MOLECULAR ECOLOGY

Molecular Ecology (2013) **22**, 508–526 doi: 10.1111/mec.12130

Crossing the uncrossable: novel trans-valley biogeographic patterns revealed in the genetic history of low-dispersal mygalomorph spiders (Antrodiaetidae, *Antrodiaetus*) from California

MARSHAL HEDIN,* JAMES STARRETT† and CHERYL HAYASHI†

*Department of Biology, San Diego State University San Diego, San Diego, CA 92182-4614, USA, †Department of Biology, University of California, Riverside, CA 92521, USA

Abstract

***Antrodiaetus riversi* is a dispersal-limited, habitat-specialized mygalomorph spider species endemic to mesic woodlands of northern and central California. Here, we build upon prior phylogeographic research using a much larger geographic sample and include additional nuclear genes, providing more detailed biogeographic insights throughout the range of this complex. Of particular interest is the uncovering of unexpected and replicated trans-valley biogeographic patterns, where in two separate genetic clades western haplotypes in the California south Coast Ranges are phylogenetically closely related to eastern haplotypes from central and northern Sierran foothills. In both instances, these trans-valley phylogenetic patterns are strongly supported by multiple genes. These western and eastern populations are currently separated by the Central Valley, a well-recognized modern-day and historical biogeographic barrier in California. For one clade, the directionality is clearly northeast to southwest, and all available evidence is consistent with a jump dispersal event estimated at 1.2–1.3 Ma. During this time period, paleogeographic data indicate that northern Sierran rivers emptied to the ocean in the south Coast Ranges, rather than at the San Francisco Bay. For the other trans-valley clade genetic evidence is less conclusive regarding the mechanism and directionality of biogeographic exchange, although the estimated timeframe is similar (approximately 1.8 Ma). Despite the large number of biogeographic studies previously conducted in central California, to the best of our knowledge no prior studies have discussed or revealed a northern Sierran to south Coast Range biogeographic connection. This uniqueness may reflect the low-dispersal biology of mygalomorph spiders, where 'post-event' gene exchange rarely erases historical biogeographic signal.**

Keywords: biogeography, California, cryptic speciation, isolation by distance, long-distance dispersal, San Joaquin marine embayment

Received 22 June 2012; revision received 12 October 2012; accepted 17 October 2012

Introduction

California is home to a rich biota, reflecting a dynamic and complex biogeographic history. This biota includes some ancient taxa, with paleoendemic genera showing biogeographic relationships to taxa from other continents (e.g. van der Meijden *et al.* 2009). At the same time, evolutionarily recent (e.g. Pleistocene) or ongoing (e.g. plate movements) processes have shaped or are currently shaping evolutionary divergence. Spatially, divergence can occur over small geographic distances, reflecting sharp environmental gradients or crisp geological boundaries (e.g. geographically adjacent exotic terranes). The richness and complexity of California has long fascinated historical biogeographers, and California is indeed one of the most biogeographically well-studied areas on Earth. This knowledge is derived from

Correspondence: Marshal Hedin, Fax: 1 619 594 5676;
E-mail: mhedin@mail.sdsu.edu

both detailed studies of individual taxa (e.g. *Ensatina* salamanders – Wake 1997; Moritz *et al.* 1992; Kuchta *et al.* 2009a,b), combined with many comparative meta-analyses (Calsbeek *et al.* 2003; Lapointe & Rissler 2005; Feldman & Spicer 2006; Rissler *et al.* 2006; Chatzimanolis & Caterino 2007; Davis *et al.* 2008; Kraft *et al.* 2010). Beyond revealing biogeographic patterns, these studies have provided general lessons in evolutionary biology, for example regarding species delimitation (Bond & Stockman 2008; Wake 2009), speciation (Pereira & Wake 2009; Satler *et al.* 2011) and the conservation of evolutionary processes (Bond *et al.* 2006; Davis *et al.* 2008; Vandergast *et al.* 2008).

Inspired by their research on salamanders (*Batrachoseps*), Martinez-Solano *et al.* (2007) argued that the study of dispersal-limited taxa might prove particularly useful in understanding Californian historical biogeography. As for most taxa, barriers to dispersal promote evolutionary divergence – such barriers might be biotic (i.e. competition with other taxa), geologic, geographic or environmental. What is atypical and special about extreme low vagility systems is that this divergence can be remarkably fine-grained spatially and can extend to very shallow timescales. Also, since post-barrier dispersal is low or nonexistent, evidence for biogeographic barriers is retained longer in these systems. As such, a multitude of barriers to dispersal can potentially be revealed, and these may vary in terms of strength of barrier (i.e. both long term and short term), geographic scale and timing. We argue that on a *per taxon* basis, dispersal-limited species provide more biogeographic information. Vertebrate examples in California include salamanders (e.g. Jockusch *et al.* 2001; Jockusch & Wake 2002; Kuchta *et al.* 2009a,b; Martinez-Solano & Lawson 2009; Rovito 2010), lizards (Leavitt *et al.* 2007; Parham & Papenfuss 2009) and fossorial rodents (Patton & Smith 1990). This study concerns another major group of Californian dispersal-limited animals, the mygalomorph spiders.

Mygalomorph spiders comprise a distinct clade within the order Araneae and include spiders such as tarantulas, trapdoor spiders and purseweb spiders. Many members of this ancient clade are notoriously dispersal-limited, with evidence for constraints on dispersal seen at many taxonomic levels [e.g. ancient vicariance promoted by continental drift (Griswold & Ledford 2001), short-range endemicity within continents (Bond 2012)]. Limited dispersal has been revealed in multiple genetic studies, which overwhelmingly show patterns of genetic fragmentation and complete population subdivision arising at localized spatial scales (e.g. Hendrixson & Bond 2005; Arnedo & Ferrandez 2007; Cooper *et al.* 2011). In California, fine-scale genetic fragmentation has been observed in multiple genera, including *Apomastus* (Bond 2004), *Promyrmekiaphila* (Stockman & Bond 2007), *Aptostichus* (Bond *et al.* 2001; Bond & Stockman 2008) and *Aliatypus* (Hedin & Carlson 2011; Satler *et al.* 2011).

Antrodiaetus riversi (O.P.–Cambridge 1883) is an antrodiaetid mygalomorph species endemic to California. This taxon occupies a disjunct distribution in California, with groups of populations found in both the central Sierra Nevada and in the Coast Ranges (Fig. 1). These spiders are moisture-sensitive microhabitat specialists, typically restricted to north-facing woodlands (Coyle 1971). In such microhabitats, these spiders live in silk-lined subterranean burrows and build silken entrance constructs that resemble small 'turrets' (Coyle 1971; Vincent 1993). Previous phylogeographic research by Starrett & Hedin (2007) investigated the genetic structuring of *A. riversi* using nuclear and mitochondrial DNA sequences, collected for 32 locations from the then known distribution. These data revealed extreme population subdivision, deep phylogeographic structuring and strict allopatry of genealogically congruent clades that likely represent cryptic species. Here, we build upon this research using a much larger geographic sample (133 sites sampled), for which we have gathered DNA sequences for a mitochondrial and nuclear gene; data for two additional nuclear genes were gathered for a representative subset of geographic samples. These data allow us to better characterize the geography of genetic clade divergence and population genetic structuring within genetic clades.

