Taylor 1940 (9). We did not find evidence that these
154 three taxa are reproductively isolated. Subsequent analyses of this cluster identified two admixed
155 populations associated with eastern and western portions of its geographic distribution (Fig. 2B).
156 Phylogenetically, CM1 was monophyletic and sister to all other central Mexican groups (Fig. 3),
157 consistent with their original placement in a distinct genus, *Rhyacosiredon* Dunn 1928 (46).
158 Analyses of CM2 grouped *A. lermaense* Taylor 1940 (9) with samples corresponding to
159 the type localities and range of *A. bombypellum* Taylor 1940 (9) and *A. granulosum* Taylor 1944
160 (47). There was no evidence for genetic isolation among any of these three taxa (Fig. 2C).
161 Within CM3, population-genetic analyses first separated samples of *A. ordinarium* Taylor
162 1940 (5) from the rest (with no hybridization detected; Fig 2D). Other samples were
163 geographically structured according to the northern and southern portions of the distribution (Fig
164 1E). The more southern cluster included two obligately paedomorphic taxa, *A. andersoni* Krebs
165 & Brandon 1984 (32) and *A. dumerilii* Duges 1870 (48), each endemic to single lake systems
166 (Lakes Zacapú and Pátzcuaro, respectively). However, our results identified admixture between
167 these lakes and other nearby localities. The samples from other CM3 localities likely represent *A.*
168 *amblycephalum* Taylor 1940 (9) and *A. flavipiperatum* Dixon 1963 (49).
169 CM4 included the axolotl, *A. mexicanum* Shaw & Nodder 1798 (50), and *A. taylori*
170 Brandon et al. 1981 (10)—both primarily considered to be near-obligate paedomorphs—as well
171 as samples assigned to the polymorphic *A. velasci* Duges 1888 (51). Subsequent DAPC and

172 STRUCTURE analyses of CM4 placed all but one *A. mexicanum* individual sampled from Lakes
173 Xochimilco and Chapultepec (i.e., the remaining habitats of the axolotl) in a distinct cluster (Fig.
174 2E, S7-S8), with evidence of some admixture in samples to the northwest of Mexico City.
175 We used a Bayesian model-testing framework (Fig. 4A) to test for genetic isolation of
176 paedomorphic populations in three central Mexico groups. Specifically, we tested for the
177 presence or absence of gene flow between (near) obligate paedomorphs (*A. andersoni* and *A.*
178 *dumerilii* in CM3, and *A. mexicanum* and *A. taylori* in CM4) and their surrounding facultatively
179 paedomorphic populations. We also tested for gene flow between *A. lermaense* [which has been
180 described as "often neotenic [= paedomorphic]" (9, pg. 427)] and remaining populations in CM2.
181 In four out of the five groups, we recovered models of bidirectional gene flow or complete
182 panmixia between putatively obligate paedomorphs and facultative paedomorphs (Fig. 4B).
183 Alternative models of no migration or unidirectional migration were rejected with decisive Bayes
184 factors (Table S3). Migration rate estimates for full-migration scenarios consistently identified
185 higher levels of gene flow from facultative to obligate paedomorphic populations. We found
186 support for genetic isolation of one paedomorphic taxon, *A. andersoni* (Fig. 4B), a species
187 endemic to Lake Zacapú, Michoacán. As we are unable to account for landscape barriers or
188 sampling gaps in these analyses, we cannot be certain whether the lack of gene flow we
189 recovered for *A. andersoni* is due to its paedomorphic life history or to its geography. Another
190 important caveat to our migration analyses is that divergence time is not included as a model
191 parameter, which prevents distinguishing between historical and contemporary gene flow. Thus,
192 our ability to detect ongoing gene flow in these tests is limited.
193 Within the U.S. and Canada, DAPC and STRUCTURE identified three genetic clusters
194 (K = 3) excluding the allopatric, more differentiated *A. californiense*, with clear signs of

