Molecular Phylogenetics and Evolution 144 (2020) 106696

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Evolutionary divergences mirror Pleistocene paleodrainages in a rapidly-evolving complex of oasis-dwelling jumping spiders (Salticidae, *Habronattus tarsalis*)

Marshal Hedin[a,*], Steven Foldi[a,b], Brendan Rajah-Boyer[a,c]

[a] Dept of Biology, San Diego State University, San Diego, CA 92182, United States
[b] Valley International Prep School, Chatsworth, CA 91311, United States
[c] St. Francis of Assisi Catholic School, Vista, CA 92083, United States

ARTICLE INFO

Keywords:
Biogeography
Ephemeral speciation
Introgression
Pleistocene
Saline Valley
Sexual selection

ABSTRACT

We aimed to understand the diversification history of jumping spiders in the *Habronattus tarsalis* species complex, with particular emphasis on how history in this system might illuminate biogeographic patterns and processes in deserts of the western United States. Desert populations of *H. tarsalis* are now confined to highly discontinuous oasis-like habitats, but these habitats would have been periodically more connected during multiple pluvial periods of the Pleistocene. We estimated divergence times using relaxed molecular clock analyses of published transcriptome datasets. Geographic patterns of diversification history were assessed using phylogenetic and cluster analyses of original sequence capture, RADSeq and morphological data. Clock analyses of multiple replicate transcriptome datasets suggest mid- to late-Pleistocene divergence dates within the *H. tarsalis* group complex. Coalescent and concatenated phylogenetic analyses indicate four early-diverging lineages (*H. mustaciata*, *H. ophrys*, and *H. tarsalis* from the Lahontan and Owens drainage basins), with remaining samples separated into larger clades from the Mojave desert, and western populations from the California Floristic Province of California and northern Baja California. Focusing on desert populations, there is a strong correspondence between RAD lineages and modern and/or paleodrainages, mirrored more finely in STRUCTURE and machine learning results. Non-metric multidimensional scaling analysis reveals strong congruence between morphological clusters and genetic lineages, whether the latter represent previously described species or *H. tarsalis* RAD lineages. Here we have uncovered a system that adds to our regional biogeographic knowledge in unique ways, using multiple types of evidence in a broadly-distributed terrestrial taxon. At the same time, we have discovered rapid evolution of both novel morphological forms and diverging genetic lineages. The hierarchical nature of variation in the *H. tarsalis* complex, the minute range sizes of many forms, the high likelihood that geographic distributions have shrunk and expanded through time, and signs of introgression all align with an ephemeral speciation model.

> "We are now convening deep below the surface level on an ancient inland sea."
>
> Hubbs and Miller (1948)

1. Introduction

Low elevation desert habitats of the western United States represent some of the harshest modern terrestrial environments on Earth. For example, average high air temperatures in Death Valley often exceed 37 °C for multiple summer months, while precipitation is essentially non-existent for over half the year (Roof and Callagan, 2003). For some taxa, this xeric harshness is a relatively recent ecological and evolutionary phenomenon, as large pluvial lakes connected by extensive riparian corridors occupied these same regions during various phases of the late Pliocene and Pleistocene. In the Death Valley region, the massive late-Pleistocene Lake Manly was fed by riparian connections from the east (Amargosa drainage), west (Owens drainage), and south (Mojave drainage; Fig. 1). Further north in the Great Basin, the Lahontan and Bonneville basins were at times filled with expansive interconnected pluvial lakes, with multiple drainages feeding into basins without outlets. Profoundly isolated basins also occur in the central Great Basin, occupied by Pleistocene pluvial lakes without connections

* Corresponding author.
E-mail address: mhedin@sdsu.edu (M. Hedin).

https://doi.org/10.1016/j.ympev.2019.106696
Received 19 August 2019; Received in revised form 14 November 2019; Accepted 27 November 2019
Available online 29 November 2019

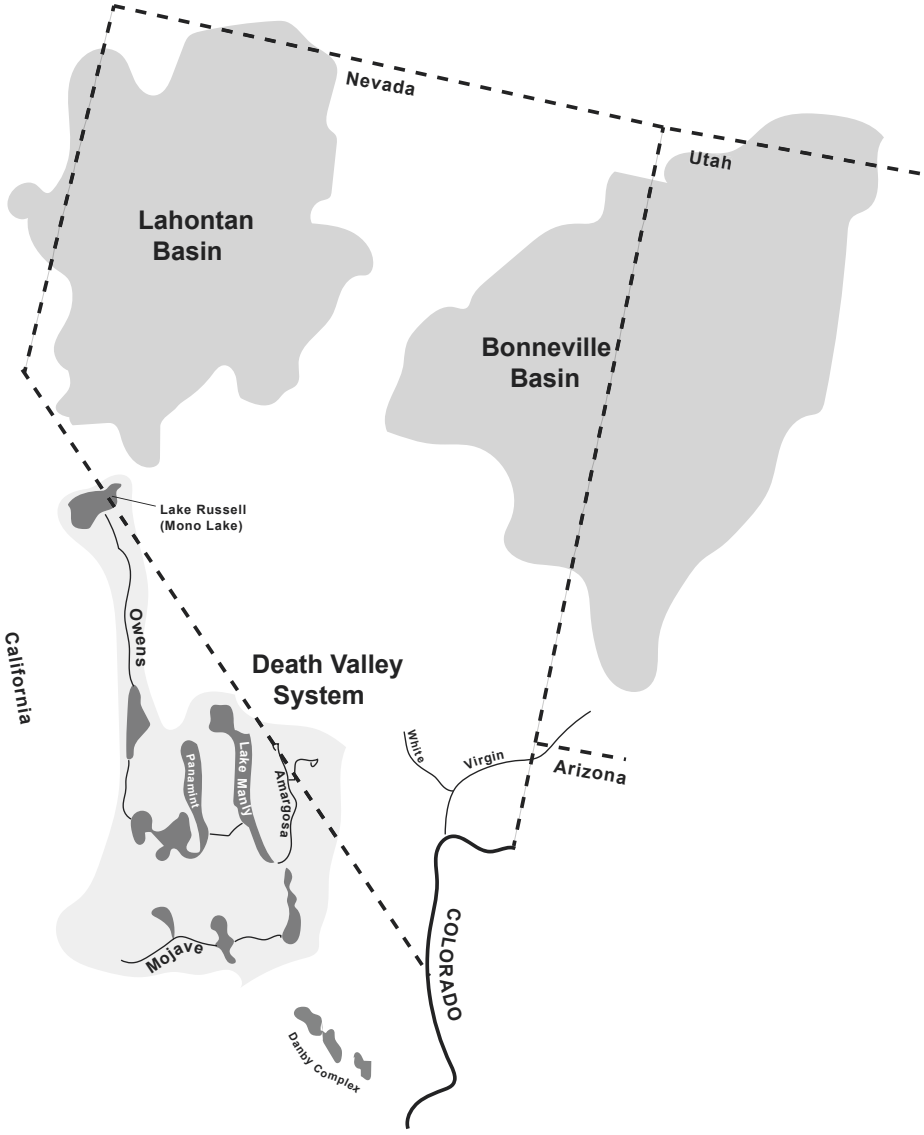


Fig. 1. Map of the western United States showing prominent Pleistocene lakes and paleodrainages as discussed in text. Geographic position and pluvial extents for Lahontan and Bonneville basins only illustrate general geographic position. Timeframe represents a non-specific high-water period from the late Pleistocene, with Lake Russell draining southwards. Modified from several sources (e.g., Knott et al., 2008; Reheis et al., 2014).

to surrounding basins, as evidenced by a lack of native fishes. Hubbs and Miller (1948) referred to these as "sterile basins". Pleistocene lakes and riparian corridors also existed in the Mojave Desert southeast of the Mojave River, but their spatial extent never matched their northern counterparts (Fig. 1; reviewed in Reheis et al., 2014).

For desert animals associated with water or associated riparian habitats, it is expected that historical bouts of isolation and connectedness, modern-day extreme isolation, and the potential for selective differences within and across basins will leave an evolutionary footprint (Hubbs and Miller, 1948; Echelle, 2008). The region is a natural continental lab for the study of potentially island-like evolutionary dynamics. Moreover, the extensive backdrop of climatic and geological evidence regarding timing and geographic context of hydrologic connectedness makes this region particularly attractive (e.g., Reheis et al., 2002a, b; Smith et al., 2002; Knott et al., 2008; Van Dam and Matzke, 2016). Many studies have corroborated hypotheses that populations from a single paleodrainage basin should share evolutionary history, and show divergence from populations found in separate basins (e.g., Chen et al., 2007; Echelle, 2008). Studies have also found evidence for divergence across isolated populations within a single paleodrainage (e.g., Williams and Wilde, 1981; Hershler et al., 2013; Houston et al., 2015; Campbell and Piller, 2017), and evidence for "sterile basins" (isolated both historically and currently) harboring populations with uniquely high levels of evolutionary divergence (Myers, 1942). Finally, some evolutionary patterns have provided evidence for historical connections that lack strong or congruent geologic or climatic evidence (Echelle, 2008; Crews and Gillespie, 2014).

Aquatic animals (e.g., fishes, springsnails, amphipods, and aquatic insects) are best represented in the studies referenced above, with only a handful of studies on terrestrial taxa more indirectly reliant upon water sources and connections (e.g., voles in riparian vegetation, Conroy et al., 2016; Krohn et al., 2018; *Saltonia* spiders on salt flats, Crews and Gillespie, 2014). Here we explore evolutionary divergence associated with paleodrainage connections in a terrestrial system with several unique and notable features. The jumping spider species *Habronattus tarsalis* is widespread in the western United States, and is particularly "widespread" in regional deserts, with a geographic distribution that extends from the lower Colorado desert northwards to the northern Great Basin (Fig. 2). In these desert habitats *H. tarsalis* is most-often associated with desert saltgrass (*Distichlis spicata*) found along intermittent riparian corridors or associated with desert springs (Foldi, 2006; personal observations), in an otherwise inhospitable desert