Of particular relevance to California biogeography is the uncovering of unexpected and replicated trans-valley biogeographic patterns in *A. riversi*, where in two separate genetic clades western haplotypes in the south Coast Ranges are phylogenetically closely related to eastern haplotypes from central and northern Sierran foothills. In both instances, these trans-valley phylogenetic patterns are strongly supported by multiple genes. These western and eastern populations are currently separated by the Central Valley, which represents a mostly flat, highly converted environment (Schoenherr 1992; Fig. 1). The Central Valley is arguably the most well-recognized modern-day and historical biogeographic barrier in California, with many upland taxa distributed in a complete or partial ring around inhospitable habitats of this valley (reviewed in Calsbeek *et al.* 2003; Lapointe & Rissler 2005; Feldman & Spicer 2006; Rissler *et al.* 2006).

Our primary hypothesis is that low-dispersal *Antrodiaetus* spiders actively or passively dispersed across the Central Valley barrier during a 2.2–0.8 Ma time window. Paleogeographic data indicate that northern Sierran rivers emptied to the ocean in the south Coast Ranges, rather than at the San Francisco Bay, during

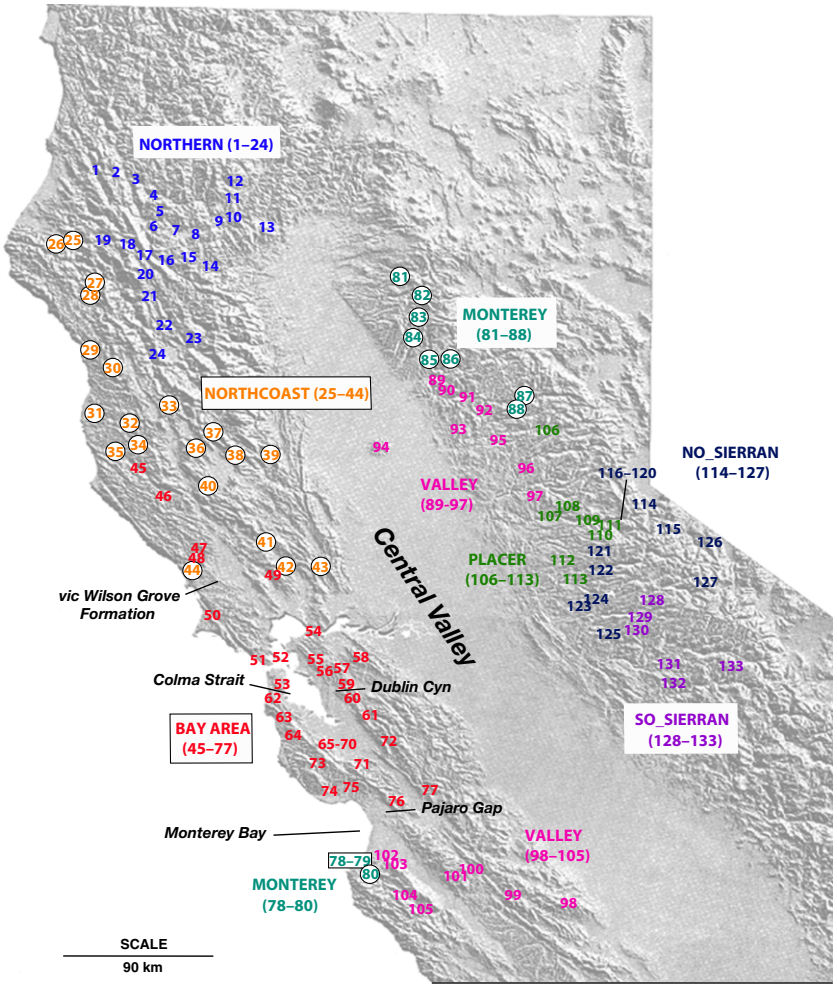

Fig. 1 Map of *Antrodiaetus riversi* collection sites in central and northern California. Site numbers listed in Supplemental Table 1. Sites are grouped into genetic clades, following results of phylogenetic analyses [Figs 2, 3 and 6, Fig. S1 (Supporting information)]. Geographic features mentioned in text are highlighted.

this time interval. We tested the spatial prediction of this hypothesis, that is, a northeast-to-southwest directionality consistent with the flow directionality of paleorivers, using multigenic phylogenetic evidence. Furthermore, we tested the temporal prediction of this hypothesis using relaxed molecular clock methods. Finally, we used coalescent and other population genetic analyses to evaluate whether trans-valley dynamics are more consistent with passive long-distance dispersal or active continuous range expansion. A comprehensive review of available literature suggests that the trans-valley patterns observed in these spiders are novel, despite the large number of biogeographic studies previously conducted in central California.

Materials and methods

Geographic sampling

Specimens were collected from 133 sites (Fig. 1, Table S1, Supporting information); at each site, a series of individuals was collected, consisting mostly of adult females, but sometimes also including immature spiders. All specimens have been assigned a unique specimen identification number (Table S1). Upon completion of our ongoing studies, representative voucher specimens will be deposited in the California Academy of Sciences, San Francisco, California. Our sample provides a comprehensive coverage of the known range of *A. riversi*, and our fieldwork led to the discovery of new northern and southern records for this species, both in the Coast Ranges and the Sierra Nevada. The species does not likely occur further north in California, where congeneric taxa appear to replace *A. riversi* (Coyle 1968). We have also searched extensively in the eastern foothills of the southern Sierra Nevada and have not found *A. riversi* populations.

Monophyly of the *A. riversi* complex was strongly supported in prior analyses using multiple outgroup taxa (Starrett & Hedin 2007). In this prior study, we identified a clade of northwestern CA samples (the 'Northern' group, see Results), which is the likely sister

group to all remaining populations. Because our emphasis in this study is not on root position, we have not included divergent outgroup sequences, and instead root gene trees assuming that the Northern group is sister to remaining genetic clades.

Gene data collection, sequence alignment & phylogenetic analyses

Genomic DNAs were used as templates in PCR amplification of four separate gene regions, including one mitochondrial (CO1) and three nuclear genes (28S, EF1G, anonymous nuclear). DNA sequences were determined using the Sanger method, edited, and aligned both manually and algorithmically. Bayesian phylogenetic analyses were conducted on individual gene matrices using best-fit models of molecular evolution. Details of all methods used are presented in Supplemental Document 1.