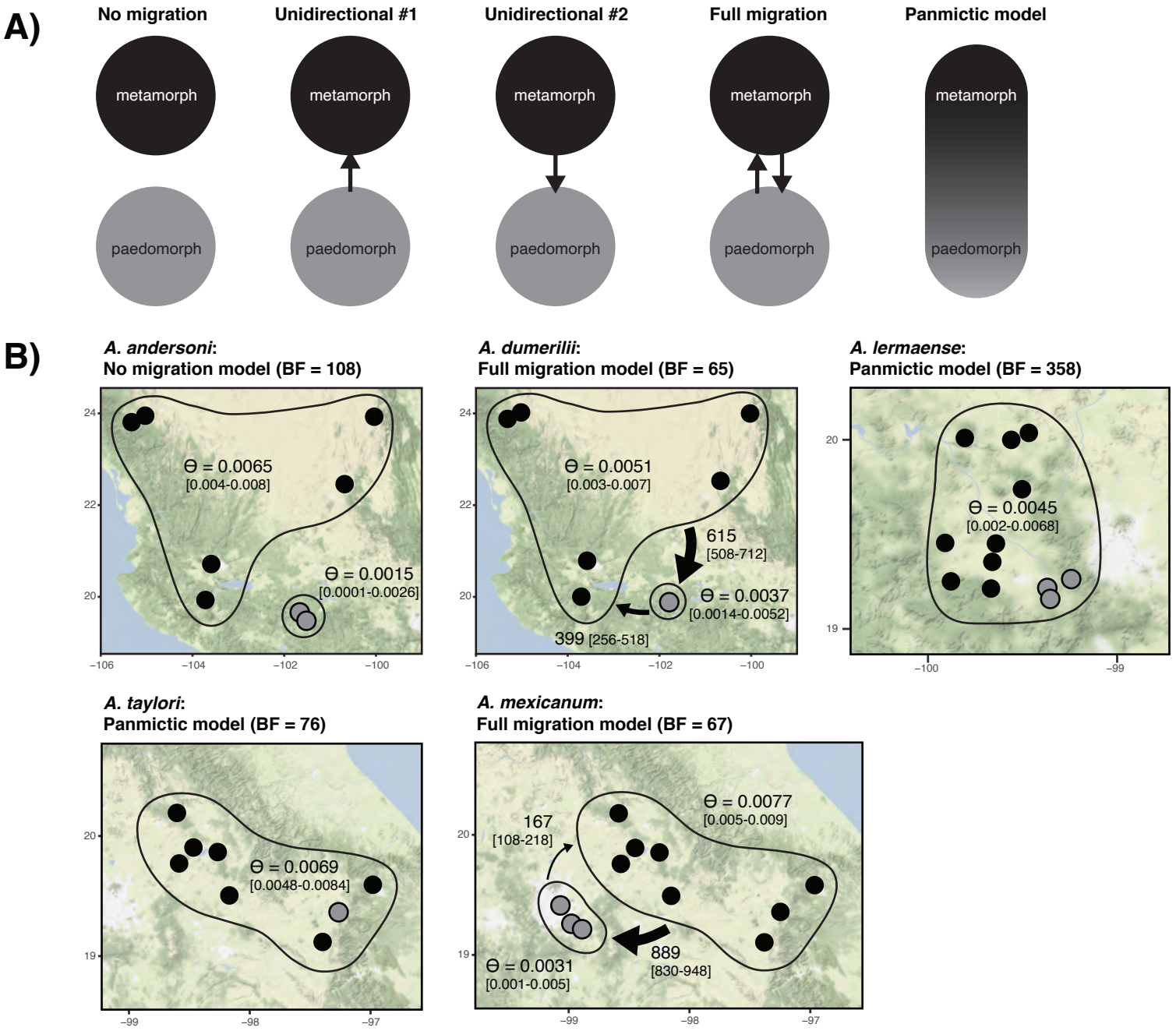

Figure 4. (A) Demographic models used to test hypotheses of gene flow. For each species considered to be obligately (*A. andersoni*, *A. dumerilii*, *A. taylori*, and *A. mexicanum*) or often (*A. lermaense*) paedomorphic, we evaluated five models of population structure and migration between that taxon and the facultatively paedomorphic taxa in the same genetic cluster (CM1-CM4). (B) Results of demographic model tests. Θ denotes the median estimated mutation-scaled effective size of each population. Values next to each arrow are median mutation-scaled migration rates ($M = m/\mu$, where m is the fraction of immigrants in each generation and μ is the mutation rate per generation per site). For both Θ and M, 95% confidence intervals are shown in brackets below the median value. Gray circles denote obligately paedomorphic lineages, while black circles denote facultatively paedomorphic lineages.

195 admixture along zones of contact (Figs. 2F, S9, S10). Geographically, these clusters are
196 associated with the eastern U.S., the central + western U.S., and the southwestern U.S. + Rocky
197 Mountains (hereafter, we refer to the latter two groups as U.S. central and U.S. Rocky
198 Mountains, respectively). Phylogenetic analyses corroborated these results, albeit with extensive
199 reticulations among clusters in the phylogenetic network (Fig. 3) and mixed support for species-
200 tree relationships among the three major groups (Fig. 3 vs. Fig. S12). Further exploration of
201 population structure within each of the three U.S. clusters recovered additional patterns of
202 differentiation (Figs. S9, S10), including northern and southern clusters in the eastern U.S. and
203 central U.S., and an allopatric cluster restricted to the Pacific Northwest. Overall, our U.S. results
204 are similar to mtDNA-based results from previous studies (11, 52), which identified haplotype
205 clades associated with the eastern U.S., the northern Great Plains, the southern Great Plains, and
206 the Pacific Northwest.

207 All analyses of the northern Mexico group identified two genetic clusters ($K = 2$; Figs.
208 2G, S11). The northern cluster corresponds to *A. rosaceum* Taylor 1941 (53), while the southern
209 cluster could represent either a second population of *A. rosaceum* or *A. silvense* Webb 2004 (54);
210 we did not include any samples from the type locality of *A. silvense* in this study. Phylogenetic
211 analyses recovered each of these two northern Mexico clusters as monophyletic, but not sister to
212 one another, and with few reticulate nodes in the phylogenetic network (Fig. 3, S12).
213

214 *The influence of life history on diversification*

215 A popular perspective on diversification in this species complex has been that the
216 evolution of obligate paedomorphosis promotes reproductive isolation and, ultimately, speciation
217 (10, 11). However, an important theme of our results is that taxa with different life history

218 strategies commonly cluster together by geography in population genetic and phylogeographic
219 analyses (Figs. 2, 3). Moreover, our migration analyses support relatively recent gene flow
220 and/or panmixia between facultative and paedomorphic populations (Fig. 4). Thus, the evolution
221 of life history extremes does not appear to be the main driver of speciation in the tiger
222 salamander complex. It is also worth noting that facultative paedomorphosis is the inferred
223 ancestral condition (8) and nearly all extant lineages in the *Ambystoma tigrinum* complex are
224 capable of metamorphosis, either in the wild or in the lab. Only *A. dumerilii* is not known to
225 survive for long after metamorphosis (55). That only one described taxon in the complex is
226 putatively fixed for paedomorphosis raises several questions: Are the rest of the primarily
227 paedomorphic populations merely *en route* to fixation, rather than actually fixed? Do facultative
228 populations fluctuate in paedomorphic frequency over time, dependent on prevailing
229 environmental and/or demographic conditions (5)? Do transforming and non-transforming
230 individuals readily interbreed outside of the laboratory setting? While these remain open
231 questions, below we provide some relevant context for exploration and future research.
232 Theory predicts that paedomorphosis in salamanders is most likely to become fixed in a
233 population when ecological conditions are favorable in the larval, aquatic environment but
234 unfavorable on land (e.g., when terrestrial habitats are arid or resource-limited) (12, 56, 57).
235 However, it is not known how long these fixations persist over evolutionary time; obligate
236 paedomorphic lineages may have arisen and gone extinct multiple times over tiger salamander
237 evolutionary history (58). One situation that would result in complete speciation is if
238 paedomorphosis became fixed in the population and migration with surrounding facultatively
239 paedomorphic populations ceased due to strong assortative mating and/or selection against
240 hybrids (59–61). However, assortative mating within life history morphs has not been

241 demonstrated in other salamander groups (62–64) and there is evidence that different tiger
242 salamander morphs do interbreed (65). In fact, the only known instance of positive assortative
243 mating in *A. tigrinum* is associated with tail length (66, 67), which might favor metamorphosed
244 males as paedomorphic individuals can have relatively short tails (32). The evidence for
245 selection against intermorph offspring is also scarce. Captive breeding using artifical fertilization
246 has produced *A. mexicanum* x *A. tigrinum* hybrids (7, 31, 68) as well as *A. dumerilii* x *A.*
247 *tigrinum* and *A. dumerilii* x *A. rivularis* hybrids, although the fertility of offspring was not
248 investigated in *A. dumerilii* crosses (31). This suggests that postzygotic barriers are weak,
249 although the extent to which divergent lineages would interbreed in the wild remains a key
250 question in need of investigation.