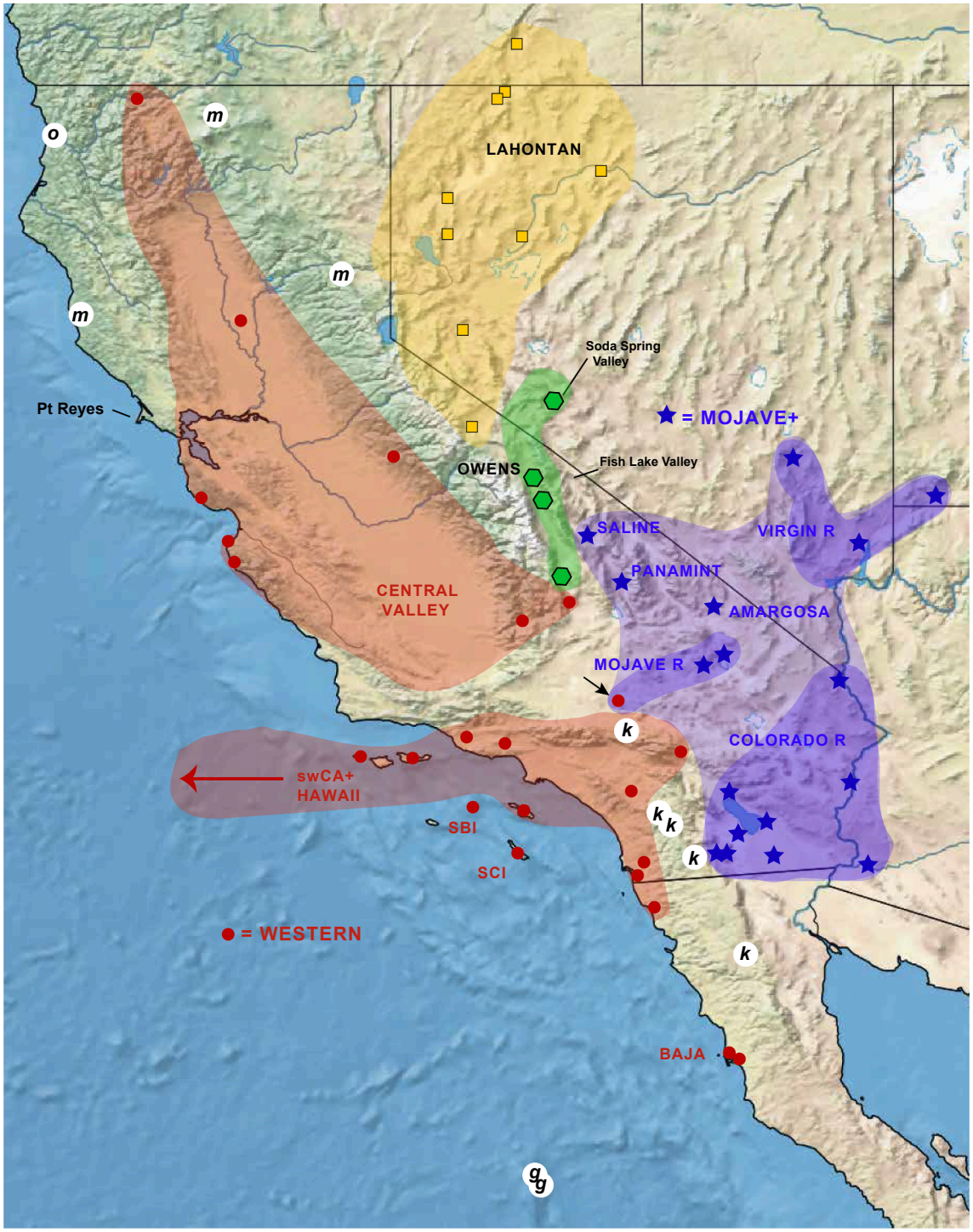


Fig. 2. Map of sampled specimens from the *H. tarsalis* complex. *Habronattus tarsalis* geographic lineages are named and colored reflecting RADSeq results (see Fig. 5). Other described species indicated by letters (**o** = *H. ophrys*, **m** = *H. mustaciata*, **k** = *H. kawini*, **g** = *H. gigas*). Locations for specific *H. tarsalis* populations discussed in text (Pt. Reyes, Soda Spring Valley, and Fish Lake Valley) are highlighted.

landscape. The species has never been collected from modern desert pinyon-juniper woodlands (that were more extensive during glacial periods of the Pleistocene), suggesting that at least in regional deserts, the species is strictly riparian-associated.

In addition to having a larger geographic distribution than previously studied terrestrial taxa, *H. tarsalis* is remarkably geographically variable in male morphology across this range (Fig. 3; Foldi, 2006). In this sense, *H. tarsalis* represents a potential microcosm of a more general evolutionary pattern in the hyperdiverse genus *Habronattus*. This genus includes over 100 species (Griswold, 1987; Maddison and Hedin, 2003; Leduc-Robert and Maddison, 2018), despite an estimated crown age of less than five million years (Bodner and Maddison, 2012). Many *Habronattus* species show extensive geographic variation in male morphology, likely explained by strong sexual selection on male courtship morphology and behavior, and perhaps also fueled by introgressive hybridization (Maddison and McMahon, 2000; Blackburn and Maddison, 2014). A well-studied example is *H. pugillis*; this species is distributed as a geographic patchwork of adjacent montane "sky-island" populations which exhibit conspicuous differences in male display morphology, with evidence for strong sexual selection overwhelming

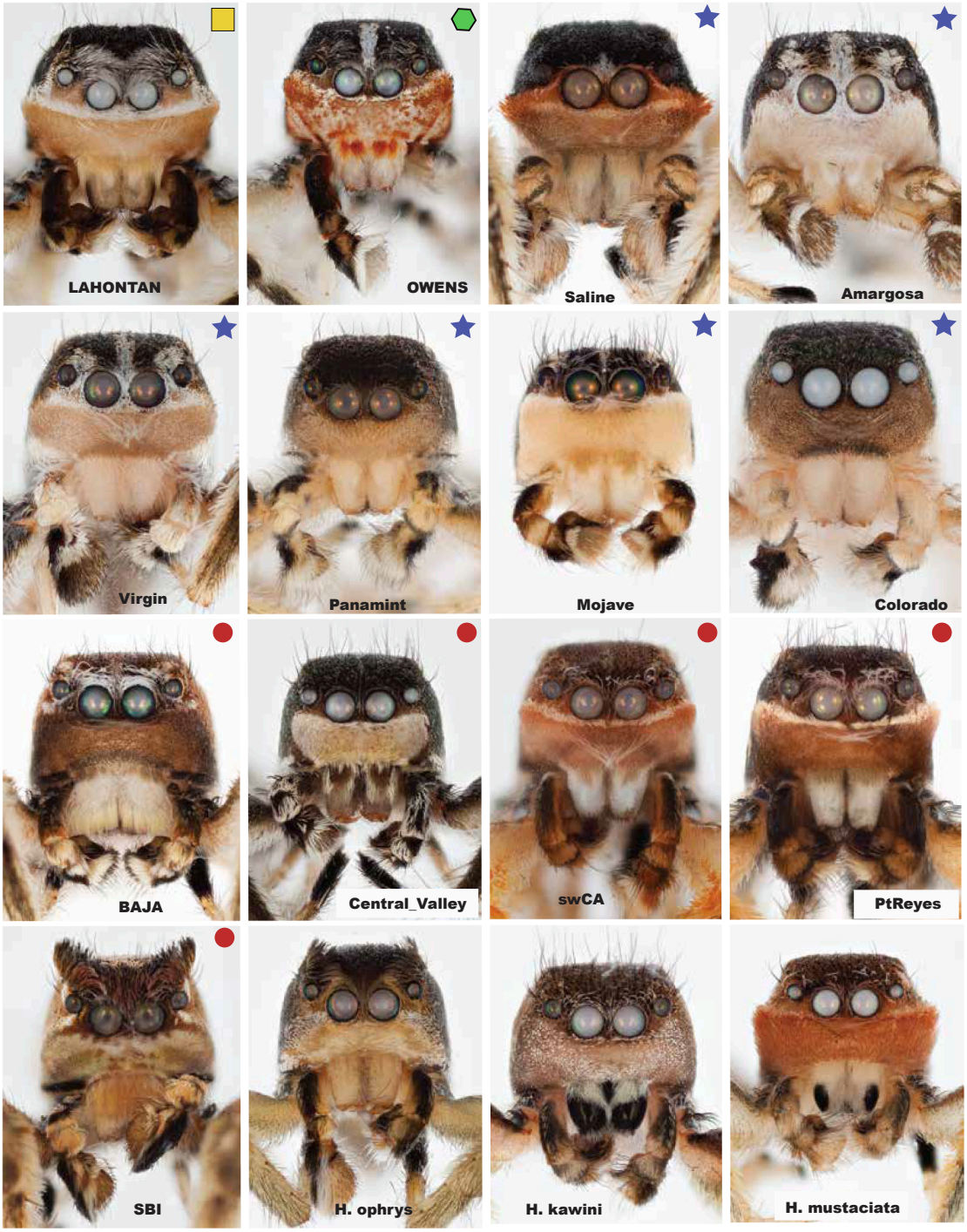


Fig. 3. Images of representative male specimens for primary *H. tarsalis* lineages and described species from the *H. tarsalis* complex. Specimens as follows (left to right, top to bottom): LAHONTAN (HA1606, Alvord Hot Springs), OWENS (HA1594, Sodaville), Saline Valley (HA1539), Amargosa (HA1525, Ash Meadows), Virgin (HA0792, Crystal Springs), Panamint Valley (G3058), Mojave (G2169, Bryman), Colorado (HA1647, Salton Sea), BAJA (HA1320, El Socorrito), Central Valley (HA1454, Lake Isabella), swCA (HA1569, Big Island), PtReyes (G3033), Santa Barbara Island (HA1621), *H. ophrys* (HA0398, Lyons), *H. kawini* (HA0764, Pine Valley), *H. mustaciata* (HA1461, San Lucas). Symbols follow genetic linages as in Fig. 2. Images not to scale. More detailed collection information for all specimens provided in Appendix B.

historical gene flow (Maddison and McMahon, 2000; Masta and Maddison, 2002). We view *H. tarsalis* as a potential inverted "sky-island" desert oasis analog, representing a powerful system to study evolutionary divergence because both morphology and genomic divergence can be used together to test Pleistocene-age evolutionary hypotheses.

Here we study *H. tarsalis* and relatives using a combination of subgenomic and morphological data for a geographically broad sample of locations. We include multiple samples from each of the Lahontan, Owens, Mojave and Colorado River systems (Fig. 2), and predict evolutionary divergence across these separated drainages and paleodrainages. Also, we include samples from currently and historically isolated basins (Saline Valley, Soda Spring Valley), and predict unique divergence in these populations. To place our *H. tarsalis* sample into a broader context, and to test the monophyly of this species, we also include genetic and morphological data for the other four described

species in the *H. tarsalis* subgroup (*H. mustaciata, H. ophrys, H. kawini,* and *H. gigas*).

2. Methods

2.1. Divergence time analyses of published transcriptome data

In an analysis including a broad sample of jumping spiders, Bodner and Maddison (2012) used a relaxed molecular clock and multiple fossils to estimate the divergence time of the most recent common ancestor (tMRCA) of *Evarcha* and *Habronattus* at 19.6 Ma, and the tMRCA of *Pellenes peninsularis* plus derived *Habronattus* at 4.8 Ma. Using BEAST v1.10.4 (Suchard et al., 2018), we used these two secondary calibrations to estimate divergence times for *H. tarsalis* subgroup taxa. We acknowledge the weaknesses in a secondary calibration approach (e.g., Schenk, 2016), but *Habronattus* currently lacks a relevant fossil record for direct fossil-based calibration.

From the primary transcriptome locus set (n = 1877) of Leduc-Robert and Maddison (2018) we filtered for loci with complete taxon coverage, also removing seven loci with one or more sequences less than 50% of the total alignment length (n = 585). We removed 13 *Habronattus* taxa from all matrices, leaving a relevant streamlined taxon sample that included *Evarcha*, *Pellenes canadensis*, and 20 *Habronattus* species including *H. ophrys* and *H. tarsalis* (from Yuma, AZ) of the *H. tarsalis* subgroup. From the 22X585 locus set we then randomly subsampled 75 loci for BEAST analyses, replicating this subsampling five times without replacement. For each locus subsample we generated XML files using unlinked site models (GTR + I), with linked clock (lognormal uncorrelated relaxed clock) and tree (speciation: calibrated Yule process) models. We used normal priors on two tMRCA nodes of interest (as above), Gamma priors for ucld.mean and yule.birthRate (shape = 0.001, scale = 1000), and left all other priors as default. BEAST analyses were run for 40 million generations, producing ESS values exceeding 200 for relevant likelihood values (Rambaut et al., 2014). Excluding the first quartile of trees as burn-in, maximum clade credibility trees were generated using TreeAnnotator (Rambaut and Drummond, 2010).

