SYNTHESIS

A legacy of geo-climatic complexity and genetic divergence along the lower Colorado River: Insights from the geological record and 33 desert-adapted animals

Greer A. Dolby | Rebecca J. Dorsey | Matthew R. Graham

1School of Life Sciences, Arizona State University, Tempe, AZ, USA
2Department of Earth Sciences, University of Oregon, Eugene, OR, USA
3Department of Biology, Eastern Connecticut State University, Willimantic, CT, USA

Correspondence
Greer A. Dolby, School of Life Sciences, Arizona State University, Tempe, AZ 85287, USA.
Email: gadolby@asu.edu

Funding information
GAD was supported by a U.S.G.S. award, a Fostering Postdoctoral Research in the Life Sciences seed grant from the School of Life Sciences at Arizona State University, and the Arizona Game and Fish Department (I7001). RJD was supported by a grant from the National Science Foundation (EAR-1546006). MRG was supported by grants from the Connecticut State University American Association of University Professors (CSU-AAUP) and the National Science Foundation (DEB-1754030).

Handling Editor: Carina Hoorn

Abstract
Aim: To review the histories of the Colorado River and North American monsoon system to ascertain their effects on the genetic divergence of desert-adapted animals.
Location: Lower Colorado River region, including Mojave and Sonoran deserts, United States.
Methods: We synthesized recent geological literature to summarize initiation phases of lower Colorado River evolution, their discrepancies, and potential for post-vicariance dispersal of animals across the river. We simulated data under geological models and performed a meta-analysis of published and unpublished genetic data including population diversity metrics, relatedness and historical migration rates to assess alternative divergence hypotheses.

Results: The two models for arrival of the Colorado River into the Gulf of California impose east-west divergence ages of 5.3 and 4.8 Ma, respectively. We found quantifiable river-associated differentiation in the lower Colorado River region in reptiles, arachnids and mammals relative to flying insects. However, topological statistics, historical migration rates and cross-river extralimital populations suggest that the river should be considered a leaky barrier that filters, rather than prevents, gene flow. Most markers violated neutrality tests. Differential adaptation to monsoon-based precipitation differences may contribute to divergence between Mojave and Sonoran populations and should be tested.

Main Conclusions: Rivers are dynamic features that can both limit and facilitate gene flow through time, the impacts of which are mitigated by species-specific life history and dispersal traits. The Southwest is a geo-climatically complex region with the potential to produce pseudocongruent patterns of genetic divergence, offering a good setting to evaluate intermediate levels of geological-biological (geobiological) complexity.

KEYWORDS
avulsion, earth-life evolution, genetic divergence, meta-analysis, monsoon, parapatric divergence, pseudocongruence, vicariance
Rivers are important physiographical features that act as barriers to migration and gene flow among terrestrial species (Hayes & Sewlal, 2004; Vences, Wollenberg, Vieites, & Lees, 2009), provide dispersal corridors and niche space for freshwater aquatic species (Burridge, Craw, & Waters, 2006; Echelle, 2008; Snorrason & Skúlason, 2009; Spencer, Smith, & Dowling, 2008), and create ecologically important estuarine and deltaic habitats where they meet the sea (Dolby, Ellingson, Findley, & Jacobs, 2018; Lau & Jacobs, 2017; Loneragan & Bunn, 1999; Swift, Findley, Ellingson, Flessa, & Jacobs, 2011). Rivers therefore influence biological evolution by structuring or isolating populations, limiting species ranges, and driving ecological adaptations. Yet rivers themselves are also dynamic systems that evolve over geological time-scales in response to many climatic and tectonic processes (e.g., Chapin, 2008; Clift & Blusztajn, 2005; Figueiredo, Hoorn, Ven, & Soares, 2009; Galloway, Whiteaker, & Ganey-Curry, 2011).

The lower Colorado River in the southwestern United States (Southwest) is a recently evolved drainage system situated in an arid and tectonically dynamic region (Figure 1). The Southwest has high levels of species richness and endemism (Mittermeier et al., 2003) among herpetofauna (Kiester, 1971; Persons & Nowak, 2007; Schall & Pianka, 1978), freshwater fishes (Spencer et al., 2008; Williams et al., 1989), plants (Baldwin et al., 2017; Kraft, Baldwin, & Ackerly, 2010; Thorndhill et al., 2017), gastropods (Hershler, Liu, & Howard, 2014; Hershler, Liu, & Mulvey, 1999), mammals (Davis, Koo, Conroy, Patton, & Moritz, 2008; Simpson, 1963), insects (McIntyre & Hostetler, 2001 and references therein), crustaceans (Witt, Threlfoff, & Hebert, 2006) and arachnids (Bryson, Riddle, Graham, Smith, & Prendini, 2013; Crews & Gillespie, 2014; Cushing, Graham, Prendini, & Brookhart, 2015; Hamilton, Hendrixson, & Bond, 2016). In seminal early studies, researchers often attributed origins of this biodiversity to formation of the Colorado River (Grinnell, 1914; Smith & Patton, 1980). Even early biogeographical work by Avise, Lamb and others (Lamb, Avise, & Gibbons, 1989; Lamb, Jones, & Avise, 1992) explored how the Colorado River impacted genetic signatures within and between species, particularly in the context of vicariant allopatric speciation, and these questions are still investigated today (Edwards et al., 2016; Graham, Wood, Henaut, Valois, & Cushing, 2017; Murphy et al., 2011; Vander Gast et al., 2013; Wood et al., 2012).

The Colorado River itself has been the focus of geological and stratigraphic investigations for over 100 years (Dutton, 1882; Powell, 1875) and today serves as a testing ground for ideas about the origin and evolution of continental-scale rivers (e.g., Dorsey, O’Connell, McDougall, & Homan, 2018; House, Pearthree, & Perkins, 2008; Howard, House, Dorsey, & Pearthree, 2015; Lucchitta, 1979; Meek & Douglass, 2001; Pearthree & House, 2014). As a result, the age and evolution of the Colorado River are relatively well-understood, with some differences between current models (see Background). Despite existing uncertainties, the wealth of geological data allows us to reconstruct palaeo-landscapes during stages of river evolution and use these models to evaluate which aspects of landscape evolution underlie present-day patterns of genetic diversity. By studying species with different ecologies and life histories in this context we can also learn how species vary in their responses to a set of shared extrinsic forces.

Despite the opportunity to integrate geological and phylogeographical evidence in this system, prior genetic studies have not fully considered how the complex history of the river has influenced population divergence, connectivity and adaptation through time (but see Graham et al., 2017). Furthermore, not all relevant geo-climatic factors have been considered when interpreting evolutionary patterns; for example, the influence of monsoon history has been largely overlooked. Incorporating the “true” geo-climatic complexity of the Southwest into biological models is challenging but necessary if we are to understand what processes have ultimately driven divergence, and among which species. Integrating these data may better enable empirical testing of speciation modes that are notoriously difficult to differentiate in nature among vertebrate species, such as parapatric and allopatric speciation (Coyne & Orr, 2004). Recent work in nearby systems has advanced the inclusion of geo-climatic complexity to understand lineage diversification. Meta-analysis of geological and genetic evolution on the Baja California peninsula (Dolby, Bennett, Lira-Noriega, Wilder, & Munguia-Vega, 2015) and assessment of 12 pairs of snake lineages across the Cochineal Filter Barrier (Myers, Hickerson, & Burbank, 2017) both illustrated the importance of individualistic responses to geo-climatic events, as well as the importance of considering multiple alternative hypotheses when determining relationships between external forces and diversification of a regional biota (Johannesson, 2010). This paper serves to integrate geological data with phylogeographical evidence to improve our mechanistic understanding of how the Colorado River has influenced biological diversification and evolution since its inception, explain sources of uncertainty and identify additional extrinsic forces that may have influenced divergence, namely monsoon evolution (Figure S2). Similar comparative phylogeographical approaches are commonly used to assess the impact of biogeographical events on regional faunas, but most focus on specific taxonomic groups (e.g., Barber & Klicka, 2010; Castoe et al., 2009; Daza, Castoe, & Parkinson, 2010; Myers et al., 2017). We performed a meta-analysis of phylogeographical data from 33 taxa (31 published, two unpublished) representing five animal groups and summarized the current state of knowledge regarding the inception and evolution of the Colorado River over the past c. 5 million years (Ma). We integrated these data to synthesize models of evolutionary history between the river and its adjacent biota during the major stages of Colorado River evolution and offer insights to guide future work in this complex region, particularly for studies using whole genome methods.

2 | GEOLOGICAL AND BIOLOGICAL BACKGROUND

2.1 | Tectonics, palaeoclimate and river evolution

The present-day Colorado River drainage covers a large area of the Colorado Plateau (Figure 1), and at 630,000 km² is the fourth largest
river catchment in the conterminous United States. Voluminous output of sediment from the Colorado River over the past c. 5 million years has filled deep basins in the Salton Trough and northern Gulf of California, exerting a major influence on plate-boundary faulting and crustal evolution (Dorsey, 2010). Intense management of the river for agriculture and urban development has dramatically
reduced the volume of river water and sediment exported to the ocean since the early 1900s. Prior to the closure of Hoover Dam in 1935, the river delivered a large annual discharge of sediment ($172 \pm 64$ Mt/yr) through Yuma, Arizona, to a network of delta distributary channels in the Salton Trough (Figure 1; Cory, 1913; Curtis, Culbertson, & Chase, 1973; Dorsey & Lazear, 2013; Meade & Parker, 1985).

The San Andreas fault is a major tectonic structure that defines the present-day transform boundary between the Pacific plate southwest of the fault, and North America plate northeast of the fault (Figure 1). Over the past ~7–8 million years, all crust southwest of the San Andreas fault has moved progressively northwest relative to North America at a rate of ~35–50 mm/yr (35–50 km/myr) (Figure 2; e.g., Bennett, Oskin, Iriondo, & Kunk, 2016; Oskin & Stock, 2003; Umhoefer et al., 2018). As a result, southwestern California, Baja California and the western Salton Trough have all translated northwest roughly 300–400 km relative to Yuma since late Miocene time. In order to interpret evolutionary patterns and infer their causal processes, it is necessary to interpret biological patterns in the context of the palaeo-landscape by using well-established tectonic reconstructions (e.g., Bennett et al., 2016; Umhoefer et al., 2018), which restore key locations to their former positions. We use these restorations to illustrate the two models of initiation (arrival to the Gulf of California) and early evolution of the Colorado River (Figure 2).