251 If biological barriers to reproduction are limited or absent, there are several potential
252 mechanisms by which dispersal could mediate gene flow between paedomorphic and
253 metamorphic populations. First, metamorphosed individuals from facultatively paedomorphic
254 populations could move into bodies of water containing paedomorphic populations. Our
255 demographic analyses support this scenario, as they recovered higher rates of migration entering
256 rather than exiting obligately paedomorphic groups (Fig. 4). However, demographic analyses
257 also estimated non-zero emigration out of paedomorphic populations; thus, "obligate"
258 paedomorphs might also occasionally metamorphose and disperse. Occasional metamorphosis
259 has been documented in several putatively paedomorphic populations (9, 10, 33, 69, 70) and
260 metamorphosis can occur at higher frequencies in dispersing males in facultative populations
261 (71). Even a low frequency of transformed adults is likely sufficient to maintain evolutionary
262 cohesion with surrounding populations (10, 29, 72). Finally, gene flow may not necessarily be
263 mediated across a terrestrial environment; genetic connectivity could also be maintained if

264 populations have been forced into contact by water level fluctuations in the groundwater system
265 of a region. The dynamic lacustrine history of the Cuenca Oriental (73), for instance, suggests
266 that isolation of some paedomorphic populations has been punctuated by broader aquatic
267 connections over short geological time scales, potentially facilitating gene flow. Collectively,
268 these mechanisms are likely to explain the maintenance of geographic genetic groups, now or in
269 the evolutionarily recent past, containing a range of life histories.

270 The genetic cohesion we detected across wide swaths of the TMVB is particularly
271 striking given other evidence for local adaptation in this region. This scenario is perhaps best
272 exemplified in the Cuenca Oriental (Fig. 2E), where some lakes contain paedomorphic
273 salamanders adapted to saline conditions, a very rare trait among amphibians (74). At one site,
274 the saline Lake Alchichica, the population is currently considered a distinct species (*A. taylori*)
275 (10). However, our results indicate that *A. taylori* is not genetically distinct from surrounding,
276 putatively non-saline-adapted populations, a result consistent with a previous microsatellite-
277 based study that found significant gene flow to and from *A. taylori* (29). While we cannot rule
278 out the potential for localized signatures of selection and adaptation in the *A. taylori* genome, our
279 results highlight that even populations adapted to aquatic conditions intolerable to most other
280 *Ambystoma* have not reached a level of isolation consistent with an independent evolutionary
281 trajectory (75).

282

283 *Comparisons to previous work*

284 This study included extensive geographic sampling of the tiger salamander complex,
285 which provided a broad spatial context to more fully understand patterns of genetic variation.
286 However, it is important to consider the possibility that our data are simply not sensitive enough

287 to detect genetic differentiation associated with lineage divergence on an extremely recent time
288 scale. The common ancestor of *A. californiense* and the remainder of the species complex dates
289 to approximately five million years (11, 37), and phylogenomic data indicate that speciation
290 across the remaining tiger salamander lineages occurred within the last one million years (76).
291 Such timing could make lineage boundaries difficult to detect (15). In this context, our genome-
292 wide markers, developed from transcriptomic resources (35), may not have an overall
293 substitution rate sufficient to detect genetic differentiation. We recommend that future research
294 uses a larger genomic dataset or faster-evolving loci with particular focus on the sampling of
295 obligate paedomorphs in this system. To this point, perhaps the strongest evidence for the genetic
296 divergence of obligately paedomorphic populations comes from past microsatellite-based work
297 indicating the genetic distinctiveness of *A. andersoni* and *A. mexicanum* from a select set of
298 populations across central Mexico (77) and the genetic distinctiveness of *A. taylori* from
299 neighboring populations in the Cuenca Oriental (29). Both of these studies recovered signatures
300 of gene flow, but identified greater overall population structure compared to this study. These
301 previous results could be due to numerous factors, including a faster microsatellite mutation rate,
302 methodological differences, and more limited locality sampling.
303

304 **CONCLUSIONS**

305 The extent to which populations within the tiger salamander complex exhibit phenotypic
306 plasticity in life history traits is remarkable and is believed to have played a role in the rapid
307 accumulation of lineages observed in the highlands of central Mexico (11, 61). While our results
308 suggest there is less species-level diversity in that region than previously recognized, there is
309 clearly more diversity in central Mexico compared to the US and Canada where there is (a) more

310 geographic space and (b) less life history variation within and between lineages. While we
311 cannot fully explain the greater diversity in Central Mexico, our results suggest that major
312 patterns of diversification are related to a complex history of geographic isolation and secondary
313 contact, in which life history strategy has played a less important role. We agree with previous
314 work (12) that the complex geological history of the TMVB, including montane uplift and
315 fluctuating drainage connectivity since the Miocene, has been the cornerstone of the evolutionary
316 history of this species complex, and that the influences of geographic isolation and
317 paedomorphosis may work synergistically to lead to the establishment of isolated populations.
318 Lingering questions notwithstanding, this large-scale genetic and geographic study has
319 established a framework for understanding the evolutionary history of the *Ambystoma tigrinum*
320 species complex. The results presented here will facilitate comparative studies of the axolotl and
321 its allies, provide direction for conservation prioritization and management, and strengthen the
322 use of the tiger salamander species complex as a model system in biology.
323