2.2. Molecular taxon sampling

Two original DNA sequence datasets were generated, including more slowly-evolving data derived from target capture of ultraconserved element (UCE) loci, and more variable RADSeq data. UCE data were collected for 10 of 11 described species in the *H. americanus* group (Griswold, 1987; Maddison and Hedin, 2003; Leduc-Robert and Maddison, 2018), including all five described species in the *H. tarsalis* subgroup and several *H. tarsalis* geographic populations (Appendix A). The RADSeq sample included more total individuals but less taxonomic breadth, focusing specifically on the *H. tarsalis* subgroup, using two *H. americanus* subgroup samples as outgroups. RADSeq *H. tarsalis* specimens were collected from localities spanning the known range, including desert, California Floristic Province, California Channel island, and Hawaiian populations (Fig. 2, Appendix A). *Habronattus tarsalis* populations are known from the White and Virgin drainages of southern Nevada, but no known records exist north of these regions (e.g., Bonneville Basin, Fig. 1). For specimens preserved in 100% EtOH at −80 °C, genomic DNA was extracted from the posterior half of the cephalothorax plus four legs (if available) using a Qiagen DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA), and quantified using a Qubit 2.0 fluorometer.

2.3. UCE data collection, matrix assembly and analysis

We used the MYbaits Arachnida 1.1 K version 1 kit (Arbor Biosciences; Faircloth, 2017) to capture UCE loci, using standard methods of library preparation and sequencing (Starrett et al., 2017; Hedin et al., 2019). Sequence reads were assembled using VELVET (Zerbino and Birney, 2008) at default settings, then processed using PHYLUCE (Faircloth, 2016). Assembled contigs were matched to probes using relatively high minimum coverage and minimum identity values (80). UCE loci were aligned with MAFFT (Katoh and Standley, 2013) and trimmed with GBLOCKS (Castresana, 2000; Talavera and Castresana, 2007) using liberal settings (– b1 0.5, –b2 0.5, –b3 10, –b4 8). Hedin et al. (2019) showed that the baits of the original arachnid UCE probeset target exons – more liberal GBLOCKS settings result in alignments that capture as much flanking intron data as possible, important for a recently diverged group such as the *H. tarsalis* complex. PHYLUCE alignments with over 50% sample occupancy were imported into Geneious 11.0.4 (Biomatters), where all individual alignments were examined – one UCE alignment with multiple divergent sequences (likely paralogs) was discarded from further analysis.

Maximum likelihood analyses were conducted using IQ-TREE software (Nguyen et al., 2015) where all loci in the concatenated UCE matrix were treated as a single partition, and a best-fit model was automatically chosen by ModelFinder. Support was assessed via 1000 ultrafast bootstrap replicates (Hoang et al., 2018), with a nearest neighbor interchange search (-bnni) to reduce the risk of overestimating branch support. An SVDQuartets analysis (Chifman and Kubatko, 2014, 2015) was conducted on the concatenated matrix using PAUP* 4.0a (Swofford, 2002), with exhaustive quartets sampling and 1000 bootstrap replicates.

2.4. RADSeq data collection and analysis

A ddRADSeq library was prepared using the Burns et al. (2017) protocol, modified from Peterson et al. (2012). The restriction enzymes *SphI* and *MluCI* were used, as this combination resulted in the highest number of shared loci for *H. tarsalis* in a pilot study of enzyme combinations (Burns et al., 2017). A library pool including 96 samples was sequenced with an Illumina NextSeq500 under the mid output 1x150bp single-end protocol at the UC Riverside IIGB Genomics Core facility. RADSeq data were demultiplexed, quality filtered, and *denovo* assembled using ipyrad v. 0.7.30 (Eaton and Overcast, 2016), with the following settings adjusted from default: max_Indels_locus = 4, trim_loci = 5 from 3′ end, clust_threshold = 0.90 (within and across).

For phylogenomic analyses we assembled RAD alignments with a minimum of 4 (min4), 10 (min10) and 20 (min20) samples (e.g., see Eaton et al., 2017; MacGuigan and Near, 2018). Maximum likelihood analyses were conducted on concatenated min4, min10 and min20 RAD matrices using IQ-TREE as above. Also, we used a phylogenetic invariants analysis (Lake, 1987; Chifman and Kubatko, 2015) to infer quartet trees and a species tree using tetrad v0.7.30, part of the ipyrad.analysis toolkit. We used a min10 SNP matrix, excluding seven samples that returned fewer than 200 loci per sample, sampling all quartets with 100 bootstrap replicates.

Focusing only on 37 desert *H. tarsalis* samples (HA1520 removed because of too few loci), we re-ran ipyrad as above, but at min18. This analysis resulted in a matrix including 696 unlinked SNPs, for which we conducted STRUCTURE 2.3.4 (Pritchard et al., 2000) runs using an admixture model with uncorrelated allele frequencies. All other priors were left as default. STRUCTURE analyses were replicated four times for individual K values ranging from 2 to 10, each run including 200,000 generations with the first 20,000 generations removed as burnin. Data were summarized using CLUMPAK (Kopelman et al., 2015), with a best-fit K chosen utilizing the prob (K) method of Pritchard et al. (2000).

Using the desert-only min18 unlinked SNPs file as above, we conducted unsupervised machine learning using Variational Autoencoders (Kingma and Welling, 2013; Derkarabetian et al., 2019). VAE was implemented utilizing the Keras python deep learning library (https://keras.io; Chollet, 2015) and the TensorFlow machine learning framework (www.tensorflow.org; Abadi et al., 2016). SNP matrices were converted to one-hot encoding with nucleotides transformed into four binary variables unique to each nucleotide, including ambiguities (Derkarabetian et al., 2019). The one-hot encoded SNP data were used to infer the distribution of latent variables, given as a normal

Table 1
Description of morphological characters and character states.

(1)BBC9	Median white stripe between AME (0 = absent; 1 = spot; 2 = stripe projects posteriorly)
(2)H1	Swirling inward hairs above AMEs (0 = reduced; 1 = slight; 2 = elaborated (e.g., Lahontan); 3 = more conspicuous erect hairs posteriorly (e.g., *mustaciata*); 4 = elaborated as crest; 5 = elaborated as divided crest)
(3)BBC5	Extensive white setae between AMEs and ALEs (0 = absent; 1 = intermediate; 2 = dense (e.g., Amargosa))
(4)BBC8	Scales below AER forming distinct lighter transverse band (0 = absent; 1 = present only under AMEs; 2 = spans entire length of AER)
(5)BBC4	Anterior facial profile (0 = typical; 1 = moderate lateral expansion; 2 = extreme lateral expansion (=mustache))
(6)H2	Mustache color (0 = typical; 1 = brightly colored (e.g., Saline))
(7)H3	Color/nature of clypeal scales (0 = uniform covering of fine dark scales, without markings (as in *gigas*); 1 = minute spatulate scales, some irridescent, pointing upwards and medially (as in *kawini*); 2 = linear, pale, upward-facing scales, without conspicuous markings; 3 = uniform, dark, slightly spatulate scales, longer overhanging dorsal cheliceral edge (as in Colorado); 4 = as state 3, but light scales; 5 = iridescent inwards and upwards scales with mottled cuticle beneath, obvious transverse band; 6 = similar to state 5, but with white band below ALEs; 7 = mottled orange and white spatulate scales (as in Sodaville); 8 = distinct oblong shape, darker medially, lighter laterally; 9 = multiple transverse light and dark bands, plus dark blotches)
(8)BBC3	Chelicerae pigmentation (0 = pale; 1 = pale with black markings on distal end; 2 = pale with black markings on distal or outer margins; 3 = black)
(9)BBC10	Cheliceral covering (0 = fine dark clypeal scales extending onto proximal base of chelicerae (as in *gigas*); 1 = long white hairs basally, sweeping inwards, then extending distally on midline (exposing shiny black chelicerae); 2 = thin, sparse basal pale hairs; 3 = thin, sparse basal pale hairs but with median and lateral extensions distally; 4 = pale and sparse throughout; 5 = concentration of dark then light hairs basally; 6 = salt & pepper basally, long lateral hairs; 7 = one long white medial hair pencil (northern *mustaciata)*; 8 = white scale covering with distal, lateral opening, exposing black cuticle (southern *mustaciata)*; 9 = scattered hair covering; 10 = basal covering with conspicuous orange scales; 11 = white whispy hairs, denser distally, exposing medial bulbous chelicerae; 12 = dark then light hairs, extending > half cheliceral length; 13 = paired bundles of long hairs, like rabbit tails; 14 = basal dark scales forming inward-facing "devils horns")
(10)H4	Shape of chelicerae (0 = linear, 1 = distal swelling; 2 = triangular; 3 = swollen proximally; 4 = bowed outwards)
(11)G135	Palpal patella with anterolateral expansion (0 = absent; 1 = present)
(12)H5	Distal median cymbium (0 = mostly light, 1 = mostly dark; 2 = mixed)
(13)H6	Extent of median cymbium white patch (0 = typical, 1 = more extensive)
(14)H7	Palpal patella silver band (0 = typical, 1 = light)
(15)H8	Palpal femur color (0 = typical; 1 = noticeably black)
(16)H9	Strength of L2 retrolateral femur fringe (0 = weak, 1 = strong)
(17)BBC14	Setae on L1 femur (0 = drab; 1 = at least some orange and/or green)
(18)*H*10	Strength of L1 retrolateral femur fringe (0 = weak, 1 = strong)
(19)*H*11	Strength of L1 prolateral femur fringe (0 = weak, 1 = median; 2 = strong)
(20)H12	Nature of spatulate scales on femur I (0 = typical, 1 = more extensive)
(21)H13	Distal leg I fringes (e.g., patella, metatarsus) (0 = weak, 1 = median; 2 = strong)
(22)H14	Extent of black pigmentation on distal tarsus I (0 = half or 3/4s; 1 = entire)
(23)H15	Black pigmentation on distal tarsus I (0 = ventral and dorsal; 1 = white scales dorsally)
(24)H16	Overall body size (0 = typical, 1 = island giant)

distribution with a mean (μ) and standard deviation (σ); the decoder was then used to map the latent distribution to a reconstruction of the one-hot encoded data, visualized in two dimensions. Analyses were replicated five times, and we used lowest loss values (discarding the first 50% of values as burnin) to choose an optimal solution.