During middle to late Miocene time (~16–6 Ma), prior to formation of the Colorado River, the Colorado Plateau was a high-standing arid region with subdued topography in which rainwater flowed to internally drained basins (e.g., Cather et al., 2008; Chapin, 2008; Lazear, Karlstrom, Aslan, & Kelley, 2013). During this time an elongate tectonic lowland formed along the future path of the lower Colorado River with no fluvial connections between the basins (Figure 2a; Dickinson, 2002; McQuarrie & Wernicke, 2005). Late Miocene intensification of monsoonal precipitation on the Colorado Plateau likely was triggered by tectonic opening and marine flooding in the Gulf of California (Gulf; Chapin, 2008), which provides one of two major moisture sources to the present-day North American monsoon system (NAMS; Adams & Comrie, 1997; Higgins, Chen, & Douglas, 1999). By ~6.5–6.3 Ma marine waters inundated a long, narrow fault-controlled basin north to Palm Springs, California and possibly as far north as Parker, Arizona (Figure 2a vs. Figure 2e; Dorsey, Housen, Janecke, Fanning, & Spears, 2011; Dorsey et al., 2018; McDougall, Poore, & Matti, 1999; Umhoefer et al., 2018; see Figure 1 for locations). Increased monsoonal precipitation and/or increased groundwater flow off the Plateau created a series of large late Miocene lakes in the Lake Mead area (Crossey et al., 2015; Faulds et al., 2016). These lakes were terminated when the first-arriving Colorado River drove a sequence of downwind-directed lake filling and spilling floods that culminated in first arrival of the through-flowing river into that tectonic lowland (Chapin, 2008; Crossey et al., 2015; House et al., 2008; Pearthree & House, 2014).

There are two prevailing models for the birth and early evolution of the Colorado River: (a) "early initiation" at ~5.3 Ma, following roughly 1 Myr of marine deposition in the lower river valley (Figure 2a–d; Busing, 1990; Dorsey et al., 2007, 2011, 2018; McDougall, 2008; McDougall & Martinez, 2014; O'Connell, Dorsey, & Humphreys, 2017); or (b) "late initiation" at ~4.8 Ma, following a very short period (<50 Kyr) of lake deposition in the lower river valley (Figure 2e–h; Bright, Cohen, Dettman, & Pearthree, 2018; Bright et al., 2016; Crow et al., 2019; Gootee, Pearthree, House, Youberg, O’Connell, et al., 2016; Gootee, Pearthree, House, Youberg, Spencer, et al., 2016; House et al., 2008; Pearthree & House, 2014; Spencer & Patchett, 1997; Spencer et al., 2013). Both models propose emergence of the integrated river network after 6.0 Ma and before 4.5 Ma, but with notable differences within that time frame. The 5.3 Ma age of river initiation in the early initiation model (Figure 2b) is based on the age of the first arrival of Colorado River-derived sands in the western Salton Trough at Split Mountain Gorge (SMG, Figures 1 and 2) determined by palaeomagnetism, micropalaeontology and U-Pb ages of volcanic tuffs (Dorsey et al., 2007, 2011). The younger 4.8 Ma age (Figure 2g) is based on tuff ages and palaeomagnetic data from deposits north of Needles, CA and $^{40}$Ar/$^{39}$Ar ages from sandstones in Split Mountain Gorge (Crow, House, et al., 2018; Crow et al., 2019). For more details see Supporting Information.

In both models, the Colorado River was an integrated drainage network by c. 4.5 Ma, establishing its possible role as a barrier to east-west dispersal since then. However, volcanic activity and changes in the river course are known or presumed to have occurred 4.5 Ma to present. Lava dams formed intermittently in the Grand Canyon during the past c. 1 Myr (Crow et al., 2015; Crow, Karlstrom, McIntosh, Peters, & Dunbar, 2008; Dalrymple & Hamblin, 1998; Fenton, Cerling, Nash, Webb, & Poreda, 2002; Fenton, Poreda, Nash, Webb, & Cerling, 2004; Hamblin, 1994). Field mapping and $^{40}$Ar/$^{39}$Ar dating show that in 17 separate events basaltic lava flows blocked the river flow and formed lakes that persisted for 10s to 1,000s of years (Crow et al., 2015). All dams were located east of Lake Mead outside the geographical range for some species, but dams could have facilitated cross-river dispersal by temporarily reducing or ceasing flow downstream (Figure 1). Additionally, the Colorado River changed course through time via channel avulsions (Howard et al., 2008; Stouthamer & Berendsen, 2007) which may have permitted pulses of cross-channel dispersal intermittently since 4.5 Ma.

### 2.2 Genetic divergence

Nearly three decades of genetic and biogeographical studies have focused on species of the Southwest, particularly emphasizing herpetofauna. In the lower Colorado River region, species exhibit different spatial and temporal patterns of genetic divergence. Geographically, species vary in the amount of genetic divergence that occurs at the river. There is complete differentiation observed for the turret-building tarantula (*Aphonopelma preticense*; Graham, Hendrixson, Hamilton, & Bond, 2015), round-tailed ground squirrel (*Xerospermophilus tereticaudus*; Bell, Hafner, Leitner, & Matocq, 2010) and leopard frogs (*Lithobates onca* & *L. yavapaiensis*; Oláh-Hemmings et al., 2010). In contrast, there is little river-based
Comparison of two models showing initiation age and early evolution of the Colorado River: (a–d) Early-initiation model (e.g., Dorsey et al., 2018); (e–h) Late-initiation model (e.g., Pearthree & House, 2014). (a) Marine incursion into northern Gulf of California, Salton Trough, and lower Colorado River corridor c. 6.3 Ma (Bouse Formation basal carbonate member). (b) Earliest through-flowing Colorado River c. 5.3 Ma (Bouse siliciclastic member). (c) Pause in river sediment output and re-flooding of southernmost Colorado River Valley for ~200–300 Kyr at 5 Ma (Bouse upper bioclastic member). (d) Resumption of through-flowing Colorado River by ~4.5 Ma (Bullhead Alluvium). (e, f, g) Deposition in alluvial valleys and basins prior to initiation of the Colorado River at about 4.8 Ma, in late-initiation model. (h) Through-flowing Colorado River by ~4.5 Ma (Bullhead Alluvium), same in both models. Schematic beneath each panel shows whether the river would be a barrier during each stage: black dotted line—barrier only in south; "no" symbol—full barrier; grey dotted line—uncertain; black double arrow—no barrier. Note that this figure does not show channel avulsions or lava dams. Abbreviations: A, Amboy; B, Blythe; LV, Las Vegas; N, Needles; P, Parker; PS, Palm Springs; SMG, Split Mountain Gorge; SD, San Diego; Y, Yuma [Colour figure can be viewed at wileyonlinelibrary.com]
differentiation in the desert pocket mouse (Chaetodipus penicillatus; Jezkova, Jaeger, Marshall, & Riddle, 2009), Arizona hairy scorpion (Hadrurus arizonensis; Graham, Jaeger, Prendini, & Riddle, 2013a), and bighorn sheep (Ovis canadensis; Nelson; Buchalski et al., 2016), among others. This disparity may reflect species-specific differences in dispersal ability or niche specificity, but there is not good evidence for such patterns at this time. Flying insects, such as wasps and moths have shown to be little-affected by the river, dispersing across it easily (Smith et al., 2011; Wilson, Clark, Williams, & Pitts, 2012). Many species exhibit divergence associated with the ecological boundary dividing the Mojave and Sonoran deserts (Mojave-Sonoran “ecotone”). Of studies referencing ecotone divergence (Inman et al., 2014; Wood, Fisher, & Vander gast, 2014), most attribute this to the vegetation transition (Devitt, 2006; Graham et al., 2015; Jezkova et al., 2009; Mulcahy, Spaulding, Mendelson, & Brodie, 2006; Wood et al., 2012) despite the likelihood of an underlying abiotic control (a proximate mechanism) on the change in plant composition. The Mojave Desert has higher average elevation, a wider range of annual temperatures, and lower annual precipitation compared to the Sonoran Desert (Norris, 1958). Most importantly, the two deserts show strong differences in seasonal precipitation patterns. The Mojave Desert receives winter rainfall from westerly storm tracks (Higgins et al., 1999). The Sonoran Desert receives some of this winter precipitation but gets a large amount of summer rainfall from the NAMS (Figure S1), which tracks storms northeastward along the coast of the southern Baja California peninsula, and/or northward through the Gulf of California (Higgins et al., 1999; Higgins & Shi, 2001). Because most phylogeographical studies have focused on river-associated vicariance, the alternative explanation of divergence via differential ecological adaptation has not been adequately tested (see Pseudocongruence and Modes of Divergence).

Temporally, genetic divergence falls into two diffuse periods (Figure 3, Table S2). Many studies find late Miocene or Pliocene-aged divergence attributed to flooding of the Gulf of California and northern marine embayments (Devitt, 2006; Graham et al., 2015, 2017), or tectonic activity that generated topography of the Basin and Range province and/or the Colorado Plateau (Douglas, Douglas, Schuett, & Porras, 2006; Graham et al., 2015; Jaeger, Riddle, & Bradford, 2005). Other studies show divergence during Pleistocene or Holocene time attributed to climatic (glacial-interglacial) processes (Douglas et al., 2006; Gottscho et al., 2017; Graham et al., 2013a; Graham, Jaeger, Prendini, & Riddle, 2013b; Jezkova et al., 2009). There is usually concordance between LGM habitat predictions and the genetic evidence for glacial refugia (Jezkova et al., 2009, 2015, 2016; Graham et al., 2013a, 2013b), but these predictions are often coarse and many studies do not have fossil records against which to verify the LGM habitat predictions (Holmgren et al., 2014). A similar disparity in divergence ages has been observed in comparative studies of divergence on the nearby Baja California peninsula as well as across the Cochise Filter Barrier (Dolby et al., 2015; Myers et al., 2017).

Disparity among divergence ages confounds the ability to infer which geo-climatic processes have driven divergence. Species could be responding to separate events of different ages, or to the same event(s) at different rates because divergence ages may not equal the barrier age (Coyne & Orr, 2004). Alternatively, inter-study variance could be due to differences among the loci analysed, molecular

**FIGURE 3** Earliest date estimates (EDD) provided for taxa with genetic divergence in the lower Colorado River region; 95% high posterior densities displayed as bars ("posterior densities not available). Vertical lines denote the onset (and continuation to present day) of Colorado River and monsoon processes discussed in detail here where EM and LM are early and late initiation models, respectively. We note the two-step flooding timing for the Gulf of California. Grey box represents the duration of high amplitude 100-kyr Northern Hemisphere glaciations (not reviewed here). The upper EDD limit for Paravaejovis spp is 17.5 Ma (†). All estimates are based on mtDNA except the Uma spp., which used nuclear RADseq data (†). Divergence in 20 of the 25 species listed here post-dates the geologic processes discussed here. Note that some markers here may be affected by homoplasy [Colour figure can be viewed at wileyonlinelibrary.com]
clock calibrations used for those loci, geographic sampling (i.e. more or less genetic variation sampled) and parameterizations used to calculate divergence ages (e.g. generation time). Taken together, the relatively well-understood palaeo-landscape of the lower Colorado River and suite of intriguing (though variable) genetic patterns suggest much can be gained by integrating these data. Furthermore, these factors may make the Southwest a good setting to incorporate an intermediate level of geo-climatic complexity where non-mutually-exclusive evolutionary hypotheses can be tested.