Phylogenetic hypothesis testing

For some recovered clades (see Results), we assessed whether trees that included taxon bipartitions conforming to a particular phylogenetic hypothesis were found in the 95% credible set of trees (see Miller *et al.* 2002; Cranston & Rannala 2007). We first trimmed the post-burn-in MrBayes trprobs file to include only the 95% credible set, then filtered this tree block in Mesquite V 2.6 (Maddison & Maddison 2009), applying filters conforming to alternative topological hypotheses.

Population structure analyses

Several measures of genetic diversity (Table 1) were calculated using DnaSP v5 (Librado & Rozas 2009). We also examined the relationship between geographic distance and CO1 genetic distance within individual primary clades or subclades (see Results) using the IBDWS v3.15 web service (Jensen *et al.* 2005). Euclidean geographic distances (in kilometres) between georeferenced sampling sites were measured using the distance measurement tool in Google Maps, at a spatial resolution of 0.9 cm = 1 km. Genetic distances between sites (using a single randomly chosen haplotype per sampled site) were calculated in PAUP* (Swofford 2002) for the unpartitioned CO1 data. To minimize the effects of mutational saturation, we calculated genetic distances using a best-fit GTR + I + G model, with I + G parameter estimates derived from jModelTest. Matrix correlations were assessed using both a Mantel test and reduced major axis (RMA) regression, based on 10 000 randomizations of the input matrices.

The program Migrate-n 3.3.2 (Beerli & Palczewski 2010; Beerli 2012) was used to estimate mutation-scaled effective population size (theta) and effective migration rate (M = immigration rate/mutation rate) values for each of two different genetic clades (Monterey and Valley, see Results). For each clade, we assumed a two population model (east vs. west of Central Valley) and based parameter estimates on multilocus data with appropriate inheritance scalars (0.25 for CO1, 1.0 for nuclear genes). Transition/transversion ratio values were derived from jModeltest (from K80 model), and the mutation rate over all loci was derived using the 'Relative' command. In initial Bayesian analyses, we used a Metropolis–Hastings sampling proposal function and default uniform priors for both theta and M. Prior bounds for M were increased (0–1000) based on results of preliminary analyses. Markov chain settings included 50 000 recorded steps with a sampling increment of 100 and 10 concurrent chains, totalling 50 million sampled parameter values per run. Four chains included static heating with default start value temperatures and a swapping interval of 10. For all analyses, we conducted two separate runs and compared modal and mean parameter estimates to gauge the reliability of results. To test alternative gene flow models, we compared M values from a full model to a symmetric M ($M_{\text{west>east}} = M_{\text{east>west}}$) model. Log marginal likelihoods were derived using thermodynamic integration and a Bezier approximation method (Beerli & Palczewski 2010), and compared using Bayes factors to evaluate alternative models.

Divergence time analyses

We used the BEAST v1.7.2 package (Drummond & Rambaut 2007; Drummond *et al.* 2012) to estimate a time-calibrated phylogeny using a relaxed molecular clock model. BEAST analyses were based on a combined data matrix assembled from geographic locations for which we had specimen data for at least 3 genes. In all cases, sequences were generated from the same specimen or different specimens from the same site (Table S1). The same sequence models used in separate Bayesian analyses were used in BEAST (including CO1 codon partitioning), with substitution and clock models unlinked across data partitions, but with a linked tree model. All analyses were conducted using an uncorrelated lognormal-relaxed clock model (Drummond *et al.* 2006). Preliminary analyses were run using a Speciation: Birth–Death tree prior, with default priors for all other parameter values. Based on initial analyses resulting in low effective sample size (ESS) values for certain parameters, some priors were changed to uniform [0, 1e100]. Because our data includes a combination of intraspecific and interspecific sampling, the choice of an appropriate tree prior is not obvious. To assess the impact of alternative tree priors on tree

Table 1 Gene diversity statistics

		N	*S*	*h*	*H*d	*P*i	*k*	Fu & Li
GROUPS								
Northern	CO1	30	172	24	0.986 (0.0001)	0.084	60.33	
Northcoast		27	135	21	0.983 (0.0001)	0.074	44.54	
Bay Area		42	138	33	0.990 (0.0001)	0.051	33.94	
Placer		15	101	8	0.905 (0.0025)	0.051	37.00	
No_Sierran		25	78	17	0.967 (0.0004)	0.042	26.79	
So_Sierran		8	42	5	0.857 (0.0117)	0.025	18.64	
Valley		23	123	17	0.976 (0.0003)	0.043	29.88	$D = -0.369$ $F = -0.690$ ($P > 0.10$)
Valley_West		12	85	8	0.939 (0.0023)	0.038	28.91	$D = 0.049$ $F = -0.098$ ($P > 0.10$)
Monterey		16	125	11	0.958 (0.001)	0.070	51.46	
Mont_West		6	7	3	0.80 (0.0148)	0.004	3.73	
Valley	28S	22	12	10	0.827 (0.0055)	0.004	3.46	$D = -0.036$ $F = -0.358$ ($P > 0.10$)
Valley_West		13	5	3	0.513 (0.0206)	0.0025	2.13	$D = 1.246$ $F = 1.386$ ($P > 0.10$)
Monterey		16	12	7	0.77 (0.0070)	0.0048	4.15	
Mont_West		7	1	2	0.57 (0.0143)	0.001	0.57	
Valley	EF1G	10	3	4	0.644 (0.0231)	0.001	0.76	$D = -0.805$ $F = -0.962$ ($P > 0.10$)
Valley_West		4	1	2	0.667 (0.0417)	0.001	0.67	
Monterey		11	10	8	0.891 (0.0084)	0.004	2.62	
Mont_West		3	0	1	0	0	0	
Valley	anon	10	7	6	0.778 (0.0189)	0.006	2.11	$D = -1.284$ $F = -1.389$ ($P > 0.10$)
Valley_West		4	4	3	0.833 (0.0495)	0.005	2	
Monterey		10	8	4	0.778 (0.0082)	0.009	3.56	
Mont_West		3	0	1	0	0	0	

Clade names follow those defined in text and figures. *N* = no of sequences; *S* = no of segregating sites; *h* = no of haplotypes; *H*d = gene diversity (variance in parentheses); *P*i = nucleotide diversity; *k* = average number of nucleotide differences. All values calculated using DnaSP v5.

topology and divergence time estimates, all analyses were replicated using Speciation: Birth–Death Incomplete Sampling, Speciation:Yule Process, and Coalescent: Constant Size tree priors. For each tree prior setting, we conducted three replicate MCMC runs each for 50 million generations, sampled every 1000 generations, using the 'auto optimize' operators option. The consistency of parameter estimates across separate runs was assessed using Tracer v1.5 (Rambaut & Drummond 2007), and the results of separate runs were combined such that ESS values exceeded 200 for all parameters. Logcombiner was used to combine separate tree files (burnin = 5 million states), with a reduced resample frequency of 10 000. From this reduced tree sample TreeAnnotater was used to reconstruct a maximum clade credibility tree, visualized using FigTree v1.3.1 (Rambaut 2009).