2.5. Morphological analysis

Seventy-three male specimens from the *H. tarsalis* subgroup (Appendix B) were scored for 24 discrete morphological characters (Table 1). We included all described species in the *H. tarsalis* subgroup, and a geographically comprehensive sample of *H. tarsalis* locations, including all primary desert drainages and paleodrainages (Lahontan, Owens, Death Valley, Mojave, and Colorado River systems). Most samples were also included in RAD genetic analyses, except for morphology-only specimens from Pt. Reyes, CA (Fig. 2). We scored characters from the face and prolateral surfaces of the first and second legs, features that males present to females during visual courtship displays (e.g., Maddison and McMahon, 2000; Elias et al., 2012). Each character consisted of at least two different character states, treated as multi-state and unordered. Almost all scorings were based on specimens also used in genomic analysis and preserved in 100% EtOH (Appendix B); the high-percentage EtOH served to preserve color important for four scored characters. Specimens were imaged using a Canon 5D Mark II with a 65 mm lens at 3X magnification, with multiple images combined into a composite image using Method C (pyramid) in HELICON FOCUS. Digital images of almost all scored specimens are deposited at the Symbiota Collections of Arthropods Network (Appendix B; https://scan-bugs.org/portal/).

Using the VEGAN R package (Oksanen et al., 2015), we calculated Bray-Curtis distances from all specimens and desert-only morphological matrices. From these we conducted a two-dimensional ordination of distance matrices using non-metric multidimensional scaling (NMDS), with random starts called using the metaMDS function.

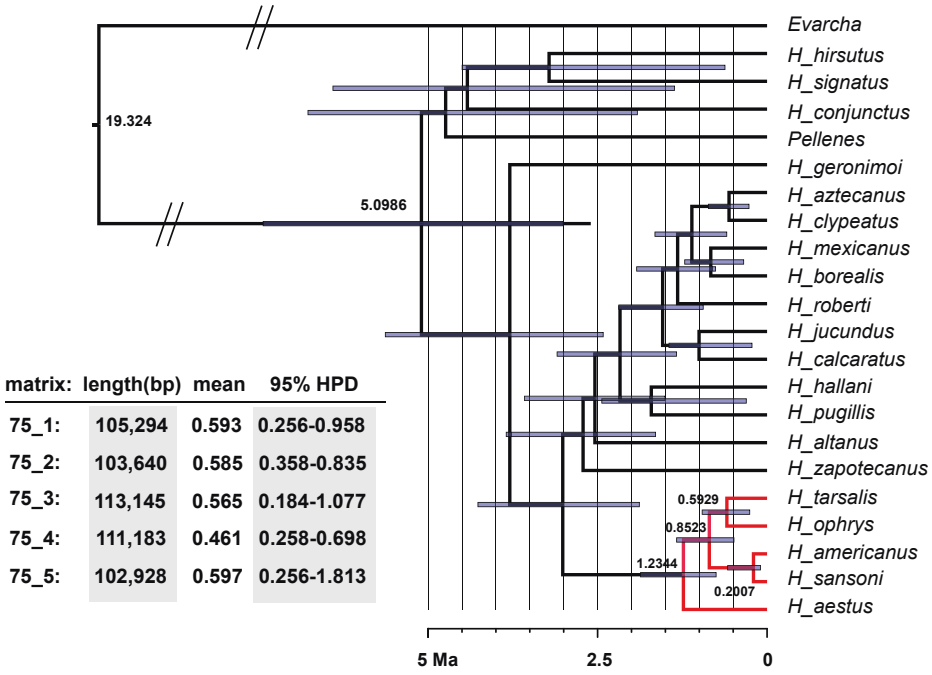
3. Results

3.1. Divergence time analyses

BEAST tree topologies from all five subsample runs are consistent with the generic-level phylogenomic results of Leduc-Robert and Maddison (2018), including the positioning of *P. canadensis* with early-diverging *Habronattus* (AAT clade), but separate from derived *Habronattus* (Fig. 4). Evidence suggests that *P. canadensis* is related to *P. peninsularis* in the subgenus *Pellenatus* (Maddison, 2017), so the secondary calibration date used here remains valid. All five subsample BEAST runs, using independent datasets each comprised of ~100,000 bp (Fig. 4 inset), indicate mean tMRCA dates for *H. ophrys* and *H. tarsalis* of approximately 0.46–0.6 Ma with 95% HPD intervals between 0.25 and 1.8 Ma (Fig. 4). We note also the results of Leduc-Robert and Maddison (2018), who suggested possible introgression between *H. ophrys* and *H. americanus* subgroup members, and between *H. tarsalis* and *H. aestus*. To the extent that this introgression is impacting the transcriptome data analyzed, this would tend to pull estimated *H. ophrys* plus *H. tarsalis* tMRCA dates towards the present (i.e., leading us to *underestimate* divergence times; Leaché et al., 2013; Wen and Nakhleh, 2018). Overall, although we acknowledge the many sources of potential error in the divergence analyses conducted, estimated dates are consistent with mid- to late-Pleistocene diversification within the *H. tarsalis* subgroup.

3.2. UCE data

A 50% occupancy matrix for 41 samples included 264 UCE loci and 1684 parsimony informative sites, with an average alignment length of approximately 420 basepairs (111,357 total). Raw reads have been submitted to the Short Read Archive (BioProject ID: PRJNA588246), and locus alignments are available at doi.org/https://doi.org/10.5061/dryad.q2bvq83fg. UCE phylogenies were rooted using *H. tuberculatus* plus *H. aestus*, as both taxa have previously been hypothesized as relatively early-



matrix:	length(bp)	mean	95% HPD
75_1:	105,294	0.593	0.256-0.958
75_2:	103,640	0.585	0.358-0.835
75_3:	113,145	0.565	0.184-1.077
75_4:	111,183	0.461	0.258-0.698
75_5:	102,928	0.597	0.256-1.813

Fig. 4. Maximum clade credibility tree resulting from BEAST analysis of 75_1 transcriptome locus set. Branch leading to *Evarcha* outgroup truncated, and without 95%_HPD value. Focal nodes with mean divergence date estimates, node bars represent 95%_HPD values. Inset: Total aligned length, *H. tarsalis* + *H. ophrys* mean tMRCA (in Ma), *H. tarsalis* + *H. ophrys* 95%_HPD (in Ma) for replicate locus sets.

diverging within the *H. americanus* group (Griswold, 1987; Leduc-Robert and Maddison, 2018; Fig. 4). IQ-TREE resolves *H. americanus* and *H. tarsalis* subgroups as reciprocally monophyletic with strong support (Supplemental Fig. S1A). Within the *H. tarsalis* subgroup four early-diverging lineages are recovered, including *H. ophrys*, *H. mustaciata*, and "LAHONTAN" and "OWENS" populations of *H. tarsalis* (Supplemental Fig. S1A). All other geographic populations of *H. tarsalis*, plus *H. gigas* and *H. kawini*, occur in a more derived but poorly-supported clade. Within this derived clade only three groups are strongly supported (we note that ultrafast bootstrap values should be interpreted as one would posterior probabilities (Hoang et al., 2018)), including *H. gigas*, three *H. tarsalis* populations from a MOJAVE+ lineage, and a Central Valley lineage. A WESTERN group, as found in RADSeq results (Fig. 5), is not resolved with the UCE data. SVDQuartets results mirror those from IQ-TREE in the separation of early-diverging lineages from all others (Supplemental Fig. S1B).

3.3. RADSeq data

After removing 13 samples that returned low raw read counts, the remaining 83 samples included between 126 and 944,000 ipyrad clusters (at mindepth_statistical = 6, clust_threshold = 0.90). min4, min10 and min20 matrices included 27,388 (3,648,477 bp), 6932 (67,388 bp) and 1665 loci (220764 bp), respectively. Raw reads have been submitted to the Short Read Archive (BioProject ID: PRJNA588237), and data matrices are available at doi.org/https://doi.org/10.5061/dryad.q2bvq83fg.

IQ-Tree results from different min_sample matrices are consistent in the recovery of over ten primary *H. tarsalis* geographic lineages, almost all strongly supported (bp > 95) in at least two of three analyses (Fig. 5). Primary lineage interrelationships are also generally consistent across different min_sample analyses, at least among *H. tarsalis* lineages and other described species. Relationships *within* lineages vary in some instances (Fig. 5, Supplemental Figs. S2 and S3), and the min20 matrix recovers different early-diverging relationships (Fig. 5). Tetrad results are congruent in overall lineage structuring and relationships, although support values are generally lower throughout (Supplemental Fig. S4).

Below we focus on congruent phylogenetic patterns, and emphasize where concatenated versus coalescent analyses differ. Four early-diverging lineages include *H. mustaciata*, *H. ophrys* (not included in tetrad analysis), LAHONTAN *H. tarsalis*, and OWENS *H. tarsalis*. Although these overall patterns are congruent with UCE results, one note of caution is that our RADSeq outgroup sample is very limited, not allowing for formal tests for introgression between *H. tarsalis* and *H. americanus* subgroup members, as hypothesized by Leduc-Robert and Maddison (2018). It is possible that the phylogenetic placement of early-diverging *H. tarsalis* subgroup RADSeq lineages reflects introgression of divergent alleles from the *H. americanus* subgroup (e.g., Eaton et al., 2015), and future studies should be designed to test this hypothesis.

Remaining RADSeq lineages are separated into two well-supported geographic clades, with multiple desert populations from the greater Mojave desert region (MOJAVE+ lineage), and WESTERN populations from the California Floristic Province of California and northern Baja California, including samples from the Channel and Hawaiian Islands (Figs. 2, 5, S4). Samples of *H. kawini* are nested within the WESTERN lineage, and *H. kawini* is not recovered as monophyletic. Geographically, the MOJAVE+ lineage is mostly distributed east of the high "central spine" of the region (including the Peninsular, Transverse, and Sierra Nevada mountain ranges), while WESTERN lineages are mostly found west of this central spine (Fig. 2).

Focusing on desert *H. tarsalis* populations, there is a strong correspondence between genetic RAD lineages and geography / paleodrainages. Members of the LAHONTAN lineage, including Mono Lake specimens, are distributed in the Lahontan basin (Figs. 1, 2, 5). Members of the OWENS lineage, including a population from Sodaville (Soda Spring Valley), are found in the Owens River drainage (Figs. 1, 2, 5). Finally, members of the MOJAVE+ lineage occur in six geographic sub-lineages, each corresponding either to a now-isolated deep valley (Saline Valley, Panamint Valley), or to a drainage basin (Amargosa River, Virgin River (*including* White River), Mojave River, Colorado River; Figs. 1, 2, 5). Specimens from the latter three drainages are intermixed on the tetrad consensus tree (Fig. S4). One upper Mojave River sample is apparently genetically misplaced, a result likely explained by introgression from western populations (see Discussion).