3 | MATERIALS AND METHODS

3.1 | Geological data

3.1.1 | Approximating the Palaeo-floodplain of the Colorado River

To determine what regions of the Colorado River channel appear most variable, we used a GIS to estimate the spatial extent of the palaeo-floodplain. We plotted the locations of five stratigraphic units deposited by the river since its inception using available geological exposure data (Beard, Kennedy, Truini, & Felger, 2011; Crow, Block, et al., 2018). The oldest mapped unit is the late Miocene to early Pliocene Bouse Formation, which records conditions prior to and during initiation of the Colorado River. Younger river deposits include the Pliocene Bullhead Alluvium and Palo Verde alluvium, the late Pleistocene Chemehuevi Formation (~70 ka), and Holocene Blythe alluvium that forms the modern floodplain. These data were plotted in the original spatial coordinates from Crow, Block, et al. (2018; GCS = North American 1983, PCS = UTM Zone 11N, NAD 1983) and used as the base map for overlaying the polygon layer for extralimital populations (i.e. those occurring outside the typical range; see Extralimital Mapping).

3.2 | Genetic data

We explored how the lower Colorado River has impacted phylogeographical patterns among co-occurring terrestrial taxa by analysing data from published and unpublished animals with distributions that span the river. The published data represent species, clades within species (i.e. the western clade of Crotalus scutulatus), and monophyletic groups of closely related species (i.e. two Homalonychus spp.). Unpublished phylogeographical data from scorpions were generated for one species (Smeringurus mesensis) and one pair of sister species (Paravaejovis confusus & P. wearini). We used five datasets from highly dispersive flying insects to act as a “null” group whose dispersal we expect to be minimally affected by the river.

We chose data from studies with dense population-level geographical sampling and included microsatellite and AFLP datasets for consistency (these criteria excluded Dipposaurus dorsalis, Trimphodon biculatus, Peragnostus amplus/longimembris, Thomomys bottae and Sauromalus obesus; Table S3). To simplify comparisons and to standardize the data, we used a single marker per study; when multiple genetic markers were available we chose the marker that maximized geographical coverage. This approach resulted in 29 published and two unpublished mtDNA datasets, as well as two published RADseq datasets. Published data were downloaded from GenBank and annotated with location data (i.e. state, east of river, west of river). Mitochondrial data were aligned in Geneious Pro v7.1.7 (Biomatters Ltd) using MUSCLE (Edgar, 2004). Additional data sources are listed in Appendix 1.

Unpublished mtDNA were generated for scorpions collected at night using ultraviolet light detection following existing protocols (Graham et al., 2013a, 2013b, 2017). Briefly, genomic DNA was extracted from leg tissue using a DNeasy Tissue Kit (Qiagen) and a fragment of mitochondrial cytochrome c oxidase subunit I (COI) was amplified using primers COImodF and LE1r. We purified PCR products using ExoSAP-IT (GE Healthcare) and sent our samples to the DF/HCC DNA Resource Core (Harvard Medical School) for bidirectional sequencing. Double-stranded DNA fragments were manually aligned using Geneious Pro v7.1.7. Sequences were deposited in GenBank with accession numbers MN124015–MN124038; MF061548; MN124039–MN124071 (Table S7).

For analyses affected by sample size (diversity, migration history), only individuals sampled from California, Nevada, Utah, Arizona, New Mexico or mainland Mexico (not the Baja peninsula) were included to standardize sampling area. We grouped individuals within each dataset based on their location north/west of the Colorado River (West), south/east of the Colorado River (East) or mainland Mexico (Mexico). For haplotype networks, all available individuals were included to generate more complete networks and reduce the number of unsampled haplotypes. In some instances, this grouping ignores previously defined population structure, but for our purposes it enables us to specifically assess genetic diversity and relatedness between populations presently separated by the Colorado River to ascertain its efficacy as a biogeographical barrier. We used PGDSpider v2.1.1.5 (Lischer & Excoffier, 2012) for file conversions; input files are available in Supporting Information.

3.2.1 | Gene trees, simulations, trait decomposition

Gene tree reconstructions
Because most of the original datasets did not include outgroups, we obtained outgroups for our phylogenetic reconstructions by searching the first sequence of each dataset that was collected from CA, NV or AZ through BLAST-n on the NCBI website. In the search we excluded the species in the original dataset and chose outgroup sequences based on alignment scores (Table S4). We then fit rate and nucleotide substitution models for each dataset using jModelTest v2 (Darriba, Taboada, Doallo, & Posada, 2012; Guindon & Gascuel, 2003) and chose the AICc-ranked best model among JC, HKY, GTR + Gamma + Invariant Sites (Table S5). For each dataset we performed Bayesian phylogenetic reconstructions using MrBayes v3.2.6 (Ronquist et al., 2012) on the CIPRES Science Gateway (Miller, Pfeiffer, & Schwartz, 2010) using the taxon-specific nucleotide substitution and rates model, nucmodel = 4by4 for 15 million generations,
four chains per run with a burn-in of 25% and sample frequency of 10,000. We generated three runs per analysis and evaluated convergence by the standard deviation of split times reaching ≤ 0.01 and/or the Potential Scale Reduction Factor (PSRF) converging on 1.0 and/or the Effective Sample Size reaching ≥ 1,000 for parameters viewed in Tracer v1.7.1 (Table S5; Rambaut, Drummond, Xie, Baele, & Suchard, 2018). Using default settings for the sump and sumt commands (25% burn-in), we aggregated across runs to produce summary statistics and a consensus tree per species.

For simulated data (next section) we partitioned the three loci and used 25 million generations, sampling frequency of 20,000 and fixed the empirical state frequencies (statefreqpr = fixed(empirical)) because the base pair composition was enriched in As. For the two RADseq datasets (Uma scoparia, U. notata) we generated maximum parsimony trees with 100 bootstrap replicates and subtree-pruning-regrafting method in MEGA v7 (Kumar, Stecher, & Tamura, 2016) with the loci concatenated because the precise loci partitions were unknown. We processed all gene trees using FigTree v1.4.3 (Andrew Rambaut, http://tree.bio.ed.ac.uk/) and rooted trees by outgroup. For both Uma species and Homalonychus theologus/selenopoides outgroups were not available, so these trees are unrooted (Figures S3k,y,ee).

Simulations
To understand how predicted changes in the river would affect tree topologies, we simulated sequence data using FASTSIMCOAL2 v2.6.0.3 (Excoffier, Dupanloup, Huerta-Sánchez, Sousa, & Foll, 2013) under different physical scenarios. We simulated five replicate datasets under each of four models: (1) Three pulsed avulsions—a series of three unidirectional gene flow events at 10,000, 50,000, 100,000 generations ago with alternating directionality (e.g. M_{1→2}, M_{2→1}, M_{1→2})—this is intended to mimic the local effects of how the course of the river would shift over time. (2) Five pulsed avulsions—same as the previous model but with two additional older avulsions of alternating directionality at 250,000 and 500,000 generations ago to see if minor changes in the number of avulsions manifest as different tree topologies. Importantly, timing of gene flow for the avulsion models is arbitrary and results could vary under other parameterizations. (3) Lava dams—modelled as periods of symmetrical low-level migration during times when lava dams were active 50,000–100,000 generations ago, and 200,000–425,000 generations ago (100–200 ka and 400–850 ka based on 40Ar/39Ar dating). (4) Complete barrier—a null model of complete isolation without gene flow.

For avulsion models (models 1 and 2), in the historical events 20% of the genes in the source population migrated to the receiving population; population sizes were held constant before and after migration events. We used this high gene flow proportion to yield an observable signal in the topologies. In reality, avulsions would probably cause a lower magnitude, higher frequency effect on gene flow than modelled here. For the lava dam model (model 3), we used a low-level symmetrical migration matrix of 0.001 migrant/generation during migratory periods. Gene flow outside of these historical events or migratory periods was set to zero. In all models, populations were merged at 2.5 million generations ago. Each of the four models were run without recombination and then run separately with recombination (rate of 1e−7) to see if recombination modified the effects we observed from the migration scenarios. In each simulated we generated 6,000 bp of sequence data (including monomorphic sites using the -S flag) split evenly across three unlinked chromosomes with a mutation rate of 1e−8 bp per generation which reflects mtDNA mutation rates, and a transition fraction of 0.33, which implies no transition/transversion bias. We assumed no demographical changes (which is unrealistic), a generation time of 2 years (used to calculate the timing of lava dam events and to merge populations at 5 Ma) and set effective population sizes of 200,000 individuals per population, recording 40 samples per population at present-day (total N = 80). Changing the generation time would affect the rate at which the signal of cross-river mating is retained in populations but should not qualitatively change the results.

Trait decomposition
As phylogenies are used here to examine how related cross-river individuals are, we sought to quantify the level of geographical mixing across those topologies. To do this we used orthonormal decomposition of variance to measure how strong the phylogenetic signal was relating to a trait of interest—in this case the trait of interest was geographical position relative of the river (i.e. "east" or "west"). Because this method assumes a quantitative trait, we encoded east and west individuals as 1 and 2, respectively, and the outgroup as 0. Categorical data should not be encoded as numeric data in general, but because our trait variable is binary (0, 1), this dummy coding system should have minimal overall effect on the results. To assess if the analysis was influenced by how the outgroup was encoded, we ran the same analysis again on a subset of five datasets where the outgroup was first encoded as an intermediate value (=1.5) and then as the opposite-endmember value (=3). We compared the results of these sensitivity tests to the original results (Table S6). With the orthogram() function in the R package adephylo v1.1–7 (Jombart & Dray, 2010) we calculated four metrics of association between tip states and phylogenetic position for eight simulated datasets and 26 species based on the rooted phylogenies from MrBayes (unrooted phylogenies were used for Homalonychus spp, Uma spp, and simulated datasets). We recorded the following statistics per analysis: Dmax (akin to Kolmogorov-Smirnov goodness-of-fit test between trait states and phylogeny), R2max (a measure of whether a single node accounts for transition in the trait value vs. the transition being diffuse across nodes), SCE (average local variation of the trait value on the gene tree) and SkR2k (a measure of whether trait variance is skewed towards the termini or root). Monte Carlo simulations (N = 999) were used to determine significance for each metric. In an idealized scenario where the river is a perfect barrier to dispersal and the lineages were fully sorted, the trait states should be entirely explained by the topology (i.e. very high phylogenetic signal). The phylogenetic signal is expected to deteriorate as cross-river mixing increases or in the case of incomplete lineage sorting. We note that because the populations are not panmictic there will automatically
be some phylogenetic signal in these test statistics. The purpose of using these tools is for relative comparisons to see: 1) whether our empirical datasets have similar or less geographical mixing across topologies (phylogenetic signal) than the null insect group, 2) whether geographical mixing (phylogenetic signal) of the simulated data is similar to or falls outside the range observed for empirical datasets.