The fossil spider *Cretacattyma raveni* from lower Cretaceous Mongolian deposits (approximately 110 Ma) has been placed in the family Antrodiaetidae (Eskov & Zonshtein 1990). However, because this fossil is much older than the *A. riversi* complex, and because phylogenetic analyses have not formally placed this fossil, we did not use fossil evidence to calibrate-relaxed molecular clock analyses. Also, because our research focuses on biogeography, we did not use biogeographic criteria in clock calibrations. Instead, for the CO1 data partition, we specified a clock rate based on a well-calibrated arthropod CO1 rate, using a normal prior with a ucld.

mean of 0.0169 ± 0.0019 (Papadopoulou *et al.* 2010; table 4). This rate is similar to a recent CO1 rate rigorously estimated for a clade of araneomorph spiders (ucld.mean of 0.0199, Bidegaray-Batista & Arnedo 2011). We also note that rigorously estimated rates of CO1 molecular evolution have not been reported for mygalomorph spiders.

Results

Data availability

Newly generated DNA sequences have been deposited to GenBank (www.ncbi.nlm.nih.gov/Genbank/), with accession numbers available in Table S1. BEAST xml files and a Google Earth kmz file (with geographic location information) have been deposited to the DRYAD data repository (doi:10.5061/dryad.121tq).

28 gene trees

28S sequences (aligned length of approximately 1270 bp) were gathered for 205 individuals representing all 133 sampling sites. Bayesian phylogenetic analyses of manual and manual + MUSCLE alignments result in very similar clade structure and support values for the primary clades discussed hereafter (data not shown); discussed below are results based on manual alignments. We define eight primary 28S clades (Fig. 2), all of which are well supported (posterior probability > 0.95). The geographic distribution of these clades is shown on Fig. 1, which reveals mostly contiguous ranges and complete allopatry of these gene tree clades. The 28S data are highly genetically structured within most primary clades. For example, in the Northern clade, there are no 28S haplotypes shared between sampling locations, suggesting limited male-based gene flow at the spatial scale sampled (Fig. 2). As further discussed below, there are exceptions to this pattern in some geographic regions for certain clades.

Two 28S gene tree clades (Monterey and Valley) have disjunct distributions across the Central Valley. Each of these trans-valley clades includes a set of eastern populations found in the western foothills of the Sierra Nevadas north of Sacramento, and a set of western populations in the south Coast Ranges near and east of Monterey (Fig. 1). 28S data for the Monterey clade suggests paraphyly of eastern populations relative to a monophyletic set of western populations (eastern 85 and 86 closely related to western 78–80; Fig. 2). Western populations carry less 28S genetic variation than in the Monterey clade as a whole (Table 1). For the Valley clade, 28S sequences from western populations (locations 98–105) do not together form a single subclade on the Bayesian majority rule consensus tree, and very few trees in the 95% credible set (17 of 74706 trees) imply monophyly of western Valley populations. Levels of 28S genetic variation are relatively high in western Valley spiders (Table 1).

CO1 gene trees

CO1 sequences (approximately 1000 bp) were gathered for 210 individuals representing all 133 sample locations. Partitioned Bayesian analysis of the CO1 data results in a tree structure very similar to the 28S gene tree data, recovering the same eight well-supported primary lineages, with two exceptions (Fig. 3). CO1 haplotypes from four peripheral Northcoast locations (locations 36, 37, 39, 43; Fig. 1) form a divergent mitochondrial subclade distantly related to other samples of the Northcoast clade. Similarly, CO1 haplotypes from four peripheral Monterey locations (locations 81–84; Fig. 1) form a divergent mitochondrial subclade. We searched the CO1 95% credible tree set (n = 96 973) for trees with a monophyletic Monterey clade – 3546 trees conforming to this hypothesis are found in this tree set.

Similar to 28S gene tree results, two mitochondrial clades – Monterey (subset) and Valley – have disjunct distributions across the Central Valley. Again, data for the Monterey (subset) clade suggest paraphyly of eastern populations relative to a monophyletic set of western populations (eastern 85 and 86 closely related to western 78–80; Fig. 3). For the Valley clade, eastern and western CO1 haplotypes are intermixed.

Levels of CO1 genetic variation within primary clades are quite high, consistent with extremely limited female-based gene exchange in these spiders. The average number of nucleotide differences (k) within clades ranges from approximately 19 to 60 (Table 1). CO1 genetic diversity in western Monterey populations is reduced, whereas CO1 genetic diversity in western Valley spiders is comparable to that observed in the clade as a whole (Table 1).

EF1G and anonymous gene trees

EF1G (850 bp) and anonymous locus (470 bp) data were gathered for 47 and 52 individuals, respectively. For individual sequences, we coded heterozygous sites using standard IUPAC ambiguity codes. All primary genetic clades (see above) are represented by multiple location samples, and we obtained sequences for almost all locations for the trans-valley Monterey and Valley clades. For EF1G, six of eight primary clades are recovered (Fig. S1, Supporting information). Placer and Northcoast clades are not recovered on the Bayesian majority rule consensus tree, but trees with such clades