324 **MATERIALS & METHODS**

325 *Geographic sampling*

326 We generated data from a total of 254 individuals sampled from across the range of the
327 *A. tigrinum* complex (Fig. S1; Table S1). These individuals were combined with 93 individuals
328 sampled in O'Neill et al. (35) to produce a data set comprising 347 individuals. To represent the
329 large geographic range of this species complex, we chose to sample a large number of localities
330 (188) with limited numbers of individuals sampled per locality (mean = 1.8, min. = 1, max. = 9).
331 This sampling included a total of 166 individuals from the US, 2 individuals from Canada, and
332 178 individuals sampled from Mexico. Individuals of *A. californiense* were sampled from

333 localities with limited to no impacts of introgression from invasive western tiger salamanders
334 (52). Additional outgroup data were generated for two species outside of the *A. tigrinum*
335 complex: *A. opacum* and *A. texanum*. Full details regarding the generation of these data can be
336 found in the Supplementary Materials.
337
338 *Data collection and sequencing*
339 We generated DNA sequence data from a panel of 95 nuclear loci developed specifically
340 for the tiger salamander species complex. A more complete description of marker development
341 can be found in O'Neill et al. (35). Genomic DNA was extracted using a DNeasy Blood and
342 Tissue kit (Qiagen). For a small number of DNA extractions, we increased DNA quantities using
343 a Repli-g whole genome amplification kit (Qiagen). We used an initial round of PCR in 96-well
344 plate format to amplify all loci from an individual, followed by a smaller second round of PCR to
345 amplify loci that did not amplify in the initial PCR. See O'Neill et al. (35) for the details of PCR
346 conditions and primer sequences.
347 PCR products from all loci were pooled for each individual in roughly equal
348 concentrations based on the intensity of amplification as visualized on an agarose gel. Indexed
349 Illumina sequencing libraries were generated for each individual using an Illumina Nextera XT
350 DNA Library Preparation Kit (Illumina). Subsequent to library preparation, indexed libraries
351 were quantified using a Qubit Fluorometer, pooled in equimolar concentrations, and checked on
352 an Agilent Bioanalyzer to assure proper fragmentation.
353 Sequencing was performed in four rounds using an Illumina MiSeq. We performed an
354 initial round using a total of seven individuals to test the compatibility of the Illumina Nextera
355 XT library kit with our PCR amplicons. Three subsequent rounds of library preparation and

356 sequencing were performed on sets of 96 individuals each. Note that some individuals were
357 sequenced in multiple rounds due to initially low read counts. All sequencing was performed
358 with paired-end 150 bp reads. Overall, we generated a total of 29,426,894 PE reads across all
359 newly sequenced individuals, with an average of 309,757 PE reads, 1,148X coverage, and 7.6%
360 missing loci per individual (Table S2).
361
362 *Bioinformatics and dataset generation*
363 All sequence reads were processed using a newly developed bioinformatic pipeline
364 written for this project (https://doi.org/10.5281/zenodo.3585970) that produces multiple
365 sequence alignments for individual loci and genome-wide SNP matrices sampled from variable
366 sites. This pipeline was developed using the Snakemake workflow management system (78),
367 linking together multiple software tools to take sequence data from raw reads to phased sequence
368 alignments for each locus. Briefly, as used in this study, demultiplexed paired-end Illumina fastq
369 files were used as input, with separate forward and reverse read files for each individual.
370 Sequence data were trimmed and filtered in Trimmomatic (79) using a sliding window of 4 base-
371 pairs and a minimum average quality score of 15. Filtered sequence reads were then aligned to
372 reference sequences from the O'Neill et al. (2013) dataset (35); specifically, we used the clean
373 sequences of an *A. ordinarium* sample that had high coverage and low amounts of missing data.
374 The resulting aligned contigs were processed using SAMtools (80) to filter and prepare data for
375 FreeBayes (81), which was used to call variable sites. Variants were filtered with VCFTools (82)
376 by removing indels and setting a quality threshold of phred score > 20 and a minimum read
377 depth of 30. The program WhatsHap (83) was used to perform read-based phasing of the data for
378 each locus x individual contig. Finally, phased haplotypes from each individual (two copies,

379 regardless of homo- or heterozygosity) were combined into an alignment of all individuals using
380 MAFFT with the default auto parameter (84). We generated fasta files of SNPs using the SNP-
381 sites program (85) and created a SNP genotype matrix by sampling variable sites from a
382 concatenated sequence alignment of all loci. Through our informatic processing and manual
383 inspections of the data, we identified three loci (E12G1, E6A11, and E7G8) as potential paralogs
384 based on high alignment error and high levels of heterozygosity for all individuals. These three
385 loci were excluded from all analyses.

386 Across the remaining 92 loci, alignment lengths ranged from 123 to 630 bp (avg. = 270
387 bp) with a total concatenated alignment of 24,788 bp. For the full dataset, including *A. texanum*
388 and *A. opacum* outgroups, single-locus alignments contained an average of 55 variable sites
389 (min. = 15, max. = 121) and an average of 41 parsimony informative sites (min. = 11, max. =
390 102). Population genetic analyses restricted to the *A. tigrinum* complex including *A.*
391 *californiense*—which were further filtered for non-biallelic SNPs and a minor allele count ≥ 3—
392 contained a total of 2,360 SNPs.

393

394 *Population structure and lineage discovery*

395 We developed hypotheses of population-level lineages across the range of the *A. tigrinum*
396 complex using a bottom-up approach, starting with the identification of major geographic
397 patterns of differentiation, and then performing a recursive set of analyses on more restricted sets
398 of individuals, as identified in the previous analytical step. In our initial round of analyses we
399 used two non-parametric methods: principal components analysis (PCA) and discriminant
400 analysis of principal components (DAPC) (36). While both analyses provide a multivariate
401 summary of genetic data, DAPC is also used to assess the fit of data to varying numbers of