3.2.2 | Population divergence and diversity

Divergence ages

We collected the earliest divergence date estimates in the Colorado River area for all published datasets from their primary literature sources. For the new scorpion datasets, we aligned the COI data in GENEIOUS using MUSCLE (Edgar, 2004) and simultaneously assessed phylogenetic relationships and divergence dates using Bayesian Inference (BI) in BEAST v1.8.0 (Drummond, Suchard, Xie, & Rambaut, 2012). `PARTITIONFINDER2` (Lanfear, Frandsen, Wright, Senfeld, & Calcott, 2016) was used to determine best-fit substitution models. We generated an xml file in BEAUTi (BEAST package) using the appropriate substitution models (TN93 + G for both taxa), uncorrelated lognormal clock models for each gene, and the coalescence constant size tree model. Divergence times were estimated by calibrating the mean substitution rates and standard deviation of the COI alignment using uniform clock-rate priors bounded at 0.00734 and 0.00393 substitutions/site/Myr, as calculated for related species (Bryson, Jaeger, Lemos-Espinal, & Lazcano, 2012; Graham et al., 2017). The MCMC runs were conducted for 50 million generations and sampled every 5,000 generations. We confirmed stationarity and convergence of Markov chains and appropriate effective sample sizes in TRACER. A maximum clade credibility tree was produced using TREEANNOTATOR (BEAST package). For each taxon, we considered “earliest date estimates” to be the age of the oldest node connecting individuals found east and west of the Colorado River.

Haplotype relatedness and genetic diversity

To assess haplotype relatedness of populations east and west of the Colorado River, we generated simple ($\nu = 0$) median-joining networks in PorART (Population Analysis with Reticulate Trees) v1.7.2 (Leigh & Bryant, 2015) using datasets that included Baja California individuals when available. The median-joining approach favours short connections and combines minimum-spanning trees into a single network, making it appropriate for phylogeographical datasets and applicable to the mtDNA and RADseq data (Bandelt, Forster, & Röhl, 1999).

To calculate standard diversity and locus statistics, we performed a suite of analyses in ARLEQUIN v3.5.2.2.2 (Excoffier, Laval, & Schneider, 2005), assessing per-species locus neutrality and signs of demographical change (Tajima’s $D$, Fu’s $F_{S}$), population structure and diversity ($\theta_W$, $\theta_S$, AMOVA in haplotypic format), population differentiation using Weir and Cockram’s pairwise $F_{ST}$, and exact test for population differentiation using the default 100,000 Markov chain steps and 10,000 dememorization steps. We excluded mainland Mexico individuals from these statistics because this area was not sampled for all species and could bias inter-species comparisons. To ascertain whether there were aggregative (i.e. cross-taxon) geographical patterns among these statistics we ran a paired $T$-test or Wilcoxon Signed Rank test to compare populations per statistic, depending on whether normality assumptions were met as determined by a Shapiro-Wilk test. For each test, values from the statistics were grouped as west or east and paired by species. The among-population (i.e. east vs. west) variances from AMOVA were analysed with a Mann–Whitney U test that compared arachnids, mammals and reptiles to the null, highly dispersing insect group to assess the effect of the river on differentiation.

3.2.3 | Historical gene flow and extralimital patterns

Gene flow

To estimate the amount of historical gene flow among populations presently separated by the river, we used MIGRATE-N v3.6 (Beerli & Palczewski, 2010) set to Bayesian Inference mode. We excluded Chaetodipus baileyi, Crotalus cerastes, Odontoloxozus longicornus, Paruroctonus becki and Uma scoparia from these analyses because they did not have adequate sampling on both sides of the river to estimate model parameters (Table S1). We excluded mainland Mexico individuals from this analysis to standardize the comparison across species. For each taxon we ran one long chain for three million generations with 20% burn-in discarded as suggested by the developers. To better sample model space, we used static heating on four chains with temperatures set manually from cold to hot: 1.0, 1.5, 3.0, 1,000,000. We ran three independent replications per analysis, assessing convergence of independent runs using the autocorrelation criterion (converged runs should achieve autocorrelation score near 1) and minimum Effective Sample Size of 1,000 per parameter. The ddRADseq dataset for Uma notata was parameterized as linked DNA sequence data instead of SNP data because it contained invariant regions and the loci partitions were unknown. We used default uniform priors on mutation-scaled population size, $\theta$ (0–0.1 with a delta of 0.01) and mutation-scaled migration parameter, $M$ (0–1,000 with a delta of 100). Based on trial runs, we increased the uniform prior limit on $\theta$ for M. flagellum, Homalonychus spp., P. platyrhinos and Paravaejovis spp. to 0–0.25 (delta of 0.01), and for D. merriami to 0–1.0 (delta of 0.1). For A. punctatus, O. canadensis and P. sordidus we increased the scaled migration parameter $M$ to 0–2,500 (delta of 200); for P. weethumpi we increased $M$ to 1–10,000 (delta 1,000), which is a large prior and probably reflects the unstructured populations and life history of this species. To determine the effective number of migrants, we multiplied each $M$ by $\theta$ of the receiving population to produce $N_m$ and $4N_m$ values as given in the manual for haploid and RADseq data, respectively; for example, $xN_mE\rightarrow W = M_{E\rightarrow W}/\theta$.

Extralimital mapping

The Colorado River limits many species ranges, but some of these species (or clades) have a small number of narrowly distributed populations on the opposite side of the river (extralimital populations). The spatial extent of such extralimital populations can reveal information about their origin. For example, if a species is narrowly
distributed along the length of the opposite side of the river it would suggest that the river is only weakly limiting dispersal (or is a complete barrier and there is incomplete lineage sorting). Conversely, a narrow patch (or "hotspot") of extralimital populations occurring in several species suggests that there are special characteristics of the river in that particular place that allows dispersal—perhaps the river is particularly passable in that location or changed course there recently. To adjudicate between these scenarios, we developed an extralimital mapping approach. For each species/clade that is geographically widespread on one side of the river except for a small number of cross-river populations, we drew a polygon around those cross-river extralimital populations of each species (Figure S6). We excluded species/clades that were distributed broadly on both sides of the river or were entirely exclusive to one side of the river. GPS data for all datasets were not available, so we manually drew polygons following the sample collection map and genetic results in the original published studies (this included results based on evidence not included here, such as microsatellite data). Populations were identified as extralimital if a clade or haplotype group was primarily confined to one side of the river but exhibited a much smaller cross-river geographical range compared to their home range, whether it admixed with neighbouring populations or not. We chose studies that sampled the range of each species reasonably well (i.e. that the extralimital populations are not a sampling artefact). To visualize if these extralimital populations from different species occurred in similar geographical areas, we used ArcMap v10.6.1 (ESRI) to sum the number of spatially overlapping extralimital polygons. We merged polygons using the geoprocessing merge tool, used feature to polygon and feature to point tools to identify individual centroids, performed a spatial join between unique centroids and the merged polygon layer, and symbolized the Join_Count field to show the number of species that had overlapping extralimital populations (Honeycutt, 2012). We overlaid this layer on the palaeo-floodplain map (see Geological Data). We note that there are clear limitations

**FIGURE 4** Bayesian gene tree reconstructions for: (a) selected empirical datasets (left), with individuals northwest of the river coloured blue, and individuals southeast of the river coloured red; (b) non-recombining, partitioned sequence data simulated under four models: three unidirectional avulsions, five unidirectional avulsions, low-level symmetrical gene flow via lava dams and complete barrier (null). Yellow bars highlight examples of nested paraphyly/mixing. Black dots are branches with a posterior value of 0.85 or above. See Figure S3 for all species and Figure S4 for the simulations with recombination [Colour figure can be viewed at wileyonlinelibrary.com]
of approaching this method manually, so we interpret these results with caution.

4 | RESULTS

4.1 | Phylogenetic relatedness and trait decomposition

Through blast we obtained outgroups for all species except Homalonychus spp. and the Uma datasets. All tree reconstructions reached convergence by at least one of our criteria (Table S5). Of the gene trees generated from 31 empirical datasets and eight simulated datasets, topologies generally exhibit two major clades, one predominantly east-of-river and one predominantly west-of-river (e.g. Crotaphytus bicintores, Xantusia vigilis, Dipodomys deserti; Table 2, Figure 4; Figure S3), while other species have two clades that are greatly mixed (e.g. Chateodipsus penicillatus, Smeringurus mesensi, Hadrurus arizonensis). The east versus west clade pattern is reflected in the Dmax statistic, which was significant in most species but often had an intermediate value, indicating that there is a phylogenetic signal in whether individuals are east or west of the river but that this signal is not absolute. Overall, it appears that the river is likely to filter gene flow to a measurable extent, but not entirely. Our trees using simulated data produced similar results to the empirical trees (Figure 4b), suggesting that nested paraphyletic relationships like those observed empirically can be achieved through pulses of unidirectional gene flow (i.e. river avulsions) or periods of low-level bidirectional gene flow (i.e. lava dams). The complete barrier scenario looked very different, and fell towards the high end of empirical values (Figure 5). Based on topological statistics from PAUP (Swofford, 2003; see Supporting Information), simulated trees were different from empirical gene trees in measures of character fit (Retention Index, Rescale Consistency Index), but were not statistically different among consensus indices that reflect topological structure and clade resolution (Figure S7; Colless weighted consensus fork, Schuh-Farris levels-sum; Miyamoto & Cracraft, 1991).

Dmax and SCE values are greater in non-insect groups than the insect group—this is expected because the insect group has higher gene flow that should erode the relationship between phylogeny and geographic position (Table 2; Figure 5, Figures S8 and S9). None of the SkR2k values were statistically significant and the values range from low (biased towards root) to high (biased towards tips), which means that there is no pattern as to whether the east-west transition is deep or shallow within the trees. The R2max scores were significant for most taxa, indicating that there is often a shift in each tree between “eastness” and “westness” concentrated in a single node rather than a diffuse pattern. Three of the five insect species had low values that were not significant for R2max, which would be expected under a scenario of high gene flow. Other species not significant for R2max were Scoloporus magister, Paruroctonus beckii, Paravaeovis confuses/waeringi, Masticophus flagellum, Dipodomys merriami, Crotalus mitchelli and both Smeringurus species, although S. vachoni was borderline ($p = .048$)—these taxa show a diffuse transition between east and west states and/or the trait is less conserved in a clade (i.e. more geographical mixing). All simulated trees were significant for R2max. Trees generated from our complete barrier models showed reciprocal monophyly in every simulation. For models 1–3 there was some stochasticity among the five simulated datasets; we show a single representative tree for each.