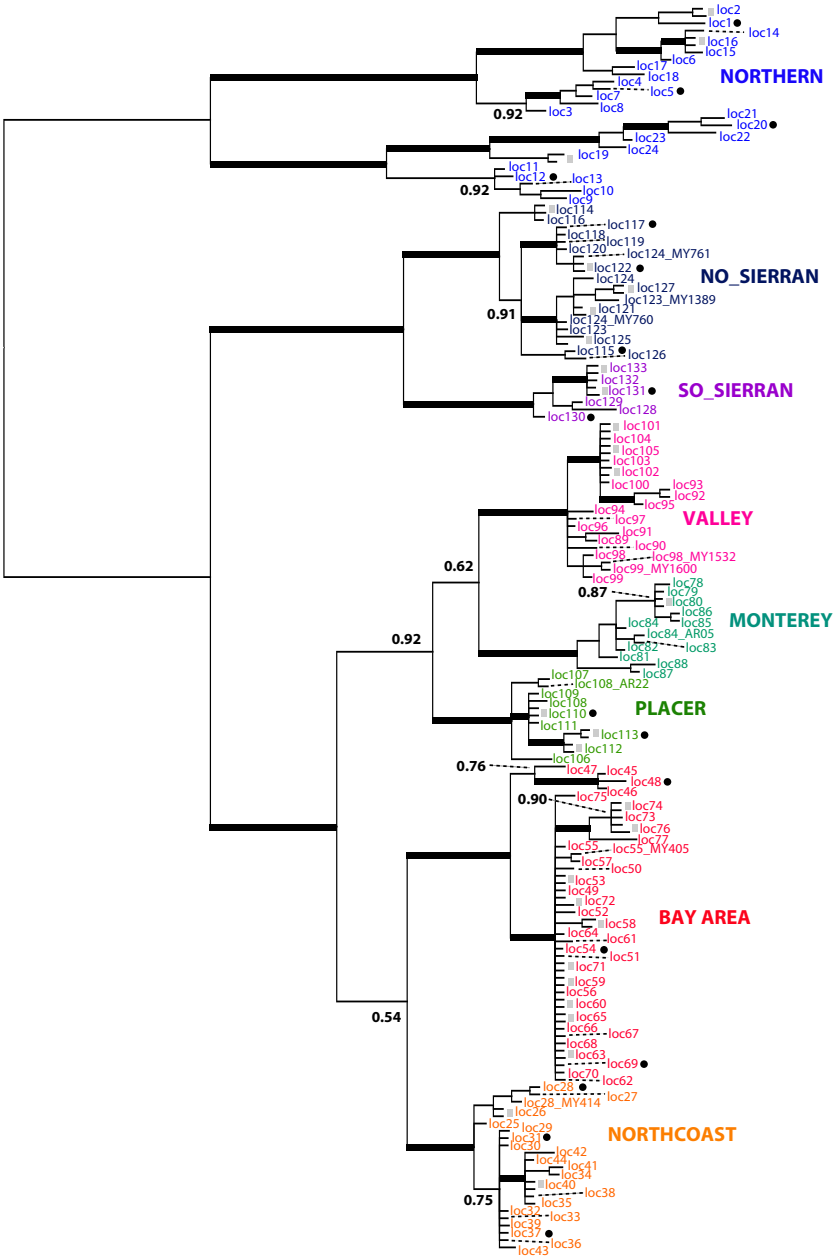

Fig. 2 28S manual alignment Bayesian consensus phylogram. Support values shown only for clades including haplotypes from three or more sampling locations, with strong support (posterior probability > 0.95) indicated by thickened branches. Support values for other relevant bipartitions indicated by exact posterior probability values. Multiple haplotypes sampled from a single location indicated by small grey bars. Haplotypes from locations with dark circle symbols included in reduced matrices for credible set hypothesis testing; all Monterey and Valley clade haplotypes were used in credible set hypothesis testing.

(n = 244, n = 1848, respectively) are found in the 95% credible set (n = 22789). The trans-valley Monterey and Valley clades are both strongly supported by the EF1G gene data, with paraphyly of eastern Monterey populations relative to a monophyletic set of western populations (eastern 85 and 86 closely related to western 78–80; Fig. S1).

Five of eight primary clades are recovered on the anonymous locus Bayesian consensus tree (Fig. S1). Closely related Sierran clades are not resolved as distinct clades by the anonymous locus data. A trans-valley Valley clade is strongly supported. The entire Monterey clade is not recovered, but the consensus tree includes a Monterey subclade of closely related haplotypes from eastern and western populations (eastern 85 and 86 closely related to western 78–80; Fig. S1). Trees with a monophyletic Monterey clade are in the 95% credible set (498 of 13308).

Monterey and valley clade networks

To better visualize patterns of allelic variation and relationships within the Valley and Monterey clades, we also reconstructed haplotype relationships using network methods. For the less-divergent 28S, EF1G and anonymous data, we used TCS v1.21 (Clement *et al.*

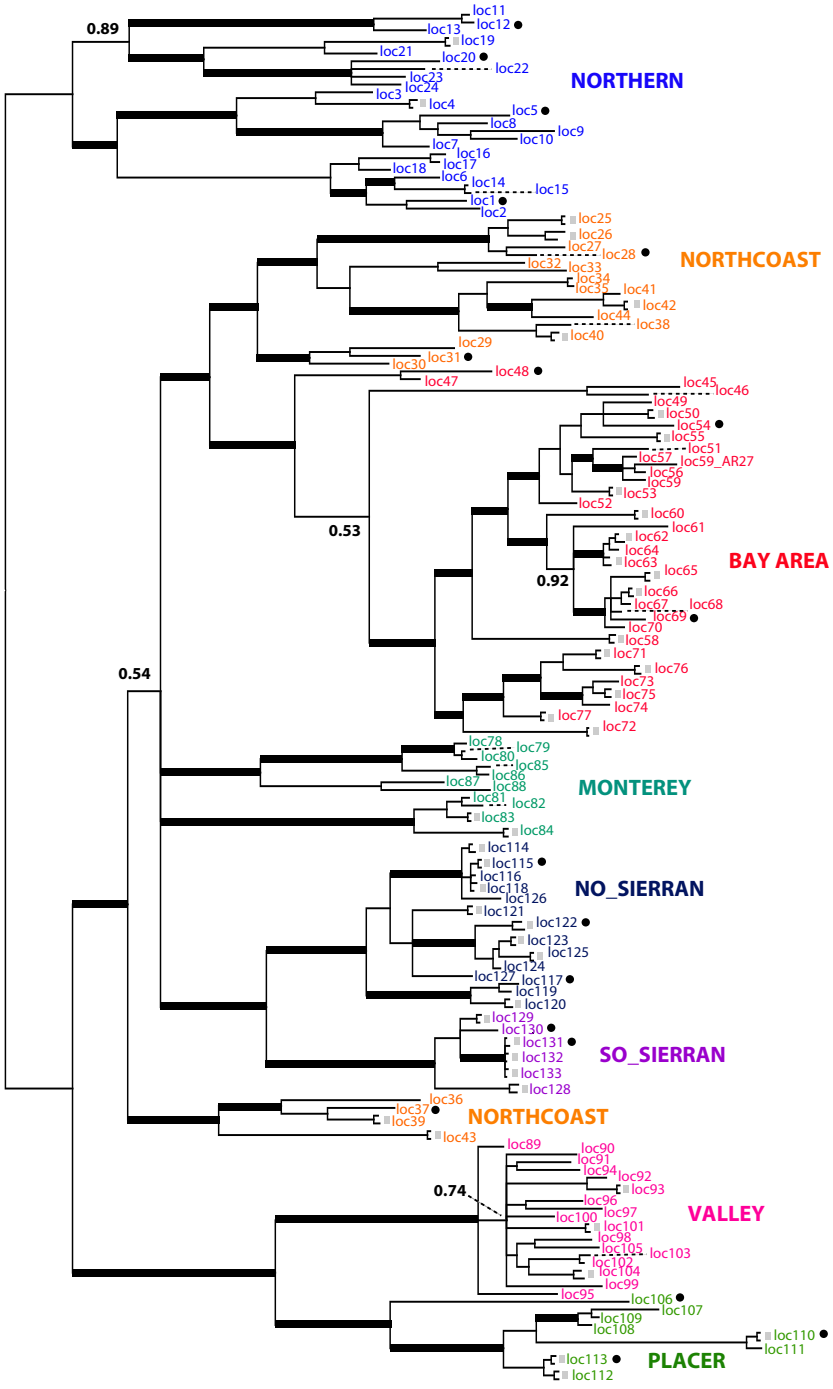

Fig. 3 CO1 partitioned Bayesian consensus phylogram. Support values shown only for clades including haplotypes from three or more sampling locations, with strong support (posterior probability > 0.95) indicated by thickened branches. Support values for other relevant bipartitions indicated by exact posterior probability values. Multiple haplotypes sampled from a single location indicated by small grey bars. Haplotypes from locations with dark circle symbols included in reduced matrices for credible set hypothesis testing; all Monterey and Valley clade haplotypes were used in credible set hypothesis testing.