402 population clusters. These analyses were applied to our full genotypic dataset including *A.*
403 *californiense* and all remaining individuals from the *A. tigrinum* complex. The PCA was
404 calculated using the function 'prcomp' in the R package stats (86), while the DAPC was
405 calculated using the package adegenet (87). The optimal number of principal components to
406 retain for DAPC was identified using cross-validation via the xvalDapc function with default
407 parameter values. DAPC was performed without prior assignment of individuals to groups and
408 an exploratory approach was used to identify patterns of differentiation in ordination space
409 across a range of cluster levels (K = 1-20). We used two metrics to identify the best estimate of
410 the primary splits in our data. First, we used the Bayesian Information Criterion (BIC) calculated
411 in the DAPC analysis to assess the fit of the data to different levels of K. We note that the level
412 of K with the absolute lowest BIC may not be a better explanation of the data than a K with a
413 slightly higher BIC (88); therefore, we applied this measure for general guidance on a range of K
414 that may describe the data well. We paired this assessment with visualizations of the first and
415 second principal components (Fig. S2, S4) and DAPC ordination plots to identify the level of K
416 at which similar clustering patterns could be observed with minimal change at successively
417 higher levels of K. DAPC of the complete tiger salamander species complex identified a
418 consistent pattern beginning at K = 5 for high differentiation of all *A. californiense* samples (Fig.
419 S3). Further DAPC analysis with *A. californiense* removed identified a consistent pattern
420 beginning at K = 5 for differentiation between clusters of populations from northern and central
421 Mexico, and three clusters of U.S. populations (two from the Western U.S. and one from the
422 Eastern U.S.; Fig. S5).
423 Using the clusters identified in the DAPC analysis of the total data set, we then used both
424 DAPC and the program STRUCTURE v.2.3.4 (89) to analyze subsequent data sets comprising

425 smaller numbers of individuals. Recursive rounds of DAPC analyses were performed as
426 described above and were stopped when BIC scores showed little improvement (ΔBIC < 2) at
427 values of $K > 1$. STRUCTURE analyses were performed using an admixture model and a total of
428 500,000 generations following a burn-in of 100,000 generations. Analyses were performed for K
429 = 1-10 with 16 replicate analyses. We identified a "best" value of K using two approaches. First,
430 we calculated ΔK using the Evanno method (91) via the CLUMPAK web tool (90). A limitation
431 of the Evanno method is that it cannot identify the best value of $K = 1$ (91), thus, we also visually
432 inspected individual group assignments and concluded a value of $K = 1$ if the corresponding
433 DAPC cluster showed little improvement (ΔBIC < 2) at values of $K > 1$, or when visualization
434 showed extensive admixture without distinct clustering.
435
436 *Tests of migration and population structure*
437 We used the coalescent-based program Migrate-N v.3.2 (92) to explicitly test for
438 population structure and gene flow in each obligately paedomorphic species of *Ambystoma* (*A.*
439 *andersoni, A. dumerilii, A. taylori,* and *A. mexicanum,* plus the "often neotenic" species *A.*
440 *lermaense*). For each model we evaluated (Fig. 4A), we treated the obligately paedomorphic
441 species as one population, and the facultatively paedomorphic individuals from the same genetic
442 cluster (CM2-CM4) as the second population. The best-fitting model was determined via Bayes
443 factors (BF), which were calculated using Bézier-corrected marginal likelihoods (Table S2); the
444 highest-ranking models have a BF = 0. Each Migrate-N analyses ran for 50 million steps,
445 recording every 10 steps, with a burn-in of five million and the default heating scheme. Suitable
446 upper bounds for priors on population size (Θ) and migration rate (M) were determined from an

447 initial test run of 10 million steps for each analysis. Median values and 95% confidence intervals
448 of Θ and *M* are reported (Fig. 4B).
449
450 *Phylogenetic reconstruction*
451 As our dataset included a large number of admixed individuals, we estimated
452 phylogenetic networks to visualize reticulation events. We used SplitsTree v. 4.14.8 (40, 93) to
453 generate four networks: one with all tiger salamander individuals, and one each for the U.S.,
454 central Mexico, and northern Mexico subgroups. Networks were constructed using uncorrected
455 p-distances and the NeighborNet algorithm (94).
456 We also used three different analytical approaches to place hypothesized population
457 lineages in a phylogenetic framework. For all analyses we used a reduced data set containing the
458 concatenated data for 2-7 representative individuals from each population genetic cluster, which
459 limited computation time and avoided violating the coalescent-model assumption of zero gene
460 flow. We first inferred the phylogeny using Bayesian Inference in BEAST v.1.8.3 (42). Analyses
461 were run for 5 million generations, sampling every 1,000 generations after the first 500,000
462 generations were removed as burn-in. Run convergence was assessed with Tracer v.1.6.0 (95).
463 Next, we inferred a maximum-likelihood phylogeny using RAxML v.8 (43). Node support was
464 assessed using a rapid bootstrap analysis with 1,000 replicates, which was summarized as a 95%
465 rule consensus tree using the program SumTrees in the DendroPy python library (96). For both
466 BEAST and RAxML, PartitionFinder (97) was used to identify the number of preferred gene
467 partitions and their substitution models, and analyses were performed on the CIPRES Science
468 Gateway server (98). Finally, we inferred phylogenetic relationships using SVDquartets (41)
469 implemented in PAUP* version 4.0a164 (99), sampling all possible quartets and assessing node

support with 1,000 bootstrap replicates. For all phylogenetic analyses, trees were visualized using FigTree v.1.4.2 (100).

DATA AVAILABILITY

Supplementary figures and tables are provided with the online version of this manuscript. Input files for all population genetic and phylogenetic analyses are available via figshare (https://figshare.com/projects/Life_history_strategy_does_not_reflect_genetic_differentiation_in_the_tiger_salamander_species_complex/74115). Sequence data are available on the NCBI Sequence Read Archive (BioProject accession PRJNA594660).

ACKNOWLEDGEMENTS

This project was supported by grants from the National Science Foundation (DEB-0949532, DEB-1355000, and a Doctoral Dissertation Improvement Grant awarded to Justin Kratovil) and by a University Research Fellowship awarded to Kathryn Everson from the University of Kentucky. We thank numerous individuals for contributing samples to this project, including Chris Beachy, Jim Bogart, Ken Carbale, Sheri Church, Jeff LeClere, Carol Hall, Paul Moler, Nancy Staub, and Ken Wray. We are grateful to Jose Bocanegra, Ben Browning, Mackenzie Humphrey, Mary Virginia Gibbs, Ricky Grewelle, Alan Lemmon, Emily Moriarty Lemmon, Deborah Lu, Stephanie Mitchell, Alex Noble, Tolu Odukoya, Ben Tuttle, and Josh Williams for their assistance with data collection and analyses.