Utilizing the prob (K) method of Pritchard et al. (2000), STRUCTURE results for desert-only samples suggest five genetic clusters (K = 5) generally following phylogenetic patterns, particularly those from quartets analysis. These clusters include LAHONTAN, OWENS, Saline Valley, Amargosa River, and remaining samples of the MOJAVE+ lineage (Fig. 6A). One individual from Mono Lake (LAHONTAN lineage) shows possible evidence for admixture with the OWENS lineage, discussed below in the context of the complex drainage history of this area. VAE machine

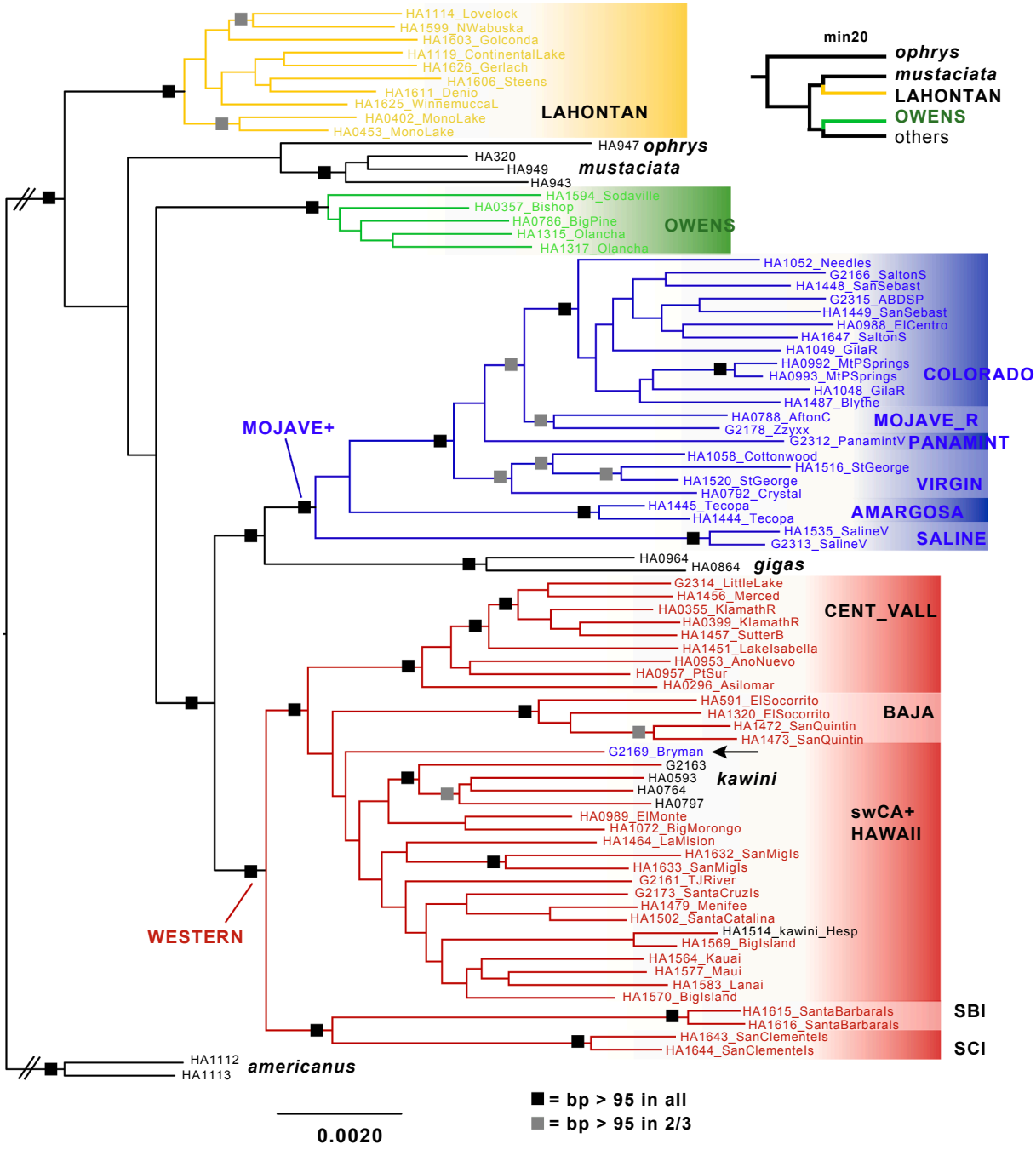


Fig. 5. RADSeq maximum likelihood IQ-TREE tree, from min4 concatenated matrix. Ultrafast bootstrap values less than 95 not shown. Upper right inset - alternative topology for early-diverging lineages from min20 matrix. Lineage names as discussed in text. Arrows point to discordant sample from upper Mojave River (Bryman).

learning results are similar to STRUCTURE results, with the possibly admixed Mono Lake sample intermediate between LAHONTAN and OWENS clusters (Fig. 6B). Saline Valley and Amargosa samples form distinct VAE clusters, and other MOJAVE+ specimens form smaller clusters that coincide with drainage basins or isolated valleys. One Virgin River sample (HA_0792) is distinct in VAE space, and shows possible evidence for genetic admixture in STRUCTURE analyses (Fig. 6A).

3.4. Morphology

NMDS analyses based on 73 scored specimens reveal a strong congruence between morphological clusters and genetic lineages (Fig. 7), whether the latter represent previously described species or *H. tarsalis* RAD lineages. In analyses including specimens from the entire complex, this congruence holds true for all but the OWENS lineage, which includes the morphologically divergent Sodaville population (Figs. 3, 7A). Analyses of desert-only samples result in clusters similar to RAD phylogenetic clades (Fig. 5) and RAD genetic clusters (Fig. 6). Distinct morphological clusters correspond to LAHONTAN, OWENS, Saline, Amargosa, Virgin, and Panamint genetic lineages. Specimens from the lower Colorado and the Mojave are morphologically similar (Fig. 7B), mirroring their genetic similarity (Figs. 5 and 6). This kinship suggests a possible historical connection between the Mojave and Colorado River basins via Bristol, Cadiz, and Danby Basins (Fig. 1). Bryman specimens from the upper Mojave are morphologically like Mojave and Colorado specimens (Figs. 3, 7B), but are genetically placed with *H. kawini* and swCA *H. tarsalis* (Figs. 5, S4), again consistent with introgression as discussed below.

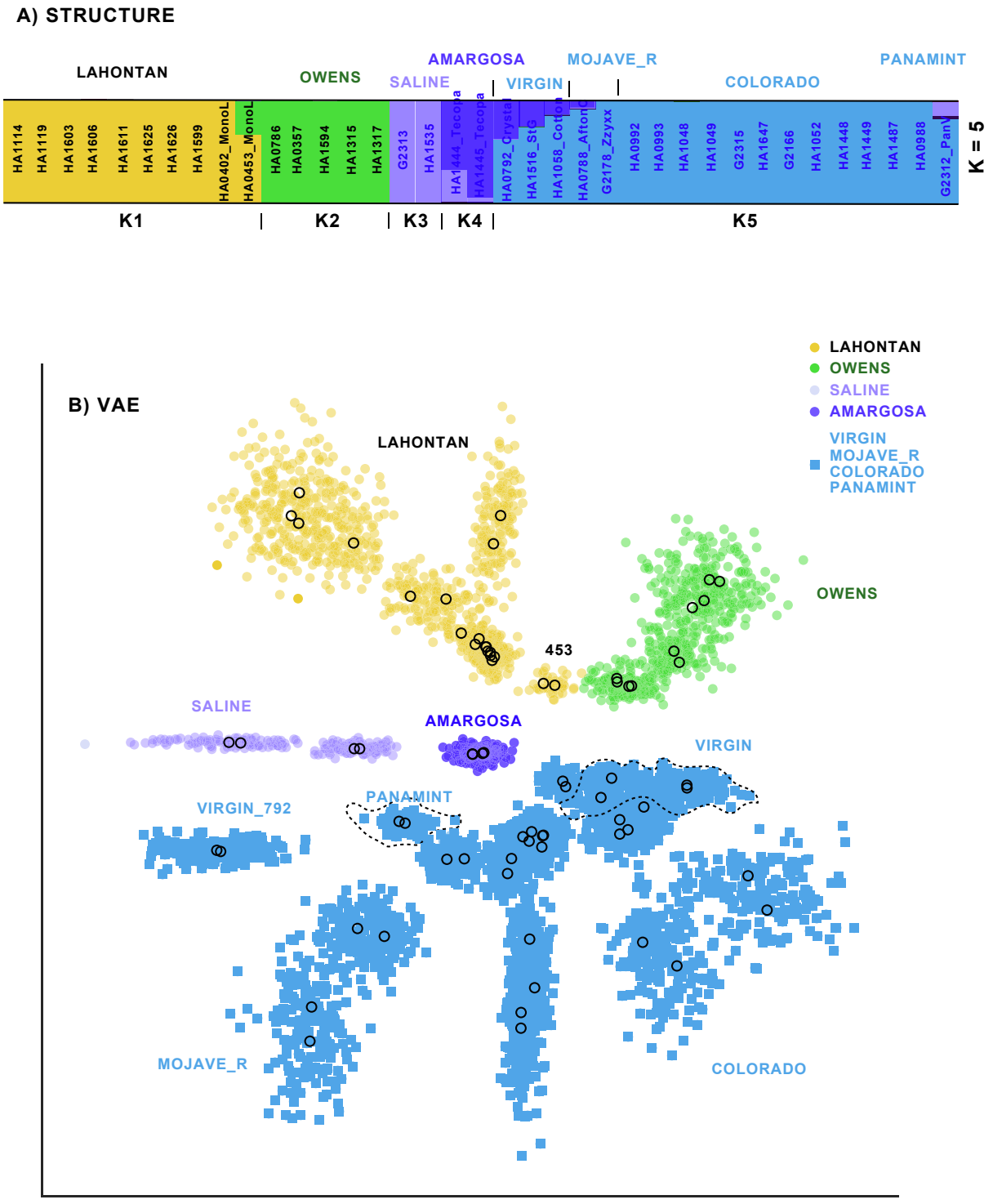


Fig. 6. Results of (A) STRUCTURE and (B) VAE machine learning analyses based on RADSeq data. VAE results with encoded mean (open circles) and standard deviation (closed circles) for each sample.

4. Discussion

4.1. Biogeographic patterns

Aquatic habitats of western North American deserts that are profoundly isolated today were intermittently connected during cyclic high-water times of the late Pliocene and Pleistocene. In this sense, modern isolation veils a deeper connectedness, and patterns of expected population kinship and connectivity can be predicted from paleoclimatic and geologic data. Of course, these latter data are themselves imperfect, and strong biological data can be used to help test and resolve contentious geologic hypotheses.