2000). Because the CO1 data were too divergent for TCS (analyses resulted in many unconnected networks), we used the parsimony splits method implemented in SplitsTree v4.11.3 (Huson & Bryant 2006). These networks clearly reveal several relevant patterns (Fig. 4). In the Monterey clade, western populations (78–80) carry limited genetic diversity for all genes, but some diversity has evolved in the rapidly evolving CO1 gene region. Gene networks clearly show that Monterey clade haplotypes from western populations 78–80 are closely related to haplotypes from eastern populations 85 and 86 for all gene regions. For the EF1G and anonymous genes, some identical haplotypes are shared across the Central Valley. This striking genetic similarity contrasts with the generally high mutational divergence seen in the Monterey clade as a whole. In the

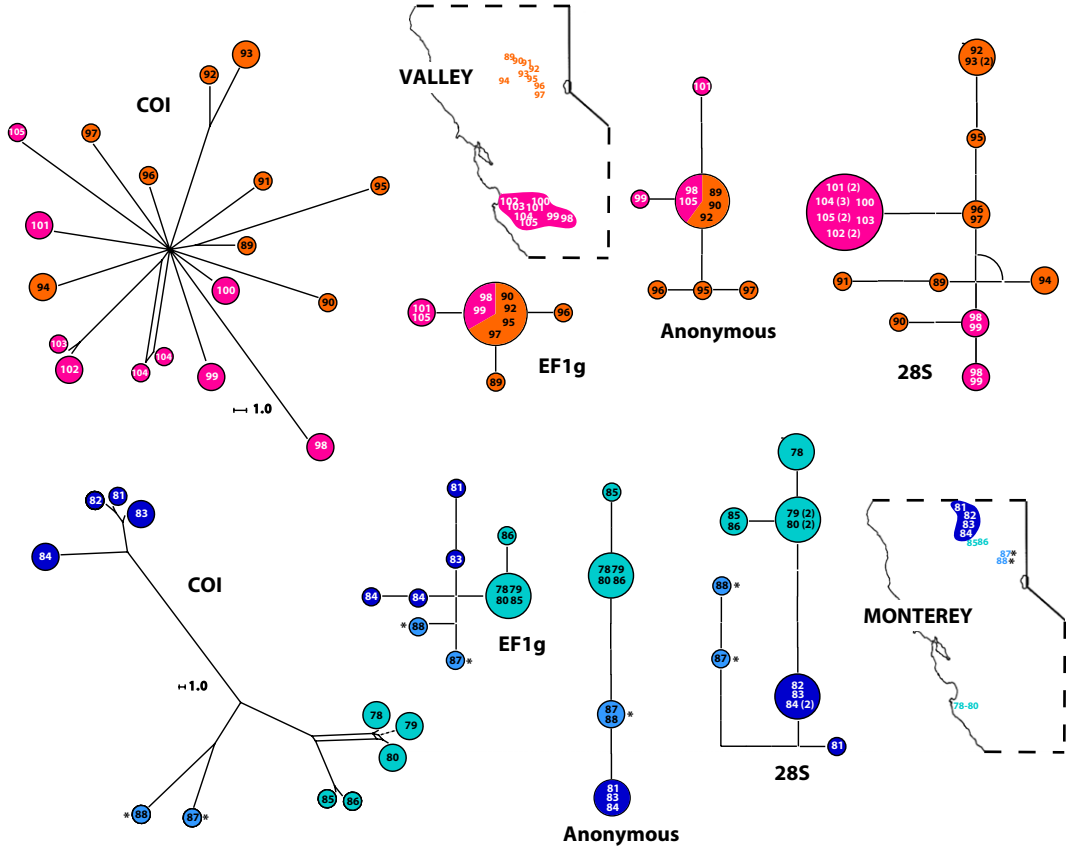

Fig. 4 Haplotype networks for members of the Valley and Monterey clades. 28S, anonymous locus and EF1G networks were reconstructed using TCS v1.21 (Clement *et al*. 2000). TCS network single branch segments correspond to one mutational step. CO1 networks reconstructed using the parsimony splits method (Huson & Bryant 2006). Haplotype circle sizes correspond to haplotype frequencies. Map insets show the geographic distribution of Valley and Monterey collection sites.

Valley clade, none of the gene networks imply regional monophyly (i.e. west vs. east reciprocal monophyly), and some EF1G and anonymous haplotypes are shared across the Central Valley. The Valley CO1 parsimony splits network is unique in revealing an array of highly divergent 'tip' haplotypes; internal haplotypes have either gone extinct or were not sampled.

Population structure analyses

Isolation by distance analyses reveal an overall strong positive relationship between geographic and CO1 genetic distances within primary clades, with more geographically distant comparisons showing higher levels of genetic divergence (Fig. 5). Mantel test *r* values range from 0.191 to 0.729, and these values are mostly statistically significant with matrix randomization (*P*-values ranging from 0.06 to 0.0001; Fig. 5). The two lowest *r* values are found for the trans-valley Monterey and Valley clades, which as expected each show two distinct clusters of points corresponding to geographic comparisons within a region (east or west), vs. longer-distance pairwise comparisons across the Central Valley. These two clades also show a second atypical pattern, this being a subset of pairwise comparisons that are relatively genetically similar despite being geographically distant (e.g. sites with genetic distances less than 10% corrected that are separated by 300–400 km). For all other primary clades (i.e. the 'typical' pattern in the *A. riversi* complex), the expected level of genetic divergence for such geographically distant sites is much higher (based on RMA slope estimates; Fig. 5). To summarize, certain distant trans-valley comparisons within the Monterey and Valley clades are *not as divergent as expected*, given isolation-by-distance dynamics elsewhere in the *A. riversi* complex.