4.2 | Population diversity and relatedness

In most species genetic variation was higher on one side of the river than the other based on $\theta_s$ and $\theta_L$ (Table 1), but this pattern was not consistent across species ($p = .3,.9$ for Wilcoxon Signed rank test, respectively), meaning one side of the river does not consistently have higher diversity. Tajima’s $D$ was negative but not significant for most populations whereas Fu’s $F_S$ was almost always negative and significant ($\alpha = .05$), suggesting that many of the single-locus datasets are affected either by purifying natural selection, linkage to neighbouring regions under selection, or that populations have undergone recent demographical expansion. There was no significant cross-species pattern for which side of the river exhibited stronger signatures of non-neutrality ($p = .8,.3$ for Tajima’s $D$).
and Fu’s $F_S$ with a Wilcoxon Signed Rank test, respectively. Many Tajima’s $D$ analyses may be underpowered relative to Fu’s $F_S$, which showed statistical significance more often. Haplotype networks show some population structure (Figure 6), but west and east haplotypes were mixed in most networks. Networks were completely segregated in Aphonopelma prenticei (arachnid), Chaetodipus baileyi and Xerospermophilus tereticaudus (mammals), Gopherus spp. (reptile), Lithobates spp. (amphibian), Odontoloxozus longicornis (insect). Our cross-river AMOVA revealed that arachnids, mammals and reptiles had genetic differentiation significantly better explained by the river than the more dispersive insects ($p = .004$; Figure 7).

4.3 Historical migration and extralimital patterns

Of the 29 species analysed for historical migration rates with Migrate-n, there was expectedly high gene flow among insects

| TABLE 1 Molecular diversity, neutrality and pairwise differentiation ($F_{ST}$) statistics for species included in this study |
|-----------------|-----------------|-----------------|---------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                 | $\theta_s$ | $\theta_n$ | Tajima’s $D$ | Fu’s $F_S$ | AMOVA (% variance among E-W) |
|                 | West | East | West | East | West | East | West | East | West | East |
| **Amphibians** |
| Anaxyrus punctatus | 4.9 | 2.5 | 4.5 | 2.4 | -0.2 | -0.08 | -25.7* | -26.8* | 20 |
| Lithobates onca/ yavapaiensis | 0.2 | 1.3 | 0.5 | 2.5 | 1.4 | -1.2 | 2.7* | 11.9* | 93 |
| **Reptiles** |
| Chionactis occipitalis | 31.3 | 28.3 | 50.0 | 20.7 | 1.8 | -1.0 | -24.2* | -24.1* | 33 |
| Crotaphytus bicinctores | 38.0 | 46.7 | 66.7 | 78.0 | -9.4 | 1.6 | -6.5* | -18.5* | 42 |
| Crotalus mitchelli | 26.2 | 6.8 | 31.9 | 5.1 | 0.9 | -1.1 | -4.1* | -9.3* | 42 |
| **Mammals** |
| Chaetodipus penicillatus | 10.2 | 14.9 | 13.1 | 12.8 | 0.7 | -0.5 | -24.2* | -24.3* | 21 |
| Dipodomys deserti | 16.0 | 17.2 | 12.1 | 18.2 | -0.8 | 0.20 | -24.0* | -24.1* | 40 |
| Dipodomys merriami | 19.9 | 20.9 | 14.1 | 15.6 | -0.9 | -0.9 | -24.0* | -24.3* | 25 |
| **Arachnids** |
| Aphonopelma prenticei | 30.7 | 18.6 | 34.5 | 22.4 | 0.45 | 0.76 | -24.1* | -21.3* | 42 |
| Homalonychus theologus/ selenopoides | 34.9 | 34.2 | 31.9 | 29.9 | -0.3 | -0.5 | -24.0* | -23.9* | 73 |
| Hadrurus arizonensis | 16.6 | 15.3 | 9.6 | 12.1 | -1.3* | -0.7 | -24.3* | -24.3* | 15 |
| Paruroctonus becki | 25.7 | 10.0 | 23.9 | 10.0 | -0.2 | 0.0 | -24.0* | 1.1 | 22 |
| **Insects** |
| Pogonomyrmex rugosus | 10.3 | 27.9 | 8.6 | 39.7 | -1.3 | 1.5 | -6.4* | -11.3* | 24 |
| Prodoxus sordidus | 5.8 | 1.4 | 36.6 | 228.9 | -1.5* | -1.5* | -2.5 | 3.1 | 17 |
| Prodoxus weethumpi | 5.0 | 5.7 | 75.1 | 36.8 | -1.3 | -0.6 | -3.6 | 0.7 | 1 |
| Sphaerophalma arota | 17.9 | 14.0 | 71.7 | 92.7 | 1.1 | 2.1 | -3.5* | -3.4* | 12 |
| Tegeticula antithetica | 16.8 | 8.2 | 19.0 | 8.0 | -0.9 | -1.3 | -24.0* | -24.9* | 3 |

Note: Asterisk (*) denotes significance ($\alpha = .05$). AMOVA column shows the percent of genetic variance explained by the river (geographical groupings “east of river” and “west of river”); these data are visualized in Figure 7.
(median Nm of 6.5), except for *Sphaeropthalma arota*. There was low gene flow (Nm ≤ ~1) in at least one direction of 13 different species (Table 3) and those species whose gene trees exhibited reciprocal monophyly, such as *Lithobates onca/yavapaiensis*, *Uma notata* and *Aphonopelma prenticei*, showed low effective migration (0.1, 0.9 and 1.0, respectively). The fact that these species have non-zero migration rates may suggest that migration rates here are slightly inflated, perhaps due to rapidly evolving mtDNA, or that our approach was not a good fit for some species (maybe due to low sample size). Six species had west-biased gene flow and seven had east-biased gene flow, reflecting no overall directionality. Four of the five mammals had east-biased gene flow, but

**Table 2** Orthogram-based decomposition of variance metrics are based on statistically comparing the fit of terminal trait states (“east” or “west” of river) with the rooted topology of each species

<table>
<thead>
<tr>
<th>Species</th>
<th>R² max</th>
<th>SkR²k</th>
<th>Dmax</th>
<th>SCE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Amphibians</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Anaxyrus punctatus</em></td>
<td>0.30</td>
<td>12.67</td>
<td>0.71</td>
<td>14.60</td>
</tr>
<tr>
<td><em>Lithobates onca/yavapaiensis</em></td>
<td>0.95</td>
<td>7.52</td>
<td>0.94</td>
<td>78.58</td>
</tr>
<tr>
<td><strong>Reptiles</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Crotalus mitchelli</em></td>
<td>0.20</td>
<td>17.10</td>
<td>0.16</td>
<td>0.19</td>
</tr>
<tr>
<td><em>Crotalus scutulatus</em></td>
<td>0.58</td>
<td>31.30</td>
<td>0.57</td>
<td>8.97</td>
</tr>
<tr>
<td><em>Lichanura trivirgata</em></td>
<td>0.49</td>
<td>10.67</td>
<td>0.75</td>
<td>22.15</td>
</tr>
<tr>
<td><em>Gopherus agassizii/mora fkai</em></td>
<td>0.53</td>
<td>6.53</td>
<td>0.59</td>
<td>4.19</td>
</tr>
<tr>
<td><em>Masticophis flagellum</em></td>
<td>0.27</td>
<td>7.75</td>
<td>0.50</td>
<td>1.88</td>
</tr>
<tr>
<td><em>Phrynosoma platyrhinos</em></td>
<td>0.14</td>
<td>80.51</td>
<td>0.29</td>
<td>6.71</td>
</tr>
<tr>
<td><em>Sceloporus magister</em></td>
<td>0.12</td>
<td>24.47</td>
<td>0.44</td>
<td>4.11</td>
</tr>
<tr>
<td><em>Uma notata</em></td>
<td>0.23</td>
<td>16.62</td>
<td>0.25</td>
<td>0.91</td>
</tr>
<tr>
<td><em>Uma scoparia</em></td>
<td>0.53</td>
<td>11.44</td>
<td>0.06</td>
<td>0.65</td>
</tr>
<tr>
<td><em>Xantusia vigilis</em></td>
<td>0.26</td>
<td>25.06</td>
<td>0.70</td>
<td>66.35</td>
</tr>
<tr>
<td><strong>Mammals</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Dipodomys deserti</em></td>
<td>0.08</td>
<td>83.19</td>
<td>0.48</td>
<td>25.73</td>
</tr>
<tr>
<td><em>Dipodomys merriami</em></td>
<td>0.06</td>
<td>60.26</td>
<td>0.34</td>
<td>10.85</td>
</tr>
<tr>
<td><em>Ovis canadensis</em></td>
<td>0.10</td>
<td>92.89</td>
<td>0.76</td>
<td>139.28</td>
</tr>
<tr>
<td><em>Xerospormophillus tereticaudus</em></td>
<td>0.52</td>
<td>5.16</td>
<td>0.46</td>
<td>0.69</td>
</tr>
<tr>
<td><strong>Arachnids</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Aphonopelma prenticei</em></td>
<td>0.13</td>
<td>32.91</td>
<td>0.24</td>
<td>2.01</td>
</tr>
<tr>
<td><em>Homalonychus theologus/selenoides</em></td>
<td>0.19</td>
<td>18.98</td>
<td>0.75</td>
<td>25.41</td>
</tr>
<tr>
<td><em>Hadrurus arizonensis</em></td>
<td>0.51</td>
<td>11.87</td>
<td>0.85</td>
<td>67.76</td>
</tr>
<tr>
<td><em>Paruroctonus becki</em></td>
<td>0.15</td>
<td>11.03</td>
<td>0.21</td>
<td>0.22</td>
</tr>
<tr>
<td><em>Parvaeosinus confuses/wearingi</em></td>
<td>0.40</td>
<td>32.12</td>
<td>0.38</td>
<td>3.05</td>
</tr>
<tr>
<td><em>Smeringurus meseaensis</em></td>
<td>0.18</td>
<td>11.27</td>
<td>0.32</td>
<td>1.41</td>
</tr>
<tr>
<td><em>Smeringurus vachoni</em></td>
<td>0.15</td>
<td>11.03</td>
<td>0.21*</td>
<td>0.22</td>
</tr>
<tr>
<td><strong>Insects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Pogonomyrmex rugosus</em></td>
<td>0.23</td>
<td>17.77</td>
<td>0.38</td>
<td>2.08</td>
</tr>
<tr>
<td><em>Prodoxus sordidus</em></td>
<td>0.23</td>
<td>22.75</td>
<td>0.34*</td>
<td>1.31</td>
</tr>
<tr>
<td><em>Prodoxus weethumpi</em></td>
<td>0.30</td>
<td>6.35</td>
<td>0.42</td>
<td>1.17</td>
</tr>
<tr>
<td><em>Sphaeropthalma arota</em></td>
<td>0.25</td>
<td>14.53</td>
<td>0.26</td>
<td>0.24</td>
</tr>
<tr>
<td><em>Tegeticula antithetica</em></td>
<td>0.16</td>
<td>21.03</td>
<td>0.41</td>
<td>1.99</td>
</tr>
<tr>
<td><strong>Simulations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complete barrier, no recombination</td>
<td>0.32</td>
<td>8.08</td>
<td>0.76</td>
<td>17.20</td>
</tr>
<tr>
<td>3 avulsions, no recombination</td>
<td>0.45</td>
<td>6.87</td>
<td>0.72</td>
<td>17.36</td>
</tr>
<tr>
<td>5 avulsions, no recombination</td>
<td>0.37</td>
<td>7.90</td>
<td>0.81</td>
<td>17.92</td>
</tr>
<tr>
<td>Lava dams, no recombination</td>
<td>0.82</td>
<td>6.21</td>
<td>0.89</td>
<td>19.95</td>
</tr>
</tbody>
</table>

Note: Bolded values are significant at the 0.01 level, bolded asterisked values are significant at the 0.05 level. All analyses are on single locus datasets except *Uma scoparia* and *U. notata*, which are based on a tree generated with RADseq data.
this is difficult to interpret without a larger sample size and more detailed study of the organisms themselves. Because mtDNA patterns are not directly comparable to large-scale nuclear data (in this case RADseq data), we include the results from Uma with caution, although they are generally consistent with the other data sets (Table 3).