The study of biogeographic patterns and correlations with pluvial pasts has a rich research history in the region, particularly for aquatic taxa. But no single biological system is expected to provide universal biogeographic insight, and studies of riparian-associated terrestrial taxa potentially add special value, but are currently scarce. Moreover, almost all previous molecular-based tests of hydrologic hypotheses and predictions, including data on both timing and phylogenetic patterns, are based only on mitochondrial evidence. One exception, Amargosa

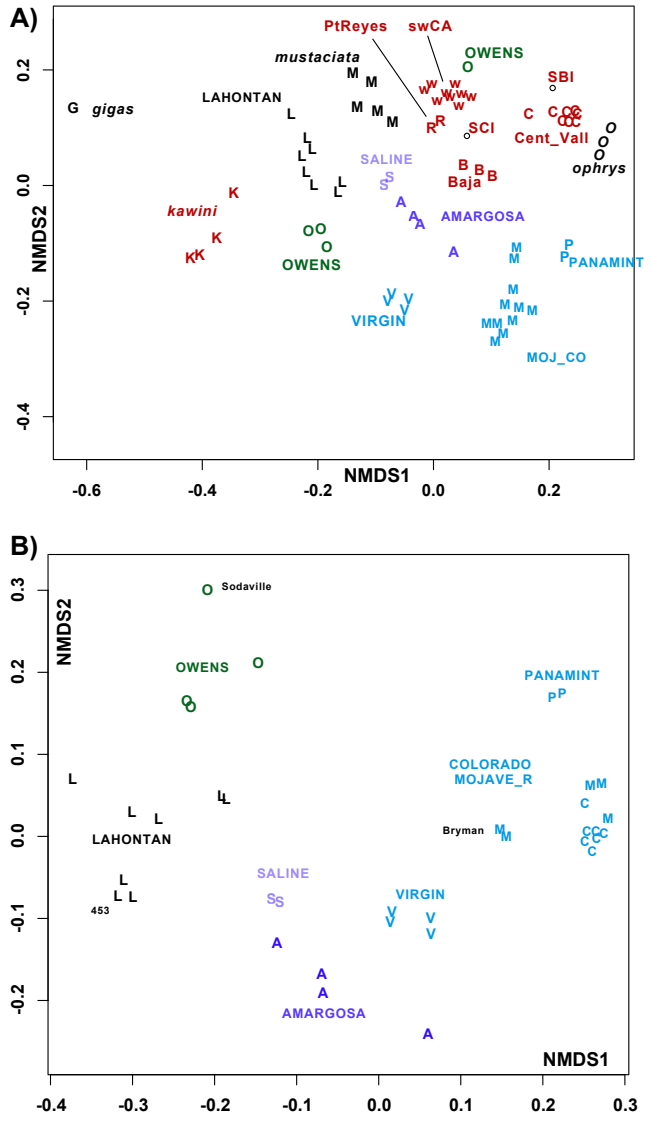


Fig. 7. Results of NMDS analysis of morphological data for (A) all described species and *H. tarsalis* RADSeq lineages, and (B) desert-dwelling *H. tarsalis* RADSeq lineages.

voles, reveals very different nuclear versus mitochondrial patterns (Conroy et al., 2016; Krohn et al., 2018), cautioning against over-interpretation of existing mitochondrial-only datasets which have often failed to find strong congruence between biological patterns and past hydrological connections/predictions (e.g., Van Dam and Matzke, 2016). Here we have uncovered a species complex that adds to our regional knowledge of desert biogeography in unique ways, because of the multiple types of evidence utilized (including nDNA and morphology) and because the taxon is terrestrial and relatively broadly distributed in the region.

Habronattus tarsalis includes a LAHONTAN lineage, supported by both genetic and morphological data, with a geographic distribution that corresponds closely to the LAHONTAN paleobasin (Figs. 1 and 2). Although the fauna of the region is clearly understudied, endemism in the LAHONTAN basin is apparently uncommon in terrestrial taxa (but see Wilson and Pitts, 2012), as most phylogeographic studies that we are aware of indicate recent range expansion from southern to northern latitudes (e.g., Jezkova et al., 2015). The LAHONTAN basin includes modern east-flowing rivers that drain from the eastern Sierra Nevada, and west-flowing rivers draining from eastern and northeastern Nevada. RADSeq data reveal sub-lineages that follow these drainages, for example sampled sites along the discontinuous modern Humboldt River (Figs. 2 and 5), but our current sample is sparse.

Samples of *H. tarsalis* from Mono Lake are phylogenetic members of the LAHONTAN lineage, despite the fact that modern Mono Lake is connected (discontinuously) southwards, to the Owens River basin (Fig. 1). However, the pluvial Lake Russell (=modern Mono Lake) had drainage connections northwards to the LAHONTAN basin before 1.3 Ma and possibly as late as 0.76 Ma (Reheis et al., 2002a, 2014). The observation that one Mono Lake *H. tarsalis* specimen is potentially genetically admixed with OWENS lineages (Fig. 6) may indicate evidence for secondary contact between original (northern) and more recently arrived (southern) lineages, as has also been suggested for tui chub fishes (Chen et al., 2007). Large genetic samples from modern Mono Lake and surroundings might be useful in further testing these drainage switching hypotheses.

Multiple times during the Pleistocene, drainages fed a massive Lake Lahontan with a lake level near 1400 m (Reheis, 1999; Reheis et al., 2002b), connecting many basins that are highly disconnected today. One geologically tenuous connection was from the southern Lahontan (southeast of Walker Lake) to basins including Lake Columbus-Rennie (Reheis et al., 2002b), including the modern Soda Spring Valley (Fig. 2). We did not find evidence for this connection in the RAD data, which instead place Sodaville specimens in the OWENS lineage (Figs. 5, 6, S4), although the Sodaville population is morphologically distinct (Figs. 2 and 7). The genetic data thus indicate a phylogeographic connection southwards, perhaps to also include Fish Lake and Deep Springs valleys. Both snail (Liu and Hershler, 2007; Hershler and Liu, 2008a) and vole (Conroy et al., 2016) samples from the latter are allied with Owens River populations, and we predict the same affinity for the Deep Springs valley population of *H. tarsalis*, which has never been sampled but surely exists.

The Saline Valley is one of the most isolated valleys in western North America, with current low passes into the valley at high elevations (~1500 m). This valley is thought to have formed via rifting 3.0–1.8 Ma (Burchfiel et al., 1987; Oswald and Wesnousky, 2002), and there is no evidence for direct aquatic connections with either the adjacent Death or Panamint Valleys during the Pleistocene. Few genetic studies of riparian-associated taxa have included samples from this valley. Amargosa voles from the Saline Valley are now extinct, and were not included in recent genetic studies (Conroy et al., 2016; Krohn et al., 2018). In springsnails, Saline Valley specimens are kin to distant southern populations, likely a result of bird dispersal (Liu et al., 2003; Hershler and Liu, 2008b). In *H. tarsalis*, specimens from Saline Valley are highly distinctive both genetically and morphologically (Figs. 3, 5–7), perhaps indicative of species status (further discussed below); this special population with an exceedingly small geographic distribution deserves further study and conservation attention.

The Owens River drainage is part of a larger Death Valley paleodrainage system (Fig. 1), with many studies indicating Pleistocene Lake Manly as an endpoint for paleorivers that flowed from the northwest (Owens), northeast (Amargosa), and the south (Mojave), although not necessarily during the same time interval. The late-Pleistocene Owens in particular is known to have flowed south then eastwards, sequentially connecting China Lake, Lake Searles, Panamint Valley then Death Valley (Jayko et al., 2008; Knott et al., 2008; Fig. 1). From this well-known history, a genetic prediction would be a Panamint < > Owens kinship, perhaps as part of a larger Death Valley lineage. In *H. tarsalis* we instead found an OWENS lineage, supported by both genetic and morphological data, separate from the remaining Death Valley system, including Panamint Valley. Some researchers have found an Owens lineage separate from greater Death Valley lineages (e.g., including Amargosa) in aquatic taxa (e.g., Echelle, 2008; Hershler and Liu, 2008a), while others have found a late-Pleistocene Owens to Amargosa connection (Sağlam et al., 2016).

Of the remaining Death Valley system lineages (not including Owens), the *H. tarsalis* Amargosa genetic lineage appears distinct. This

special divergence is mirrored in many other taxa from this now highly-discontinuous drainage, including *Saltonia* spiders (Crews and Gillespie, 2014), Amargosa voles (Krohn et al., 2018), the iconic Devil's hole pupfish (Echelle, 2008; Sağlam et al., 2016), springsnails (Hershler and Liu, 2008b; Hershler et al., 2013), and others. This divergence may reflect tenuous connections with other paleodrainages during the early to mid-Pleistocene (e.g., see Fig. 18, Knott et al., 2008). Studies of other taxa have also revealed divergence across populations *within* this discontinuous drainage (e.g., Liu et al., 2003; Hershler and Liu, 2008a). Although not tested here with genetic data, the *H. tarsalis* morphological data do suggest internal divergence (Fig. 7B), and detailed genetic sampling along the Amargosa would be insightful. Paleo-reconstructions of the upper Amargosa generally do not indicate extensive pluvial lakes, perhaps indicating more persistent levels of habitat fragmentation over deep time.

Although not a focus of this paper, biogeographic patterns for populations west of the high elevation spine of California are similarly informative. Our data indicate that the California Channel Islands, including Isla Guadalupe from central Baja, have been colonized four separate times by members of the *H. tarsalis* complex (Figs. 2, 5, S4). Southern islands, including Santa Barbara, San Clemente, and Isla Guadalupe each house endemic populations that are uniquely divergent both genetically and morphologically (Figs. 3, 5–7). Conversely, the northern Channel Islands, the disjunct Pt. Reyes population, and the Hawaiian Islands all show limited divergence from southern California mainland populations, as suggested by RADSeq data and/or morphology (Figs. 3, 5, 7). Natural dispersal to the northern Channel Islands is likely, but as argued by other authors (Prószyński, 2002), recent human transport to the Hawaiian Islands is most parsimonious.

4.2. Divergence in Habronattus

Habronattus includes a large number of species having evolved in a relatively short period of time, illustrating the potential for rapid diversification at the level of species. But this rapid evolution applies also to what is currently considered as geographic variation within species, including morphologically divergent sky island populations of *H. pugillis* (Maddison and McMahon, 2000; Masta and Maddison, 2002), divergent montane morphological forms of *H. americanus* (Blackburn and Maddison, 2014), and distinct ecological forms of *H. ustulatus* (Hedin and Lowder, 2009). Other examples of putative geographic variation were discussed by Griswold (1987), but have yet to be formally studied. For both the inter- and intraspecific divergences discussed above, a combination of geographic isolation, strong sexual selection on male morphology and behavior, and occasional gene flow across population and species boundaries is hypothesized to fuel rapid diversification (Maddison and McMahon, 2000; Hedin and Lowder, 2009; Leduc-Robert and Maddison, 2018).

Here we have revealed a similar example of rapid evolution of both novel morphological forms and diverging genetic lineages. We emphasize the term "diverging" here, acknowledging that most of our presented analyses presume a branching (phylogenetic) pattern of divergence, but that at the same time, the lineages discovered (including described species) could be connected by varying degrees of present and historical gene flow. While we do not view the entire complex as phylogeographic, with divergence explained by isolation by distance, many parts of the complex almost certainly agree with such a dynamic (e.g., MOJAVE+ lineage, not including Saline Valley and Amargosa River, etc.). But to the extent that the described species are unique and have evolved under a mostly divergent history, then phylogenetic analyses and resultant interpretations are justified. Also, the finding of similar patterns of north < > south divergence on either side of a central montane spine is consistent with divergence and not isolation by distance. As noted below in the Conclusions, and several other places in this manuscript, what is ultimately needed here is more fine-grained geographic sampling, including more individuals per collection location.

In the context of estimated mid- to late-Pleistocene divergence times (Fig. 4), the extent of phenotypic diversity discovered is noteworthy. Our data indeed suggest that male courtship morphology is sometimes evolving as fast or faster than rapidly-evolving RAD data, for example, the finding of different morphological clusters (Virgin, Panamint, Mojave & Colorado; Fig. 7B) within a single inferred STRUCTURE cluster (Fig. 6A). The *H. tarsalis* subgroup also presents a situation where different nuclear genetic lineages (including some previously described species) correspond rather closely with morphological units. This congruence contrasts slightly with prior studies of *Habronattus* geographic variation. In Arizona sky island *H. pugillis,* mitochondrial clades are not strictly congruent with range-specific morphological groups, explained by incomplete lineage sorting of mitochondrial haplotypes (Masta, 2000). Conversely, single nuclear genetic clusters within montane *H. americanus* contain multiple divergent morphological forms (Blackburn and Maddison, 2014), suggesting active divergence with gene flow in this recently-diverged group (Fig. 4).