For unconstrained analyses (theta and *M* values free to vary), Migrate-n estimates of theta posterior distributions are similar across replicate runs (Table 2). For the Monterey clade, theta 95% highest posterior density interval (HPD) values do not overlap (west 0–0.0035, east 0.0047–0.0198; Table 2), indicating a reduced theta in the west. Theta west is less than theta east for the Valley clade, but the posterior

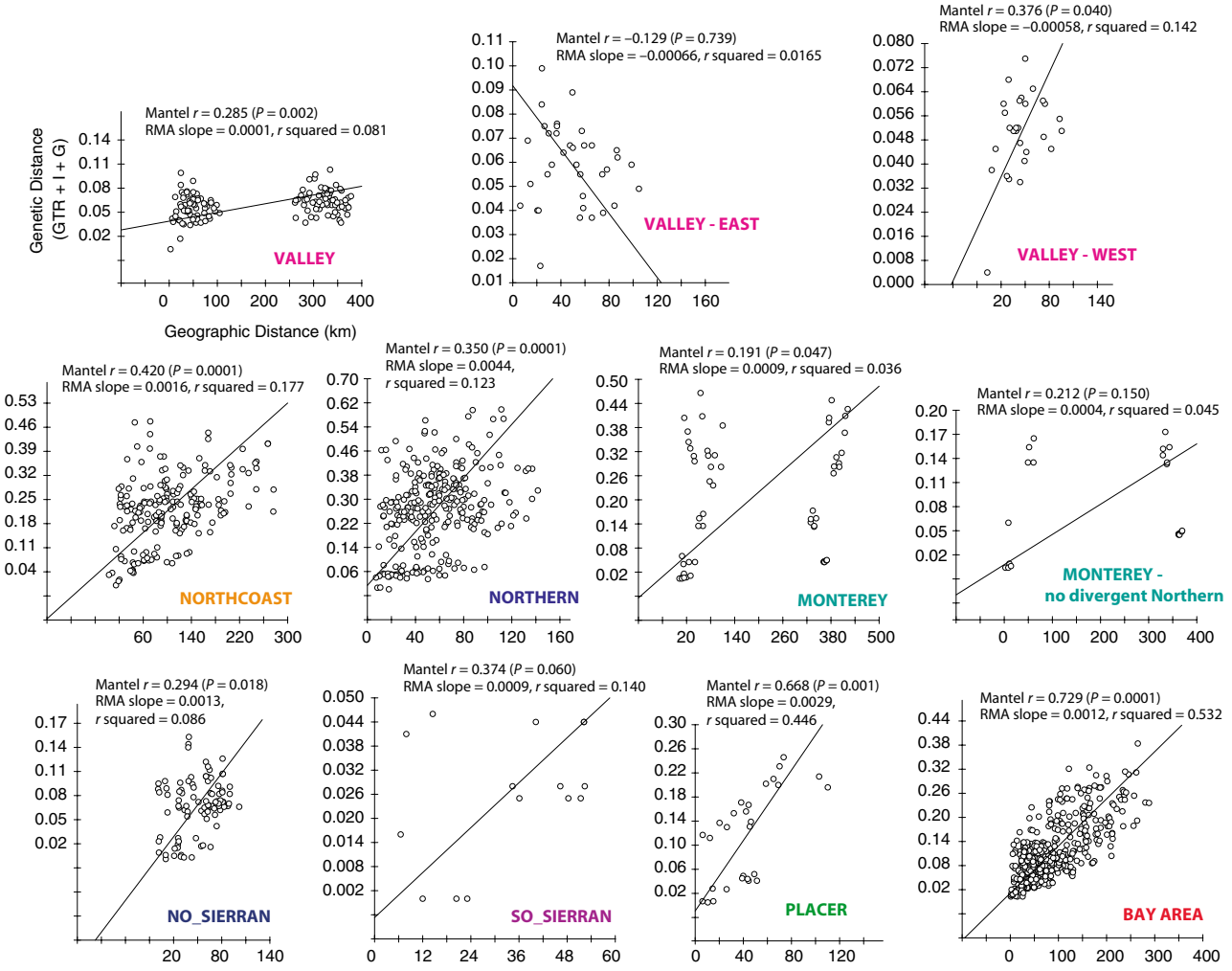

Fig. 5 Results of CO1 isolation-by-distance analyses for primary genetic clades, including Mantel test and RMA regression results. 'Monterey – no divergent Northern' analyses conducted excluding data for populations 81–84, which are recovered as a separate clade in CO1 Bayesian analyses (Fig. 3).

distributions overlap (Table 2). Modal values for the effective migration rate (M) suggest higher east-to-west gene flow directionality for both clades, but the posterior distributions for these values are very broad (Table 2) and right-skewed (Fig. S2 Supporting information). For both clades, log marginal likelihood (l mL) values for a symmetric M ($M_{west>east} = M_{east>west}$) model are similar to those for an unconstrained model (Table 2), but we note that l mL values vary slightly across runs and differ slightly for thermodynamic integration vs. Bezier-corrected values (Table 2). Because Bezier-corrected values might be more reliable with relatively short Migrate-n runs (Beerli & Palczewski 2010), we preferred these l mL values. Log Bayes factors based on least different Bezier l mL values (Monterey, 1.84; Valley, 2.6) do not provide strong or decisive support for alternative migration models (Kass & Raftery 1995).

Divergence time analysis

Relaxed clock Bayesian analyses were conducted on a combined data matrix including specimens from 58 locations. Analyses conducted using a Speciation: Birth–Death tree prior result in phylogenetic trees that provide strong support (posterior probabilities near 1.0) for eight primary clades (Fig. 6), and strong support for certain interclade relationships (i.e. Northcoast and Bay, Placer and Valley, Northern and Southern Sierran). Root position inferred using the relaxed molecular clock places the Northern clade as sister to all other clades, consistent with prior hypotheses of root position in the complex using outgroups (Starrett & Hedin 2007). Estimates for the age of the most recent common ancestor (MRCA) of the entire *A. riversi* complex (mean approximately 9.74 Ma, 95% HPD 7–13 Ma) are generally consistent with molecular clock estimates

Table 2 Results of Migrate-n analyses

	Theta – west	Theta – east	*M* east > west	*M* west > east	Raw thermodynamic l mL	Bezier l mL
Monterey						
Run 1 – full	0.0010 (0–0.0035)	0.0096 (0.0047–0.0193)	18 (0–649)	2 (0–307)	−4282.94	−3867.25
Run 2 – full	0.0011 (0–0.0035)	0.0098 (0.0047–0.0198)	20 (0–632)	0.3 (0–279)	−4283.03	**−3867.16**
Run 1 – symmetrical *M*	0.0014 (0–0.0037)	0.0096 (0.0045–0.0189)	50 (5.3–179)	50 (5.3–179)	−4283.74	−3865.89
Run 2 – symmetrical *M*	0.0014 (0–0.0037)	0.0092 (0.0044–0.0188)	52 (8-180)	52 (8–180)	−4287.03	**−3866.21**
Valley						
Run 1 – full	0.0025 (0.0001–0.0050)	0.0099 (0.0041–0.0214)	16 (0–768)	0.3 (0–750)	−5798.91	**−5429.5**
Run 2 – full	0.0024 (0–0.0059)	0.0094 (0.0043–0.0215)	14 (0-691)	0.3 (0–578)	−5798.82	−5429.1
Run 1 – symmetrical *M*	0.0026 (0–0.0057)	0.0093 (0.0042–0.0202)	228 (88–820)	228 (88–820)	−5808.86	**−5430.8**
Run 2 – symmetrical *M*	0.0025 (0–0.0058)	0.0102 (0.0045–0.0208)	281 (63–709)	281 (63–709)	−5814.81	−5431.4

Modal and 95% HPD (in parentheses) theta and *M* values shown for ALL loci (see Beerli 2012).