References

1. S. R. Palumbi, Marine speciation on a small planet. *Trends Ecol. Evol.* **7**, 114–118 (1992).
2. S. Mopper, S. Y. Strauss, *Genetic Structure and Local Adaptation in Natural Insect Populations: Effects of Ecology, Life History, and Behavior* (Springer US, 1998).
3. S. J. Gould, *Ontogeny and Phylogeny* (Harvard University Press, 1977).
4. D. B. Wake, Paedomorphosis. *J. Herpetol.* **14**, 80 (1980).
5. H. H. Whiteman, Evolution of facultative paedomorphosis in salamanders. *Q. Rev. Biol.* **69**, 205–221 (1994).
6. R. B. Page, *et al.*, Effect of thyroid hormone concentration on the transcriptional response underlying induced metamorphosis in the Mexican axolotl (Ambystoma). *BMC Genomics* **9**, 78 (2008).
7. S. R. Voss, Genetic Basis of Paedomorphosis in the Axolotl, Ambystoma mexicanum: A Test of the Single-Gene Hypothesis. *J. Hered.* **86**, 441–447 (1995).
8. S. R. Voss, H. B. Shaffer, Adaptive evolution via a major gene effect: paedomorphosis in the Mexican axolotl. *Proc. Natl. Acad. Sci. USA* **94**, 14185–9 (1997).
9. E. H. Taylor, New salamanders from Mexico with a discussion of certain known forms. *Univ. Kansas Sci. Bull.* **26**, 407–439 (1939).
10. R. Brandon, E. Maruska, W. Rumph, A new species of neotenic *Ambystoma* (Amphibia, Caudata) endemic to Laguna Alchichica, Puebla, Mexico. *Bull. South. Calif. Acad. Sci.* **80**, 112–125 (1981).
11. H. B. Shaffer, M. L. McKnight, The polytypic species revisited: Genetic differentiation and molecular phylogenetics of the tiger salamander *Ambystoma tigrinum* (Amphibia: Caudata) complex. *Evolution* **50**, 417–433 (1996).

514 12. H. B. Shaffer, Evolution in a paedomorphic lineage. I. An electrophoretic analysis of the
515 Mexican ambystomatid salamanders. *Evolution* **38**, 1194–1206 (1984).
516 13. H. B. Shaffer, Phylogenetics of model organisms: The laboratory axolotl, *Ambystoma*
517 *Mexicanum*. *Syst*. *Biol*. **42**, 508–522 (1993).
518 14. H. B. Shaffer, Evolution in a paedomorphic lineage. II. Allometry and form in the
519 Mexican ambystomatid salamanders. *Evolution* **38**, 1207–1218 (1984).
520 15. H. B. Shaffer, R. C. Thomson, Delimiting Species in Recent Radiations. *Syst*. *Biol*. **56**,
521 896–906 (2007).
522 16. P. R. Grant, Speciation and the adaptive radiation of Darwin's finches. *Am*. *Sci*. **69**, 653–
523 663 (1981).
524 17. O. Seehausen, Explosive speciation rates and unusual species richness in haplochromine
525 cichlid fishes: Effects of sexual selection. *Adv*. *Ecol*. *Res*. **31**, 237–274 (2000).
526 18. M. Fegan, P. Prior, "How complex is the *Ralstonia solanacearum* species complex" in
527 *Bacterial Wilt Disease and the Ralstonia Solanacearum Species Complex*, C. Allen, P.
528 Prior, A. C. Hayward, Eds. (American Phytopathological Society Press, 2005), pp. 449–
529 461.
530 19. J. Mallet, Hybrid speciation. *Nature* **446**, 279–283 (2007).
531 20. M. C. Fontaine, *et al*., Extensive introgression in a malaria vector species complex
532 revealed by phylogenomics. *Science* **347**, 1258524 (2015).
533 21. J. Geml, G. A. Laursen, K. O'Neill, H. C. Nusbaum, D. L. Taylor, Beringian origins and
534 cryptic speciation events in the fly agaric (*Amanita muscaria*). *Mol*. *Ecol*. **15**, 225–239
535 (2005).
536 22. K. S. Jones, D. W. Weisrock, Genomic data reject the hypothesis of sympatric ecological

537 speciation in a clade of *Desmognathus* salamanders. *Evolution* **72**, 2378–2393 (2018).
538 23. R. Newman, Adaptive plasticity in amphibian metamorphosis. *Bioscience* **42**, 671–678
539 (1992).
540 24. H. H. Whiteman, S. A. Wissinger, W. S. Brown, Growth and foraging consequences of
541 facultative paedomorphosis in the tiger salamander, *Ambystoma tigrinum nebulosum*.
542 *Evol. Ecol.* **10**, 433–446 (1996).
543 25. P. J. Fernandez, J. P. Collins, Effect of environment and ontogeny on color pattern
544 variation in Arizona tiger salamanders (*Ambystoma tigrinum nebulosum* Hallowell).
545 *Copeia* **1988**, 928 (1988).
546 26. S. P. D. Riley, H. B. Shaffer, S. R. Voss, B. M. Fitzpatrick, Hybridization between a rare,
547 native tiger salamander (*Ambystoma californiense*) and its introduced congener. *Ecol.*
548 *Appl.* **13**, 1263–1275 (2003).
549 27. D. W. Weisrock, H. B. Shaffer, B. L. Storz, S. R. Storz, S. R. Voss, Multiple nuclear gene
550 sequences identify phylogenetic species boundaries in the rapidly radiating clade of
551 Mexican ambystomatid salamanders. *Mol. Ecol.* **15**, 2489–2503 (2006).
552 28. G. Parra-Olea, *et al.*, Conservation genetics of threatened Mexican axolotls (*Ambystoma*).
553 *Anim. Conserv.* **15**, 61–72 (2012).
554 29. R. Percino-Daniel, E. Recuero, E. Vázquez-Domínguez, K. R. Zamudio, G. Parra-Olea,
555 All grown-up and nowhere to go: Paedomorphosis and local adaptation in *Ambystoma*
556 salamanders in the Cuenca Oriental of Mexico. *Biol. J. Linn. Soc.* **118**, 582–597 (2016).
557 30. R. A. Brandon, Hybridization between the Mexican salamanders *Ambystoma dumerilii*
558 and *Ambystoma mexicanum* under laboratory conditions. *Herpetologica* **28**, 199–207
559 (1972).