The extralimital mapping analysis revealed that the Mohave Valley hosts an exceptionally high number of extralimital populations east of the river (Figure 8: west of the Black and Mohave Mountains, AZ). The Yuma region hosts a smaller number of extralimital populations. Species with extralimital populations east of the river include two mammals (Chaetodipus penicillatus, Ovis canadensis), seven reptiles (Chionactis annulata, Chionactis occipitalus, Crotaphytus bictores, Gopherus spp.: based on microsatellite data; Edwards et al., 2015, Lichanura trivirgata, Phrynosoma platyrhinos, Uma scoparia), two arachnids (Smeringurus spp. Hadrurus arizonensis Clade III), and two insects (Prodoxys sor didus, P. weethumpi). There were very limited extralimital populations west of the river and neither of the two amphibians had extralimital populations.
4.4 | Limits of mitochondrial data

All results we obtained in this study are heavily reliant on mtDNA data, which has inherently low effective population sizes because it is haploid and mostly maternally inherited. Additionally, the mitochondrial genome does not recombine so the effects of linkage, hitchhiking and selection can be particularly impactful, and the effects of isolation and gene flow are particularly acute relative to signals retained in nDNA. The rapidly evolving nature of these markers also mean that homoplasy may affect divergence ages and measures of relatedness, as well as the confidence intervals surrounding them, though the sequence evolution should be accounted for in Bayesian Inference analysis.

5 | DISCUSSION

5.1 | Palaeo-landscape and genetic divergence

5.1.1 | Age of divergence

The two prevailing models for early Colorado River evolution make different predictions about the age of river arrival and depositional setting of the southern valley immediately prior to initiation (see Background; Figure 2), but both models agree on the stepped lake spillover process of initiation. For terrestrial species, the important difference between the two geological models is the river initiation age, which would determine the onset of east-west (E-W) isolation.

In the early initiation model, E-W populations would be isolated between Yuma and Parker, Arizona during northward incursion of the Gulf of California starting c. 6.3 Ma (Figure 2a). For subtropical species with ranges extending only a short distance north of the present-day head of the Gulf (“Northern Taxa”, Figure 10 in Dolby et al., 2015), this may have imposed complete E-W isolation through today. For species with ranges extending farther north (e.g. through southern Nevada) the incursion would have been an incomplete barrier—becoming a complete E-W barrier for all species when the river connected to the Gulf at 5.3 Ma. In contrast, the late initiation model proposes no northward marine incursion beyond Yuma, Arizona and therefore a divergence age of c. 4.8 Ma is expected (a difference of 0.5 Myr between models, which is probably biologically negligible).

The divergence ages of species studied here vary considerably, but 20 of the 25 divergence ages post-date flooding of the northern Gulf and initiation/strengthening of the NAMS (6.3 Ma) and/or Colorado River initiation (5.3–4.8 Ma); the five exceptions are Paravacajovis spp., Aphonopelma prenticei, Chionactis occipitalis, Sphaerophalma arota and Tegeticula antithetica (two arachnids, one reptile, and two insects, respectively; Figure 3). The three mammals exhibit much younger ages than other groups (≤1.6 Ma) and seven species in total show divergence younger than 0.8 Ma, which marks the onset of high-amplitude Northern Hemisphere glacial cycles. Importantly, these age estimates are almost entirely based on rapidly evolving mtDNA markers and are only as accurate as the calibrations used. There can be issues with saturation (homoplasy) that bias estimates towards younger values because not all mutations are counted. We note, however, that the nuclear RADseq data for Uma also produced very young ages (0.31–0.69 Ma). Ages from older work regarding the “Bouse embayment” have been used to either calibrate or calculate divergence ages in Gopherus (Avise, Bowen, Lamb, Meylan, & Bermingham, 1992; Edwards et al., 2015, 2016), Xeroperomorphus (Bell et al., 2010), Pterygomeirus (Bryson et al., 2012) and Crotalus (Bryson, Murphy, Graham, Lathrop, & Lazcano, 2011; Castoe, Spencer, & Parkinson, 2007). New and ongoing geological work reviewed here should be used to revise the use and application of such datums in the future. Given the small age difference between models of river initiation and the variance surrounding genetic divergence ages (0.21–11.89 Ma), a geo-genomic approach (Baker et al., 2014; Fritz & Baker, 2017) using molecular clocks to independently evaluate the river initiation models appears infeasible, even if they are well-calibrated. However, considering the geological uncertainty in these models is important if river-based vicariance is being used to estimate node splits, mutation rates or constrain model priors (e.g. in ABC analyses).

A major complication for calculating divergence ages comes from the potential for post-divergence gene flow (see Modes of Post-Initiation Dispersal), which would render E-W populations more similar than expected under a strict allopatric model, biasing molecular clocks towards younger ages (Leaché, Harris, Rannala, & Yang, 2013). Our tree reconstructions (Figure 4), trait decomposition analysis (Table 2), migration analyses (Table 3), haplotype networks (Figure 6) and extralimital mapping analysis (Figure 8) together indicate cross-river mixing of some degree, which can be explained either by genetic mixing after river initiation or incomplete lineage sorting. Directly testing between these is not possible without additional data. Our simulations under strict vicariance reveal reciprocal monophony (i.e. no incomplete lineage sorting; Figure 4), but because these simulations are idealized, they offer only limited support for post-vicariance mixing over incomplete lineage sorting. If post-vicariance mixing is a correct interpretation, then the river should be considered a leaky barrier that filters gene flow without eliminating it. The few empirical datasets without mixing (i.e. reciprocal monophony) occur within all four dispersal-limited groups: A. prenticei, C. baileyi, Lithobates spp., U. notata (excluding Mexico) and X. tereticaudus, but not insects.

If we assume the post 6.3-Ma genetic divergence ages (Figure 3) are fairly accurate and not due to incomplete lineage sorting and/or homoplasy, then they can be easily explained if the Colorado River initiation and/or strengthening of the NAMS began filtering (i.e. reducing) gene flow between 4.8–6.3 Ma and has allowed some low level of gene flow since. In this scenario the variance surrounding divergence dates is reasonable as we expect species’ differences to affect these measurements. Generation times would affect the rate at which genetic signals of outbreeding are retained (Ohta, 1992; Stolle, Kidner, & Moritz, 2013) and differences in real effective population size (ne) or the result of marker choice (e.g. mtDNA vs. nDNA) would also influence the time to coalescence and divergence age estimates. For example, the signal from short-lived gene flow would be retained longer in a species that is long-lived with a long-generation
TABLE 3  Median values for number of effective migrants (Nm) from Migrate-n and directionality of historical gene flow (Uma estimates are 4Nm and based on RADseq data)

<table>
<thead>
<tr>
<th>Species</th>
<th>Median Nm_E→W</th>
<th>Median Nm_W→E</th>
<th>Directionality (≥25% bias)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Amphibians</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaxyrus punctatus</td>
<td>12.5</td>
<td>3.1</td>
<td>Westward</td>
</tr>
<tr>
<td>Lithobates onca/yavapaiensis</td>
<td>0.1</td>
<td>0.1</td>
<td>–</td>
</tr>
<tr>
<td><strong>Reptiles</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chionactis occipitalis</td>
<td>2.4</td>
<td>1.4</td>
<td>–</td>
</tr>
<tr>
<td>Crotaphytus bicintores</td>
<td>1.0</td>
<td>1.0</td>
<td>–</td>
</tr>
<tr>
<td>Crotalus micchelli</td>
<td>5.7</td>
<td>0.5</td>
<td>Westward</td>
</tr>
<tr>
<td>Crotalus scutulatus</td>
<td>0.6</td>
<td>1.4</td>
<td>–</td>
</tr>
<tr>
<td>Lichanura trivirgata</td>
<td>0.5</td>
<td>1.3</td>
<td>–</td>
</tr>
<tr>
<td>Gopherus agassizii/morafkai</td>
<td>0.6</td>
<td>0.5</td>
<td>–</td>
</tr>
<tr>
<td>Masticophis flagellum</td>
<td>3.7</td>
<td>9.9</td>
<td>–</td>
</tr>
<tr>
<td>Phrynosoma platyrhinus</td>
<td>12.3</td>
<td>10.7</td>
<td>–</td>
</tr>
<tr>
<td>Sceloporus magister</td>
<td>1.6</td>
<td>5.9</td>
<td>Eastward</td>
</tr>
<tr>
<td>Uma notata</td>
<td>0.6</td>
<td>1.2</td>
<td>–</td>
</tr>
<tr>
<td>Xantusia vigilis</td>
<td>3.8</td>
<td>1.9</td>
<td>–</td>
</tr>
<tr>
<td><strong>Mammals</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chaetodipus penicillatus</td>
<td>0.7</td>
<td>33.4</td>
<td>Eastward</td>
</tr>
<tr>
<td>Dipodomys deserti</td>
<td>1.2</td>
<td>4.2</td>
<td>Eastward</td>
</tr>
<tr>
<td>Dipodomys merriami</td>
<td>5.7</td>
<td>217.7</td>
<td>Eastward</td>
</tr>
<tr>
<td>Ovis canadensis</td>
<td>14.6</td>
<td>15.6</td>
<td>–</td>
</tr>
<tr>
<td>Xeropherophilus tereticaudus a,b</td>
<td>0.9</td>
<td>16.1</td>
<td>Eastward</td>
</tr>
<tr>
<td><strong>Arachnids</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aphonopelma prenticei</td>
<td>1.1</td>
<td>0.8</td>
<td>–</td>
</tr>
<tr>
<td>Homalonychus spp.</td>
<td>1.1</td>
<td>1.0</td>
<td>–</td>
</tr>
<tr>
<td>Hadrurus arizonensis</td>
<td>2.5</td>
<td>13.0</td>
<td>Eastward</td>
</tr>
<tr>
<td>Paravaeovis confuses/wearingi</td>
<td>3.0</td>
<td>2.3</td>
<td>–</td>
</tr>
<tr>
<td>Smeringurus meseensis</td>
<td>9.9</td>
<td>1.7</td>
<td>Westward</td>
</tr>
<tr>
<td>Smeringurus vachoni</td>
<td>1.2</td>
<td>2.5</td>
<td>–</td>
</tr>
<tr>
<td><strong>Insects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pogonomymex rugosus</td>
<td>1.7</td>
<td>8.7</td>
<td>Eastward</td>
</tr>
<tr>
<td>Prodoxus sordidus</td>
<td>126.2</td>
<td>1.0</td>
<td>Westward</td>
</tr>
<tr>
<td>Prodoxus weethumpf</td>
<td>340.4</td>
<td>32.0</td>
<td>Westward</td>
</tr>
<tr>
<td>Sphaerophalma arot</td>
<td>0.9</td>
<td>0.5</td>
<td>–</td>
</tr>
<tr>
<td>Tegeticula antithetica</td>
<td>15.0</td>
<td>4.2</td>
<td>Westward</td>
</tr>
</tbody>
</table>

Note: Bias is determined by averaging the number of effective migrants per species and calculating whether directional estimates deviate by at least 25% (dash indicates unbiased). Migration estimates that are low (Nm ≤ 1 after rounding to the nearest whole number) are bolded.

a Small sample size, likely underpowered.
b Eastern population is only from Sonora, MX.
c Very high numbers may be due to panmixia or low sample size.