The montane spine of California acts as strong biogeographic barrier, with essentially all lineages confined to either side of this barrier. One exception includes the Bryman population along the Mojave River, downstream of eastern populations of *H. kawini* along the upper Mojave (Fig. 2). These eastern *H. kawini* are the only known populations of this species east of the montane crest. Introgression appears to be occurring along the upper Mojave where these divergent lineages meet, as also found in fishes (Hubbs and Miller, 1943). Bryman specimens display a Mojave morphology (Figs. 3, 7B) but are phylogenetically nested within a RAD lineage that includes *H. kawini* (Figs. 5, S4). Blackburn and Maddison (2014) showed recent nuclear gene flow across phenotypically divergent populations of *H. americanus*. These authors argued that this phenotypic divergence is likely maintained by sexual selection on male morphologies presented to females during courtship, and tentatively suggested genome-wide genetic divergence underlying this phenotypic divergence. In the case of the Bryman population, the mismatch between genome-wide RADSeq data versus phenotypes might instead be more consistent with smaller genomic "islands of divergence" (Feder et al., 2012), but more data are needed to test this hypothesis.

Introgression is also possibly evidenced by the phylogenetic placement of some or all *H. kawini* specimens, which are morphologically distinct from swCA *H. tarsalis* (Figs. 3, 7B), but fall within the larger swCA RAD lineage (Figs. 5, S4). *Habronattus kawini* and *H. tarsalis* are largely parapatric in southern California and northern Baja, with *H. kawini* found at higher elevations than adjacent *H. tarsalis* populations, and we have seen specimens from intermediate elevations (Lake of the Woods, 4000 ft elevation) labeled by Griswold as "*H. tarsalis* X *H. kawini*?". Another possible area of introgression is found at the terminus of the Owens drainage, where a member of the Central Valley lineage is found east of the dividing crest (at Little Lake; Figs. 2 and 5). In all of the cases discussed above, introgression might be occurring predominantly along linear riparian corridors, allowing interaction between western versus eastern lineages, or low- versus high-elevation lineages. The geographic linearity of these interactions may facilitate and simplify more detailed studies of introgression processes.

The data presented here have possible taxonomic implications. Accepted at face value, phylogenetic results reveal four described species nested within a highly variable *H. tarsalis* (Fig. 5). Even if *H. mustaciata* and *H. ophrys* fall phylogenetically outside of *H. tarsalis* (Fig. 5 inset, S4), *H. gigas* and *H. kawini* are still nested inside. All five species in this complex were originally described based on diagnostic morphologies, although Griswold (1987) noted the challenges of the possibly "paraphyletic entity" that is *H. tarsalis*. Our more comprehensive data indicates that the complex indeed includes a series of mostly allopatric lineages which are both genetically and morphologically distinct to varying degrees. Under many (most?) species concepts, these lineages would constitute species, but the challenge of neatly delimiting species in *Habronattus* is undeniable. Novel geographic forms, traditionally treated as intraspecific variants, appear to evolve

rapidly (e.g., Blackburn and Maddison, 2014); but in the same system, gene flow across species boundaries is apparent even across clearly distant and morphologically distinct species (Maddison and Hedin, 2003; Leduc-Robert and Maddison, 2018).

The *H. tarsalis* complex, and *Habronattus* more generally, may well match properties of the ephemeral speciation model (Rosenblum et al., 2012). Under this model, speciation is common and rapid, but most new species do not persist over deeper evolutionary time. The lack of species persistence might relate to extinction of small populations, or to reticulation and reabsorption from other species, both amplified in geographic regions where environmental conditions have drastically changed through time. The hierarchical nature of variation in the *H. tarsalis* complex (nested morphological and genetic lineages), the extremely small range sizes of many forms (e.g., Saline Valley, Soda Spring Valley, Santa Barbara Island, etc.), the high likelihood that geographic distributions have shrunk and expanded through time, and signs of introgression all align with an ephemeral speciation model. Regarding desert populations in particular, Smith et al. (2002) discussed a similar model for Great Basin fishes, arguing that isolated populations in desert basins arise frequently, but then have a higher likelihood of extinction than speciation. Finally, the model predicts an uneven distribution of geographic variants ("incipient forms") within species, with some species containing many, while others contain few (Fig. 1B, Rosenblum et al., 2012). This prediction matches empirical observations across *Habronattus*, as many widespread species in fact lack pronounced geographic variation. Overall, we propose the *H. tarsalis* complex as a system in which to understand the demographic processes of isolation and persistence, both key parameters in speciation research (Harvey et al., 2019), in a simplified, island-like continental setting. For desert *H. tarsalis* in particular, measurements of population size are highly tractable (potentially estimated from the extent of available habitat), and geologic data help bolster estimates of timing and spatial context of population isolation (and thus population persistence).

5. Conclusions

Here we have discovered a rather compelling species complex for additional studies of regional phylogeographic history and island-like speciation dynamics. Our current sample of desert populations barely scratches the surface, with capacity in all geographic areas and drainages to increase the spatial density of sampling. This is true for all primary basins discussed herein (Lahontan, Owens, Amargosa, Mojave, etc.), but might also extend to areas where *H. tarsalis* populations are currently unknown, but possibly exist (e.g., Bonneville basin, etc.). We expect that increased sampling, combined with modern genetic methods, will provide enhanced geographic resolution and ability to further test specific Pleistocene-age biogeographic hypotheses.

Hubbs and Miller (1948) were captivated by the Great Basin, envisioned as an area of exceedingly rapid fish speciation since isolation after the last glacial maximum. Although this timing paradigm has been questioned (again reflecting mostly mitochondrial datasets; e.g., Smith et al., 2002), the idea of active evolutionary divergence in now isolated basins or along discontinuous modern drainages remains the same. Desert populations in the *H. tarsalis* complex present an obvious arena for evolutionary divergence because the formation of discrete populations through isolation (over space and through time) is universally present, and because rapid character evolution relevant to reproductive closure is overlaid upon this isolation. In this sense, emphasis on whether or not the isolated and diverging populations represent species versus distinct populations somewhat misses the point. Active and rapid divergence (and sometimes reticulation) is happening, the processes are there to study.

Acknowledgements

Permits for specimen collection were granted by Death Valley National Park, Ash Meadows National Wildlife Refuge, and Anza-Borrego Desert State Park. We thank James Starrett and Mercedes Burns for assistance in RADSeq data collection, Shahan Derkarabetian and Tierney Bougie for assistance in UCE experiments, and Shahan Derkarabetian for facilitating machine learning analyses. Many individuals helped to collect specimens, including Tierney Bougie, Mercedes Burns, Joe Deas, Damian Elias, Robin Keith, Steve Lew, Michael Lowder, Jose Macias, Pierre Paquin, James Starrett, Steven Thomas, and Dustin Wood. Comments of Guilherme Azevedo, Wayne Maddison, and an anonymous reviewer helped to improve the manuscript. Funding was provided by grants to SF (American Arachnological Society, Mojave Desert Research Fund), BRB (Anza-Borrego Foundation, Howie Wier Memorial Conservation Grants Program, Theodore Roosevelt Memorial Fund of the American Museum of Natural History), and MH (NSF DEB 1754591).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2019.106696.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., 2016. Tensorflow: a system for large-scale machine learning. 12th USENIX Symposium on Operating Systems Design and Implementation 16, pp. 265–283. www.tensorflow.org.

Blackburn, G.S., Maddison, W.P., 2014. Stark sexual display divergence among jumping spider populations in the face of gene flow. Mol. Ecol. 23, 5208–5223.

Bodner, M.R., Maddison, W.P., 2012. The biogeography and age of salticid spider radiations (Araneae: Salticidae). Mol. Phylo. Evol. 65, 213–240.

Burchfiel, B.C., Hodges, K.V., Royden, L.H., 1987. Geology of Panamint Valley-Saline Valley pull-apart system, California: Palinspastic evidence for low-angle geometry of a Neogene range bounding fault. J. Geophysical Res.: Solid Earth 92 (B10), 10422–10426.

Burns, M., Starrett, J., Derkarabetian, S., Richart, C.H., Cabrero, A., Hedin, M., 2017. Comparative performance of double-digest RAD sequencing across divergent arachnid lineages. Mol. Ecol. Res. 17, 418–430.

Campbell, D.C., Piller, K.R., 2017. Let's jump in: a phylogenetic study of the great basin springfishes and poolfishes, *Crenichthys* and *Empetrichthys* (Cyprinodontiformes: Goodeidae). PloS one 12 (10), e0185425.

Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17 (4), 540–552.

Chen, Y., Parmenter, S., May, B., 2007. Introgression between Lahontan and endangered Owens tui chubs, and apparent discovery of a new tui chub in the Owens Valley. California. Cons. Gen. 8 (1), 221–238.

Chifman, J., Kubatko, L., 2014. Quartet inference from SNP data under the coalescent model. Bioinformatics 30 (23), 3317–3324.

Chifman, J., Kubatko, L., 2015. Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes, site-specific rate variation, and invariable sites. J. Theoret. Biol. 374, 35–47.

Chollet, F., 2015. Keras. https://keras.io.

Conroy, C.J., Patton, J.L., Lim, M.C., Phuong, M.A., Parmenter, B., Höhna, S., 2016. Following the rivers: historical reconstruction of California voles *Microtus californicus* (Rodentia: Cricetidae) in the deserts of eastern California. Biol. J. Linn. Soc. 119, 80–98.

Crews, S.C., Gillespie, R.G., 2014. Desert salt flats as oases for the spider *Saltonia incerta* Banks (Araneae: Dictynidae). Ecol. Evol. 4, 3861–3874.

Derkarabetian, S., Castillo, S., Koo, P.K., Ovchinnikov, S., Hedin, M., 2019. A demonstration of unsupervised machine learning in species delimitation. Mol. Phylo. Evol. 139, 106562.

Eaton, D.A.R., Overcast, I., 2016. ipyrad: interactive assembly and analysis of RADseq data sets. http://ipyrad.readthedocs.io.

Eaton, D.A., Hipp, A.L., González-Rodríguez, A., Cavender-Bares, J., 2015. Historical introgression among the American live oaks and the comparative nature of tests for introgression. Evolution 69 (10), 2587–2601.

Eaton, D.A., Spriggs, E.L., Park, B., Donoghue, M.J., 2017. Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. Syst. Biol. 66 (3), 399–412.

Echelle, A.A., 2008. The western North American pupfish clade (Cyprinodontidae: *Cyprinodon*): mitochondrial DNA divergence and drainage history. Spec. Pap. Geol. Soc. Am. 439, 27.

Elias, D.O., Maddison, W.P., Peckmezian, C., Girard, M.B., Mason, A.C., 2012. Orchestrating the score: complex multimodal courtship in the *Habronattus coecatus* group of *Habronattus* jumping spiders (Araneae: Salticidae). Biol. J. Linnean Soc. 105 (3), 522–547.

Faircloth, B.C., 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32, 786–788.

Faircloth, B.C., 2017. Identifying conserved genomic elements and designing universal bait sets to enrich them. Meth. Ecol. Evol. 8, 1103–1112.