31. R. A. Brandon, Interspecific hybridization among Mexican and United States salamanders of the genus *Ambystoma* under laboratory conditions. *Herpetologica* **33**, 133–152 (1977).

32. S. Krebs, R. Brandon, A new species of salamander (family Ambystomatidae) from Michoacan, Mexico. *Herpetologica* **40**, 238–245 (1984).

33. J. B. Armstrong, G. M. Malacinski, *Developmental Biology of the Axolotl* (Oxford University Press, 1989).

34. L. L. Bailey, T. R. Simons, K. H. Pollock, Estimating site occupancy and species detection probability parameters for terrestrial salamanders. *Ecol. Appl.* **14**, 692–702 (2004).

35. E. M. O'Neill, *et al.*, Parallel tagged amplicon sequencing reveals major lineages and phylogenetic structure in the North American tiger salamander (*Ambystoma tigrinum*) species complex. *Mol. Ecol.* **22**, 111–129 (2013).

36. T. Jombart, S. Devillard, F. Balloux, Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genet.* **11**, 94 (2010).

37. H. B. Shaffer, G. B. Pauly, J. C. Oliver, P. C. Trenham, The molecular phylogenetics of endangerment: Cryptic variation and historical phylogeography of the California tiger salamander, *Ambystoma californiense*. *Mol. Ecol.* **13**, 3033–3049 (2004).

38. J. W. Petranka, *Salamanders of the United States and Canada* (Smithsonian Books, 2010).

39. , AmphibiaWeb. *Univ. California, Berkeley, CA* (2020) (May 11, 2020).

40. D. H. Huson, SplitsTree: Analyzing and visualizing evolutionary data. *Bioinformatics* **14**, 68–73 (1998).

41. J. Chifman, L. Kubatko, Quartet inference from SNP data under the coalescent model.

Bioinformatics **30**, 3317–3324 (2014).

42. A. J. Drummond, A. Rambaut, BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).

43. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

44. A. A. Dugès, *Description d'un axolotl des montagnes de las cruces (*Amblystoma *[sic]* altamirani*, A. Dugès)* (Imprimerie du Ministère de "Fomento," 1895).

45. E. H. Taylor, Herpetological novelties from Mexico. *Univ. Kansas Sci. Bull.* **29**, 343–358 (1943).

46. E. Dunn, A new genus of salamanders from Mexico. *Proc. New Engl. Zool. Club* **10**, 85–86 (1928).

47. E. H. Taylor, A new ambystomid salamander from the Plateau Region of Mexico. *Univ. Kansas Sci. Bull.* **30**, 57–61 (1944).

48. A. A. Dugès, Una nueva especie de ajolote de la Laguna de Patzcuaro. *La Nat.* **1**, 241–244 (1870).

49. J. Dixon, A new species of salamander of the genus Ambystoma from Jalisco, Mexico. *Copeia* **1**, 99–101 (1963).

50. G. Shaw, F. Nodder, *The Naturalist's Miscellany; or Coloured Figures of Natural Objects Drawn and Described Immediately from Nature*, 9th Ed. (Nodder & Co., 1798).

51. A. A. Dugès, Erpetología del valle de México. *La Nat.* **2**, 97–146 (1888).

52. J. R. Johnson, R. C. Thomson, S. J. Micheletti, H. B. Shaffer, The origin of tiger salamander (*Ambystoma tigrinum*) populations in California, Oregon, and Nevada: Introductions or relicts? *Conserv. Genet.* **12**, 355–370 (2011).

53. E. H. Taylor, Two new ambystomid salamanders from Chihuahua. *Copeia* **3**, 143–146 (1941).

54. R. Webb, Observations on tiger salamanders (*Ambystoma tigrinum* complex, family Ambystomatidae) in Mexico with description of a new species. *Bull. Maryl. Herpetol. Soc.* **40**, 122–143 (2004).

55. R. A. Brandon, Spontaneous and induced metamorphosis of *Ambystoma dumerilii* (Dugès), a paedogenetic Mexican salamander, under laboratory conditions. *Herpetologica* **32**, 429–438 (1976).

56. H. M. Wilbur, J. P. Collins, Ecological aspects of amphibian metamorphosis: Nonnormal distributions of competitive ability reflect selection for facultative metamorphosis. *Science* **182**, 1305–1314 (1973).

57. R. M. Bonett, M. A. Steffen, S. M. Lambert, J. J. Wiens, P. T. Chippindale, Evolution of paedomorphosis in plethodontid salamanders: Ecological correlates and re-evolution of metamorphosis. *Evolution* **68**, 466–482 (2014).

58. C. R. Marshall, E. C. Raff, R. A. Raff, Dollo's law and the death and resurrection of genes. *Proc. Natl. Acad. Sci. USA* **91**, 12283–12287 (1994).

59. M. J. West-Eberhard, Phenotypic Plasticity and the Origins of Diversity. *Annu. Rev. Ecol. Syst.* **20**, 249–278 (1989).

60. M. J. West-Eberhard, Alternative adaptations, speciation, and phylogeny (A Review). *Proc. Natl. Acad. Sci. USA* **83**, 1388–92 (1986).

61. B. M. Fitzpatrick, Underappreciated consequences of phenotypic plasticity for ecological speciation. *Int. J. Ecol.* **2012**, 1–12 (2012).

62. M. Denoël, P. Poncin, J.-C. Ruwet, Sexual compatibility between two heterochronic

morphs in the Alpine newt, *Triturus alpestris*. *Anim. Behav*. **62**, 559–566 (2001).

63. J. D. Krenz, P. A. Verrell, Integrity in the midst of sympatry: Does sexual incompatibility facilitate the coexistence of metamorphic and paedomorphic mole salamanders (*Ambystoma talpoideum*)? *J. Zool*. **258**, S0952836902001589 (2002).

64. H. H. Whiteman, J. D. Krenz, R. D. Semlitsch, Intermorph breeding and the potential for reproductive isolation in polymorphic mole salamanders (*Ambystoma talpoideum*). *Behav. Ecol. Sociobiol*. **60**, 52–61 (2006).