Time (e.g., bighorn sheep, desert tortoises) than short-lived species (e.g., mice, lizards). We also expect that putative modes of cross-river dispersal would be unequal across species—highly dispersive species, those that can swim or survive periods underwater, as well as ecological generalists would be expected to have higher rates of cross-river mixing than their counterparts.

We should note that our meta-analysis overlooks some species-specific nuances. For example, Crotaphytus collaris (not included here) has been shown to introgress into the C. bicintores lineage (included here) in southwestern Arizona (McGuire et al., 2007). Such species-specific attributes would contribute to interspecific variance but should (in theory) be stochastic relative to the cross-species patterns shown here.

5.1.2 Pseudocongruence and modes of divergence

There is considerable interest in documenting and quantifying how changes in the physical landscape promote speciation, and which extrinsic processes are most impactful (Antonelli, Ariza, et al., 2018; Antonelli, Kissling, et al., 2018; Dolby et al., 2015, 2018; Hoorn et al., 2010; Jacobs, Haney, & Louie, 2004). The timing at which isolating events occur is important because it can be difficult to disentangle the biological impacts of co-occurring physical processes. We note that flooding of the Gulf of California, onset (or strengthening) of the palaeo-monsoon, and initiation of the Colorado River into the Gulf all occurred at similar times (Figure 3). The fact that it is unlikely to differentiate between these events using genetic divergence ages suggests the events can be considered biologically synchronous. While this co-occurrence is challenging, it is also an opportunity as the two processes would exert appreciably different evolutionary effects.

Southwestern divergence has largely been attributed to allopatric vicariance (but see Wood et al., 2012). In our meta-analysis, the Colorado River explains some of this divergence (Figures 5 and 7), but relatedness and migration metrics (Table 3, Figures 4 and 5) indicate that either the gene lineages studied here have not fully sorted despite several million years of complete isolation, or there has been a nontrivial amount of post-initiation gene flow that has occurred. If the river has been a leaky barrier, then the observed divergence would have actually occurred with reduced gene flow between adjacent populations (i.e., in parapatry as defined by Coyne & Orr, 2004, sensu stricto). Assuming E-W individuals can be considered discrete populations, then the stepping-stone model of parapatric divergence may fit (Coyne & Orr, 2004) in which populations diverge with reduced gene flow in the face of differential adaptation to environmental differences (we suggest seasonal precipitation disparity; Figure S1). Many of our tests for neutrality were rejected, suggesting that there is either a common pattern of demographical change or selection in the species studied here, but this cannot be resolved with available data (Table 1). As with sympatric speciation, the prospect of parapatric divergence is controversial (Bolnick & Fitzpatrick, 2007; Fitzpatrick, Fordyce, & Gavilates, 2009) as it requires divergence to overcome the influence of gene flow and there are challenges in measuring and documenting this phenomenon in natural populations (Coyne, 2007). That said, we suggest that the lower Colorado River region is an intriguing natural laboratory in which to test such a hypothesis because knowledge of the external factors affecting divergence can be well-constrained.
Coyne and Orr (2004) note that explanations of parapatric divergence require knowledge that the differential selection pressure(s) existed for the duration of divergence. Geological evidence shows that the NAMS has been in effect since at least 6.3 Ma and its influence on precipitation disparity between Mojave and Sonoran deserts has been operating during the period when 20 of the 25 species diverged. Precipitation asynchrony has been shown to drive genetic divergence in New World birds (Quintero, González-Caro, Zalamea, & Cadena, 2014), and Mojave-Sonoran precipitation asynchrony could initiate differences in behaviour, mating patterns, energy balance and osmotic regulation (Barrows, 2011; Davis & DeNardo, 2007, 2010). For example, the Mojave and Sonoran desert tortoises exhibit differences in timing of activity, foraging and reproduction that coincide with seasonal precipitation differences (Esque, Drake, & Nussar, 2014). These species also hybridize at the Mojave-Sonoran ecotone east of the Mohave Valley (Figure 8b) where precipitation patterns are intermediate (Figure S1). In this contact zone the hybrids show an intermediate habitat preference to the two parent species (Edwards et al., 2015). The Ambrosia shrub (not studied here) has different drought mortality in the Mojave versus Sonoran
deserts that is best explained by a natal precipitation index, also suggesting adaptation to different precipitation regimes (McAuliffe & Hamerlynk, 2010). Finally, Bowers (2005) found that there were some phenological differences between how Mojave and Sonoran wildflowers responded to El Niño events, and animals have diverged based on phenological differences in other systems (Thomassen, Freedman, Brown, Buermann, & Jacobs, 2013). These examples may represent differential adaptation and/or prezygotic isolation barriers that contribute to reinforcement, a process critical to developing "good" species. Similar work on the Cochise Filter Barrier and Baja California peninsula both emphasized the importance of ecological factors in limiting species ranges and reinforcing divergence (Dolby et al., 2015; Myers et al., 2017).

Differentiating between parapatric speciation and allopatric speciation with secondary contact is notoriously difficult (Coyné & Orr, 2004). The genetic data in this study cannot adjudicate between these scenarios, but future studies using whole-genome sequencing have promise. Methods such as ≥α = (Gutenkunst, Hernandez, Williamson, & Bustamante, 2009) and ≥IM (Costa & Wilkinson-Herbots, 2017) can be used to assess the fit of speciation models, such as speciation in isolation, with initial migration or with continuous migration. Methods to differentiate between incomplete lineage sorting and post-vicariance admixture can be achieved by statistically comparing consistency of tree topologies (e.g. GSI statistic, Cummings, Neel, & Shaw, 2008). Analysis of neutral regions to estimate demographical parameters in concert with examining genes for signatures of differential selection would reveal the degree to which local adaptation has occurred and whether monsoon-driven precipitation differences are a likely explanation. Finally, using redundancy analysis (RDA; Legendre & Fortin, 2010), distance-based RDA (DBRDA; Legendre & Anderson, 1999; e.g., Kierępka and Latch, 2015), and partial RDA (pRDA; Borcard et al., 2011) with genomic data can help disentangle the importance of ecological/climatic factors from geographical distance. Methods that directly quantify the relative amount of adaptive versus neutral divergence will become more important as researchers seek to determine the relative contribution of specific geo-climatic factors to lineage divergence (e.g. BEDASSLE which uses genetic and ecological covariance matrices, Bradburd, Ralph, & Coop, 2013). The next methodological frontier will be to develop metrics that achieve this in a way that can be compared across different geographical and taxonomic systems. This is necessary in order to determine what global processes most generate new species both today and over deep time (e.g., Antonelli, Ariza, et al., 2018; Hoorn et al., 2010), and resolve how these patterns vary by taxonomic group. A simple way to achieve this in the lower Colorado River region is to take advantage of geography. Because the Mojave-Sonoran ecotone and Colorado River are not coincident, dense sampling in future studies could triangulate "neutral" (river-based) and adaptive (monsoon-based) divergence by strategically sampling populations with Mojave-style precipitation east of the river and Sonoran-style precipitation west of the river (Figure S1). This strategy makes a time-for-space assumption that the populations have evolved in situ and abiotic conditions have not majorly changed, which may be unrealistic. Either way, through ongoing geological and biological work the Southwest is an emerging model setting to explore evolutionary pseudocongruence (different processes producing similar effects; Figure S2) and embrace intermediate levels of ecological-biological (i.e. geobiological) complexity.

### 5.2 Modes of post-initiation dispersal

Tree reconstructions of simulated data show that low-level or pulsed gene flow can produce the effect of nested paraphyly observed in the empirical trees (Figure 4b; Figures S4 and S7), but they do not give insight as to which mode is more likely. The Dmax statistic reveals that trees from simulated data have slightly less geographical mixing than the empirical data (but not significantly so)—this provides a useful benchmark to suggest that on average the empirical datasets may result from more cross-river gene flow than we modelled in our simulations. Whole genome sequencing should better resolve the amount of admixture, its age, and whether key genomic regions are reinforcing divergence in the face of gene flow (Cruckshank & Hahn, 2014; Feder, Flaxman, Egan, Comeault, & Nosil, 2013; Guerrero & Hahn, 2017; Nosil & Feder, 2011; Roda, Mendes, Hahn, & Hopkins, 2017; Wang & Hahn, 2018). We evaluate possible modes of cross-river dispersal below to help guide future studies.

#### 5.2.1 River avulsions and lava dams

Channel avulsions as a means of cross-river dispersal is not new (Graham et al., 2017; Jezkova et al., 2009; Mulcahy et al., 2006). Viewed on small temporal and spatial scales, a single avulsion has the potential to facilitate a pulse of unidirectional gene flow (Figure 9). Aggregated over space (i.e. the length of the river) and time (thousands to millions of years), avulsions result in low levels of bi-directional gene flow. Avulsion is a primary process of aggrading river floodplains that may occur at frequencies ranging from c. 50 to 1,400 years or longer (e.g., Slingerland & Smith, 2004; Stouthamer & Berendsen, 2001, 2007). The Rhine-Weisse delta in the Netherlands is a model system for studying avulsions; authors there have measured 91 avulsions over the Holocene alone (Stouthamer & Berendsen, 2007). In our context, cross-river migration by any single avulsion would be a rare event but aggregated over the total number of events would yield non-trivial opportunity for cross-river dispersal, and these can be weighted by species-specific characteristics.