Feder, J.L., Egan, S.P., Nosil, P., 2012. The genomics of speciation-with-gene-flow. Trends

Genet. 28 (7), 342–350.
Foldi, S.E., 2006. Morphology-based species delimitation and phylogeny reconstruction within the Habronattus tarsalis species complex of california and surrounding states. San Diego State Univ. Master's Thesis.
Griswold, C.E., 1987. A revision of the jumping spider genus *Habronattus* F.O.P. Cambridge (Araneae; Salticidae), with phenetic and cladistic analyses. Univ. California Pub. Ent. 107, 1–344.
Harvey, M.G., Singhal, S., Rabosky, D.L., 2019. Beyond reproductive isolation: demographic controls on the speciation process. Ann. Rev. Ecol. Evol. Syst. 50.
Hedin, M., Lowder, M.C., 2009. Phylogeography of the *Habronattus amicus* species complex (Araneae: Salticidae) of western North America, with evidence for localized asymmetrical mitochondrial introgression. Zootaxa 2307 (1), 39–60.
Hedin, M., Derkarabetian, S., Alfaro, A., Ramírez, M.J., Bond, J.E., 2019. Phylogenomic analysis and revised classification of atypoid mygalomorph spiders (Araneae, Mygalomorphae), with notes on arachnid ultraconserved element loci. PeerJ 7, e6864.
Hershler R., Liu H.-P., 2008a. Ancient vicariance and recent dispersal of springsnails (Hydrobiidae: Pyrgulopsis) in the Death Valley system, California-Nevada. In: Reheis M.C., Hershler R., Miller D.M. (Eds.) Late Cenozoic drainage history of the southwestern Great Basin and lower Colorado River region. Geol. Soc. Am. Spec. Paper 439, pp. 91–101.
Hershler, R., Liu, H.-P., 2008b. Phylogenetic relationships of assimineid gastropods of the Death Valley region: relicts of a late Neogene marine incursion? J. Biogeog. 35, 1816–1825.
Hershler, R., Liu, H.-P., Bradford, C., 2013. Systematics of a widely distributed North American springsnail, *Pyrgulopsis micrococcus* (Caenogastropoda, Hydrobiidae), with descriptions of three new congeners. ZooKeys 330, 27–52.
Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S., 2018. UFBoot2: improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35, 518–522.
Houston, D.D., Evans, R.P., Shiozawa, D.K., 2015. Pluvial drainage patterns and Holocene desiccation influenced the genetic architecture of Relict Dace, *Relictus solitarius* (Teleostei: Cyprinidae). PloS One 10 (9), e0138433.
Hubbs, C.L., Miller, R.R., 1943. Mass hybridization between two genera of cyprinid fishes in the Mohave Desert California. Michigan Acad. Sci., Arts Lett.
Hubbs, C.L., Miller, R.R., 1948. The zoological evidence: correlation between fish distribution and hydrographic history in the desert basins of Western United States. Bull. Univ. Utah Biol. Series 10 (38), 17–166.
Jayko, A.S., Forester, R.M., Kaufman, D.S., Phillips, F.M., Yount, J., McGeehin, J., Mahan, S.A., Reheis, M.C., Hershler, R., Miller, D.M., 2008. Late Pleistocene lakes and wetlands, Panamint Valley, Inyo County, California. Geol. Soc. Am. Spec. Papers 439, 151.
Jezkova, T., Riddle, B.R., Card, D.C., Schield, D.R., Eckstut, M.E., Castoe, T.A., 2015. Genetic consequences of postglacial range expansion in two codistributed rodents (genus *Dipodomys*) depend on ecology and genetic locus. Mol. Ecol. 24 (1), 83–97.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30 (4), 772–780.
Kingma, D.P., Welling, M., 2013. Auto-encoding variational Bayes. In: Proceedings of the International Conference on Learning Representations (ICLR) arXiv:1312.6114v10 [stat.ML].
Liddicoat, J.C., David, B.T., Ebbs, V.M., 2008. Reconstructing late Pliocene to middle Pleistocene Death Valley lakes and river systems as a test of pupfish (Cyprinodontidae) dispersal hypotheses. Late Cenozoic drainage history of the southwestern Great Basin and lower Colorado River region: geologic and biotic perspectives 439, 1.
Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., Mayrose, I., 2015. CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across K. Mol. Ecol. Res. 15 (5), 1179–1191.
Krohn, A.R., Conroy, C.J., Pesapane, R., Bi, K., Foley, J.E., Rosenblum, E.B., 2018. Conservation genomics of desert dwelling California voles (*Microtus californicus*) and implications for management of endangered Amargosa voles (*Microtus californicus scirpensis*). Cons. Genet. 19 (2), 383–395.
Lake, J.A., 1987. A rate-independent technique for analysis of nucleic acid sequences: evolutionary parsimony. Mol. Biol. Evol. 4 (2), 167–191.
Leaché, A.D., Harris, R.B., Rannala, B., Yang, Z., 2013. The influence of gene flow on species tree estimation: a simulation study. Syst. Biol. 63 (1), 17–30.
Leduc-Robert, G., Maddison, W.P., 2018. Phylogeny with introgression in *Habronattus* jumping spiders (Aranae: Salticidae). BMC Evol. Biol. 18.
Liu, H.-P., Hershler, R., 2007. A test of the vicariance hypothesis of western North American freshwater biogeography. J. Biog. 34, 534–548.
Liu, H.P., Hershler, R., Clift, K., 2003. Mitochondrial DNA sequences reveal extensive cryptic diversity within a western American springsnail. Mol. Ecol. 12 (10), 2771–2782.
MacGuigan, D.J., Near, T.J., 2018. Phylogenomic signatures of ancient introgression in a rogue lineage of Darters (Teleostei: Percidae). Syst. Biol. 68 (2), 329–346.
Maddison, W.P., 2017. New species of *Habronattus* and *Pellenes* jumping spiders (Araneae, Salticidae, Harmochirina). ZooKeys 646, 45.
Maddison, W.P., Hedin, M., 2003. Phylogeny of *Habronattus* jumping spiders (Araneae:Salticidae), with consideration of genital and courtship evolution. Syst. Ent. 28, 1–21.
Maddison, W., McMahon, M., 2000. Divergence and reticulation among montane populations of a jumping spider (*Habronattus pugillis* Griswold). Syst. Biol. 49, 400–421.
Masta, S.E., 2000. Phylogeography of the jumping spider *Habronattus pugillis* (Araneae: Salticidae): recent vicariance of sky island populations? Evol. 54 (5), 1699–1711.
Masta, S.E., Maddison, W.P., 2002. Sexual selection driving diversification in jumping spiders. Proc. Nat. Acad. Sci. 99 (7), 4442–4447.
Myers, G.S., 1942. The black toad of Deep Springs valley, Inyo County, California. Occas. Papers Mus. Zool. Univ. Michigan 460, 3–13.
Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. Mol. Biol. Evol. 32, 268–274.
Oksanen, J., Blanchet, F.G., Kindt, R., Legendere, P., Minchin, P.R., O'Hara, R.B., 2015. Vegan: Community Ecology Package. R Package Vegan, Version 2.5-5.
Oswald, J.A., Wesnousky, S.G., 2002. Neotectonics and Quaternary geology of the Hunter Mountain fault zone and Saline Valley region, southeastern California. Geomorphology 42, 255–278.
Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADSeq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE 7, e37135.
Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
Prószyński, J., 2002. Remarks on Salticidae (Aranei) from Hawaii, with description of *Havaika* gen.n. Arthropoda Selecta 10, 225–241.
Rambaut, A., Drummond, A.J., 2010. TreeAnnotator version 1.6. 0. Univ. Edinburgh, Edinburgh, UK.
Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J., 2014. Tracer v1.6. http://beast.bio.ed.ac.uk/Tracer.
Reheis, M., 1999. Highest pluvial-lake shorelines and Pleistocene climate of the western Great Basin. Quat. Res. 52 (2), 196–205.
Reheis, M.C., Stine, S., Sarna-Wojcicki, A.M., 2002a. Drainage reversals in Mono Basin during the late Pliocene and Pleistocene. Geol. Soc. Amer. Bull. 114 (8), 991–1006.
Reheis, M.C., Sarna-Wojcicki, A.M., Reynolds, R.L., Repenning, C.A., Mifflin, M.D., 2002b. Pliocene to middle Pleistocene lakes in the western Great Basin: ages and connections. Great Basin Aqu. Syst. hist. 22, 53–108.
Reheis, M.C., Adams, K.D., Oviatt, C.G., Bacon, S.N., 2014. Pluvial lakes in the Great Basin of the western United States—A view from the outcrop. Quat. Sci. Rev. 97, 33–57.
Roof, S., Callagan, C., 2003. The climate of Death Valley, California. Bull. Amer. Met. Soc. 84 (12), 1725–1739.
Rosenblum, E.B., Sarver, B.A.J., Brown, J.W., Des Roches, S., Hardwick, K.M., Hether, T.D., Eastman, J.M., Pennell, M.W., Harmon, L.J., 2012. Goldilocks meets Santa Rosalia: an ephemeral speciation model explains patterns of diversification across time scales. Evol. Biol. 39 (2), 255–261.
Sağlam, I.K., Baumsteiger, J., Smith, M.J., Linares-Casenave, J., Nichols, A.L., O'Rourke, S.M., Miller, M.R., 2016. Phylogenetics support an ancient common origin of two scientific icons: Devils Hole and Devils Hole pupfish. Mol. Ecol. 25, 3962–3973.
Schenk, J.J., 2016. Consequences of secondary calibrations on divergence time estimates. PLoS ONE 11 (1), e0148228. https://doi.org/10.1371/journal.pone.0148228.
Smith, G.R., Dowling, T.E., Gobalet, K.W., Lugaski, T.S.D.K., Shiozawa, D.K., Evans, R.P., 2002. Biogeography and timing of evolutionary events among Great Basin fishes. Great Basin Aqu. Syst. hist. 33, 175–234.
Starrett, J., Derkarabetian, S., Hedin, M., Bryson, R.W., McCormack, J.E., Faircloth, B.C., 2017. High phylogenetic utility of an ultraconserved element probe set designed for Arachnida. Mol. Ecol. Res. 17, 812–823.
Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. 4 (1), vey016.
Swofford, D.L., 2002. PAUP*. Phylogenetic Analysis using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Mass.
Talavera, G., Castresana, J., 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst. Biol. 56 (4), 564–577.
Van Dam, M.H., Matzke, N.J., 2016. Evaluating the influence of connectivity and distance on biogeographical patterns in the south-western deserts of North America. J. Biogeogr. 43 (8), 1514–1532.
Wen, D., Nakhleh, L., 2018. Coestimating reticulate phylogenies and gene trees from multilocus sequence data. Syst. Biol. 67 (3), 439–457.
Williams, J.E., Wilde, G.R., 1981. Taxonomic status and morphology of isolated populations of the White River springfish, *Crenichthys baileyi* (Cyprinodontidae). Southwestern Nat. 485–503.
Wilson, J.S., Pitts, J.P., 2012. Identifying Pleistocene refugia in North American cold deserts using phylogeographic analyses and ecological niche modelling. Div. Dist. 18 (11), 1139–1152.
Zerbino, D.R., Birney, E., 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18 (5), 821–829.