65. H. Whiteman, J. Gutrich, R. Moorman, Courtship behavior in a polymorphic population of the tiger salamander, *Ambystoma tigrinum nebulosum*. *J. Herpetol*. **33**, 348–351 (1999).

66. R. . Howard, R. . Moorman, H. . Whiteman, Differential effects of mate competition and mate choice on eastern tiger salamanders. *Anim. Behav*. **53**, 1345–1356 (1997).

67. D. H. Bos, R. N. Williams, D. Gopurenko, Z. Bulut, J. A. DeWoody, Condition-dependent mate choice and a reproductive disadvantage for MHC-divergent male tiger salamanders. *Mol. Ecol*. **18**, 3307–3315 (2009).

68. C. E. Nelson, R. R. Humphrey, Artificial interspecific hybridization among *Ambystoma*. *Herpetologica* **28**, 27–32 (1972).

69. J. M. Velasco, Anotaciones y observaciones al trabajo del Sr. D. A. Weismann sobre la transformación del ajolote Mexicano en Amblistoma. *La Nat*. **5**, 58–84 (1880).

70. H. M. Smith, The Mexican axolotl: Some misconceptions and problems. *Bioscience* **19**, 593–615 (1969).

71. A. G. E. Mathiron, J.-P. Lena, S. Baouch, M. Denoël, The ‘male escape hypothesis’: Sex-biased metamorphosis in response to climatic drivers in a facultatively paedomorphic amphibian. *Proc. R. Soc. B Biol. Sci*. **284**, 20170176 (2017).

652 72. C. L. Morjan, L. H. Rieseberg, How species evolve collectively: Implications of gene flow
653 and selection for the spread of advantageous alleles. *Mol. Ecol.* **13**, 1341–1356 (2004).
654 73. J. L. Tamayo, R. C. West, "The hydrography of middle America" in *Handbook of Middle*
655 *American Indians*, R. Wauchope, R. C. West, Eds. (University of Texas Press, 1964), pp.
656 84–121.
657 74. G. R. Hopkins, E. D. Brodie, Occurrence of amphibians in saline habitats: A review and
658 evolutionary perspective. *Herpetol. Monogr.* **29**, 1–27 (2015).
659 75. K. De Queiroz, Species concepts and species delimitation. *Syst. Biol.* **56**, 879–886 (2007).
660 76. P. M. Hime, *et al.*, Phylogenomics reveals ancient gene tree discordance in the amphibian
661 tree of life. *Syst. Biol.* (2020) https:/doi.org/10.1093/sysbio/syaa034 (May 12, 2020).
662 77. G. Parra-Olea, *et al.*, Conservation genetics of threatened Mexican axolotls (*Ambystoma*).
663 *Anim. Conserv.* **15**, 61–72 (2012).
664 78. J. Koster, S. Rahmann, Snakemake: A scalable bioinformatics workflow engine.
665 *Bioinformatics* **28**, 2520–2522 (2012).
666 79. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina
667 sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
668 80. H. Li, A statistical framework for SNP calling, mutation discovery, association mapping
669 and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**,
670 2987–2993 (2011).
671 81. E. Garrison, G. Marth, Haplotype-based variant detection from short-read sequencing.
672 *arXiv* **1207.3907** (2012).
673 82. P. Danecek, *et al.*, The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158
674 (2011).

675 83. M. Martin, *et al*., WhatsHap: Fast and accurate read-based phasing. *bioRxiv* **10.1101/08**
676 (2016).
677 84. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7:
678 Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
679 85. A. J. Page, *et al*., SNP-sites: Rapid efficient extraction of SNPs from multi-FASTA
680 alignments. *Microb. genomics* **2**, e000056 (2016).
681 86. R. C. Team, R: a language and environment for statistical computing (2012).
682 87. T. Jombart, adegenet: A R package for the multivariate analysis of genetic markers.
683 *Bioinformatics* **24**, 1403–1405 (2008).
684 88. T. Jombart, C. Collins, A tutorial for discriminant analysis of principal components
685 (DAPC) using adegenet 2.0. 0. *adegenet.r-forge.r-project.org* (2015) (May 30, 2019).
686 89. J. K. J. Pritchard, M. Stephens, P. Donnelly, Inference of population structure using
687 multilocus genotype data. *Genetics* **155**, 945–959 (2000).
688 90. N. M. Kopelman, J. Mayzel, M. Jakobsson, N. A. Rosenberg, I. Mayrose, Clumpak: A
689 program for identifying clustering modes and packaging population structure inferences
690 across K. *Mol. Ecol. Resour.* **15**, 1179–1191 (2015).
691 91. G. Evanno, S. Regnaut, J. Goudet, Detecting the number of clusters of individuals using
692 the software structure: A simulation study. *Mol. Ecol.* **14**, 2611–2620 (2005).
693 92. P. Beerli, M. Palczewski, Unified framework to evaluate panmixia and migration direction
694 among multiple sampling locations. *Genetics* **185**, 313–326 (2010).
695 93. D. H. Huson, D. Bryant, Application of phylogenetic networks in evolutionary studies.
696 *Mol. Biol. Evol.* **23**, 254–267 (2006).
697 94. D. Bryant, V. Moulton, Neighbor-Net: An agglomerative method for the construction of

phylogenetic networks. *Mol. Biol. Evol.* **21**, 255–265 (2003).

95. A. Rambaut, A. Drummond, D. Xie, G. Baele, M. Suchard, Tracer v1.7 (2018).

96. J. Sukumaran, M. Holder, DendroPy: A Python library for phylogenetic computing. *Bioinformatics* **26**, 1569–1571 (2010).

97. R. Lanfear, B. Calcott, S. Y. W. Ho, S. Guindon, PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol Biol Evol* **29**, 1695–1701 (2012).

98. M. A. Miller, W. Pfeiffer, T. Schwartz, Creating the CIPRES Science Gateway for inference of large phylogenetic trees in *2010 Gateway Computing Environments Workshop (GCE)*, (IEEE, 2010), pp. 1–8.

99. D. Swofford, PAUP 4.0 b10: Phylogenetic analysis using parsimony (2002) (December 21, 2013).

100. A. Rambaut, FigTree, a graphical viewer of phylogenetic trees (2007) (May 8, 2014).