It is difficult to distinguish cross-river dispersal by avulsions from dispersal in the absence of avulsions (the genetic consequences are equivalent). Empirical evidence for avulsion-mediated transport would come from aquaphobic (water-fearing) species showing cross-river dispersal, or spatial concordance between areas of extralimital populations and places where the floodplain may be prone to avulsion, such as wide areas. We expect a single channel avulsion (like any dispersal event) would manifest a clear phyleogeographical pattern (i.e. Figure 9). Aggregated over space and time this could result
FIGURE 9  Schematic showing how a westward channel avulsion could move individual(s) eastward of the river and leave a genetic signature. This is an example of a local, singular event. Extrapolated over the length of the river and over time, this process would result in low levels of bidirectional gene flow. In theory, an individual avulsion would leave a strong, observable phylogeographical signal for a short period after the event, but this signal would be difficult to distinguish from a scenario where dispersal is localized to a specific part of the river simply because that section is particularly passable (i.e. localized dispersal unaided by avulsions) [Colour figure can be viewed at wileyonlinelibrary.com]

in nested paraphyly, or greater genetic similarity among cross-river clades than expected under strict allopatry.

Our mapping of extralimital populations in this study reveals a strong pattern of “western” genotypes located just east of the river near the Mohave Valley. Nine species have cross-river extralimital populations in this general area. Our mapping of published data (Beard et al., 2011; Crow, Block, et al., 2018) shows the Mohave Valley region has a wide palaeo-floodplain (Figure 8b) and may be an area that has been particularly passable or is a region of common (in deep time) and/or recent river avulsions. Howard et al. (2008) described evidence for at least four Pliocene and Pleistocene avulsions in this area. A westward shift of the main channel in the Mohave Valley would explain these extralimital populations. The localized nature of this multi-species pattern suggests that it could be the result of a series of recent dispersal events, avulsion-mediated or not, although a single avulsion is a parsimonious explanation. The Basin and Range-style topography of N-S trending mountain ranges could limit further dispersal eastward. This area also coincides with the Mojave-Sonoran ecotone and could alternatively reflect a hybrid zone where differentially adapted populations come into contact and are limited ecologically. Differentiating between primary divergence and secondary contact could be achieved through detailed analysis of how patterns of neutral and adaptive alleles change across this zone (Coyne & Orr, 2004). We detect very few extralimital populations west of the river. Our historical gene flow results (Table 3), which are expected to reflect a long-term historical signal, show no aggregate directionality.

River avulsions can be considered an opportunistic dispersal event and species living proximal to the river channel would be more likely to translocate during such events. For instance, the desert pocket mouse Chaetodipsus penicillatus inhabits sandy slopes near the river and shows very high levels of cross-river introgression (Jezkova et al., 2009). In contrast, highly philopatric or specialized organisms, such as lizards adapted to live in rock crevices (Leavitt, Bezy, Crandall, & Sites, 2007) or on aeolian-derived dunes (Gottscho, Marks, & Jennings, 2014) may be less likely to disperse. Same is true for sparsely distributed species; for example, River Island State Park near Parker, AZ sits at the heart of a small Pleistocene avulsion that affected an area less than 1.5 km² (see Figure 5c in Howard et al., 2008).

Lava dams are known to have occurred over the last 0.8 Ma towards the northeast of our study area. Though ephemeral, these dams could have mediated cross-river dispersal as short-lived land bridges, or by reducing flow downstream of the dam, rendering the river more passable. The precise Ar-Ar dates can be used to parameterize evolutionary models (as done here) to directly compare how well lava dams versus other historical scenarios explain genetic diversity patterns. Assuming model parameters can be reasonably known (generation time, mutation rate), the coalescent age for cross-river clades should be younger when dispersal was mediated by lava dams rather than avulsions or stochastic dispersal events (which would occur over the river’s life span).

5.2.2  Changes in river flow

Rafting of terrestrial fauna on storm debris has long been considered an important method of long distance dispersal, particularly for colonizing distant oceanic islands (Bell et al., 2015; Gillespie et al., 2012; Queiroz, 2005), and has also been implicated as a means of cross-river dispersal (Grinnell, 1914; Hedin, Starrett, & Hayashi, 2013). Rafting is more likely to occur during floods when river competency is higher and able to transport larger debris. River discharge is often seasonal and tied to rainfall patterns or the outflow of
glacial meltwater (Morehead, Syvitski, Hutton, & Peckham, 2003). These changes can modify habitat communities (Rosado, Morais, & Tockner, 2014) and many species have evolved to adapt to changes in flow (Boedeltje, Bakker, Brinke, Groenendaal, & Soesbergen, 2004; Lylte & Poff, 2004). Given its large watershed, the Colorado River likely channelled high volumes of meltwater from the Rocky Mountains (Miller, Susong, Shope, Heilweil, & Stolp, 2014) as high-elevation montane glaciers retreated at the end of major glacial cycles over the past c. 0.8 Myr (Figure 3), and should be considered a viable cross-river dispersal hypothesis (Higgins, Nathan, & Cain, 2003).

Prior to damming, the Colorado River experienced large variations in water discharge due to its arid setting (e.g., Minckley, 1979). As such, cross-river dispersal could also have taken place during periods of low river discharge. During dry periods, the river held little suspended sediment and formed wide regions of braided streams (Carlson & Muth, 1989), which may be easier for some animals to disperse across than a flowing channel. Changes in flow could also feed back onto other processes, such as channel avulsions. Some animals, such as scorpions, can survive underwater for considerable durations (M. R. Graham, personal observation) and could opportunistically disperse across the river during periods of reduced flow. Historical photographs record river behaviour prior to damming and reveal crossable regions that were narrow and shallow (Unknown, 1920b, 1930), as well as regions that would never be crossable even during periods of reduced flow (Huber, 1925; Unknown, 1920a, 1931).

5.3 Tributaries and biogeographical patterns

Tributaries of the Colorado River may also act as migration barriers. In Arizona, the Bill Williams and Gila rivers feed into the Colorado at approximately Parker and Yuma, respectively, and both are sites of genetic divergence. The Gila River marks either the geographical limit or site of genetic divergence in lizards (Crotaphytus nebris, McGuire et al., 2007; Uma rufopunctatus, U. cowlesi, and U. sp., Gottscho et al., 2017), snakes (Lichanura trivirgata, Wood, Fisher, & Reeder, 2008), rodents (Chaetopidus penicillatus, Jezkova et al., 2009) and scorpions (haplotype groups 5 and 6 of Hadrurus arizonensis, Graham et al., 2013a). The Bill Williams river marks a similar pattern for snakes (Chionactis annulatus and C. occipitalis, Wood et al., 2014), toads (haplogroups within Anaxyrus punctatus, Jaeger et al., 2005), lizards (Scloporus magister, S. “uniformis”, Leaché & Mulcahy, 2007; Uma spp, Gottscho et al., 2017), mammals (Nelson and Mexican groups of Ovis canadensis, Buchalski et al., 2016) and arachnids (haplotype group 3 of Aphonopelma mojave, Graham et al., 2015). Geomorphic studies show that the Colorado and Gila rivers experienced increased discharge and gravel transport during Pleistocene glacial periods of higher rainfall, and channel deposits were incised by erosion during the drier interglacial periods (Anders et al., 2005; Chadwick, Hall, & Phillips, 1997; Sharp, Ludwig, Chadwick, Amundson, & Glaser, 2003; Waters, 2008; Waters & Haynes, 2001). We suggest that high water discharge in the Bill Williams and Gila rivers during glacial periods may have created barriers to migration that could explain the observed genetic divergence associated with these tributaries and reflect additional, subtler ways that this palaeo-landscape has changed. The Bill Williams River happens to also coincide with the Mojave-Sonoran ecotone (Figure 1; Figures S1 and S2), so this is another setting where evolutionary pseudocongruence may be occurring in which neutral and adaptive forces are at work.

6 CONCLUSIONS

Our meta-analysis of 33 species suggests that there is quantifiable divergence between populations east and west of the Colorado River. The river, however, is best characterized as a leaky barrier that filters rather than prevents gene flow. Analysis of gene tree topologies and variance decomposition, forward simulations, haplotype networks and historical migration suggests that there has been cross-river gene flow in many species since initiation of the Colorado River into the Gulf of California, perhaps aided by lava dams and channel avulsions. The age differences in the two models of river initiation are unlikely to impact divergence assessments, but this uncertainty should be accounted for in genetic models. The region east of the Mohave Valley hosts a large number of extralimital populations and may be a location where the river channel is most passable. Considering the river to be a leaky barrier reframes our understanding of southwestern divergence—raising the possibility that divergence occurred in parapatry and may be aided by differential adaptation to monsoon precipitation across the Mojave-Sonoran ecotone in the face of ongoing gene flow. While this may be controversial, the role of ecological adaptation should be tested with whole genome approaches in future studies (Johannesson, 2010). The genetic patterns and detailed geo-climatic knowledge available in the Southwest makes it an excellent setting to explore intermediate levels of geobiological complexity where measuring different drivers of divergence may be plausible across a suite of desert-adapted taxa.

ACKNOWLEDGEMENTS

We thank three anonymous reviewers and K Kusumi, P Peathree, K House, R Crow, S Beard, SEK Bennett, M Darin, D DeNardo, S Pratt, R Johnson and K Whipple for sharing data and/or fruitful conversations. We thank R Ayrey, S Couceiro, P Cushing, G Graham, R Jones, R Kaiser, WB Mattingly, M Pinto and D. Wood for assistance with fieldwork, laboratory procedures and permits.

CONFLICT OF INTEREST

We have no conflicts to declare.

DATA AVAILABILITY STATEMENT

Input files, control files and raw output files can be found at Harvard Dataverse: https://dataverse.harvard.edu/dataverse/GADLCRR. Accession numbers are in Tables S4 and S7.
REFERENCES


Dorsey, R. J., Housen, B. A., Janecke, S. U., Fanning, C. M., & Spears, A. L. F. (2011). Stratigraphic record of basin development within the San Andreas fault system: Late Cenozoic Fish Creek-Vallecito basin,


BIOSKETCH

Our research team studies how species evolve with their landscapes through time. We integrate geologic and genetic data to account for geo-climatic complexity of the palaeo-landscape to better understand the forces that drive genetic divergence, the means by which divergence occurs, and the variability in how species respond to these forces. We all share an interest in the natural history of North American deserts.

Author contributions: GAD, RJD and MRG co-conceived of the project and analysed the data. MRG collected the published data and generated the unpublished data; GAD and MRG conducted genetic analyses. GAD and RJD drafted the manuscript and figures. All authors revised the manuscript.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Dolby GA, Dorsey RJ, Graham MR. A legacy of geo-climatic complexity and genetic divergence along the lower Colorado River: Insights from the geological record and 33 desert-adapted animals. J Biogeogr. 2019;00:1–27. https://doi.org/10.1111/jbi.13685

APPENDIX 1

ADDITIONAL DATA SOURCES