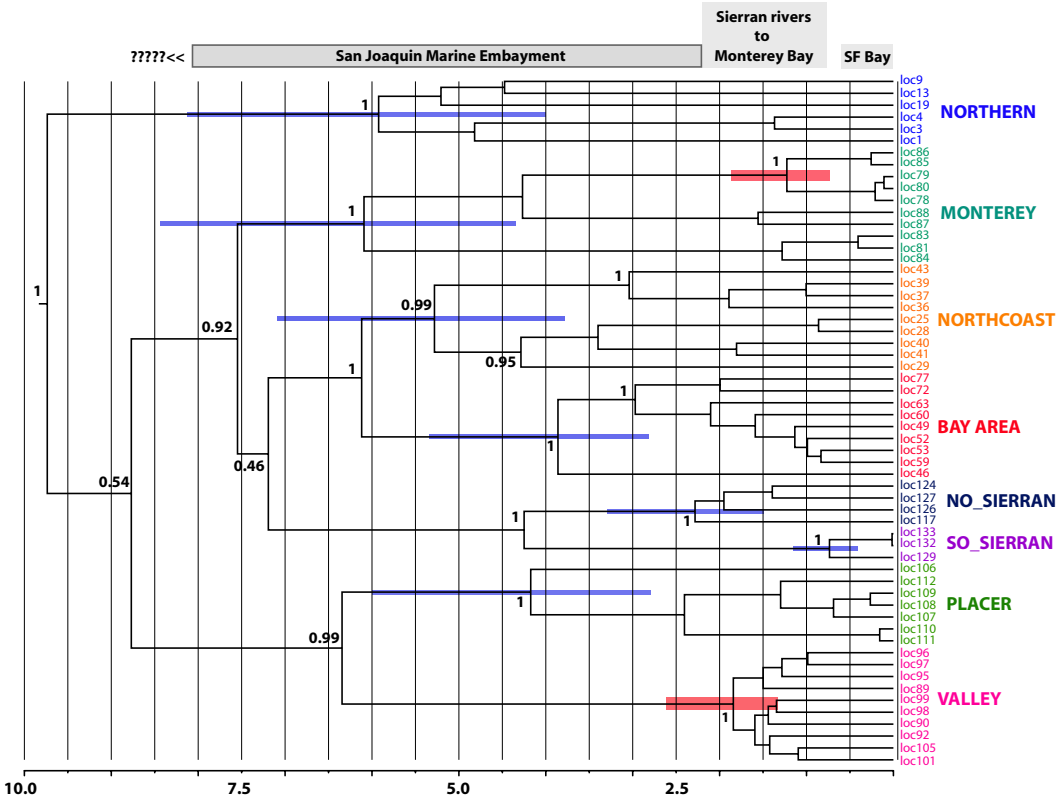

Fig. 6 Maximum clade credibility tree resulting from BEAST analysis using a Birth Death tree prior. Posterior probability values included for all primary genetic clades, nodes uniting primary genetic clades and other relevant nodes within genetic clades. HPD values included for all primary genetic clades. HPD values for trans-valley Valley clade and Monterey subclade thickened. Paleogeographic reconstructions shown above chronogram (see Discussion).

reported by Starrett & Hedin (2007), conducted using penalized likelihood and biogeographic calibrations. The mean age estimate for the MRCA of the Monterey subclade that includes both western and eastern populations (populations 78–80, 85–86) is 1.22 Ma (95% HPD 0.71–1.86 Ma), while the mean age estimate for

the MRCA of the entire Valley clade is 1.84 Ma (95% HPD 1.28–2.57 Ma). BEAST analyses conducted using alternative tree priors are consistent in always placing the Northern clade as sister to remaining clades and in recovering eight primary clades with strong support (results not shown). The use of alternative tree priors has minimal impact on MRCA mean age estimates for the Monterey subclade and entire Valley clade (Table 3), and as discussed below, these age estimates fall within a time window when northern Sierran rivers were flowing to the Monterey Bay (Fig. 6).

Discussion

General biogeographic patterns

A significantly expanded geographic and genetic sample uncovers biogeographic patterns not apparent in our original study of the *A. riversi* complex (Starrett & Hedin 2007) and reveals this system as being particularly informative for understanding the biogeographic history of California. The *A. riversi* complex has apparently existed in California for a long period of time (Fig. 6), and because of the unique biology of these animals, genetic differences arise and persist at small spatial scales. These spiders live a sedentary, fossorial existence, closely tied to favourable mesic microhabitats (Coyle 1968, 1971; Vincent 1993). Spiderlings leaving maternal burrows are small and prone to desiccation, are not known to balloon and typically settle close to maternal burrows. Adult males appear to represent the only conduit for potential gene flow, but even here, the patterns observed for the densely sampled nuclear 28S data demonstrate that the spatial scale of male wandering is likely limited. We have shown here that populations separated by more than 30 km are almost always genetically distinct for 28S and CO1, and in many cases, even more geographically proximate populations are genetically distinct. This finding of microgeographic genetic structuring is the rule for mygalomorph spiders in California (Bond *et al.* 2001; Bond 2004; Stockman & Bond 2007; Bond & Stockman 2008; Hedin & Carlson 2011; Satler *et al.* 2011).

Here, we briefly highlight several clade-specific biogeographic patterns, more extensively discussing trans-valley biogeographic patterns below. Three clades in the north Coast Ranges (Northern, Northcoast, Bay Area) have distributions oriented in a SE-NW trending parallel manner, with northern populations of the Northcoast clade extending past southern populations of the Northern clade, and northern populations of the Bay Area clade extending past southern populations of the Northcoast clade. Much of this pattern may result from range expansion both northwards (locations 25–28) and southwards (locations 41, 42, 44) of the Northcoast clade (Figs. 1 and 2). The northernmost locations occupied by the Northcoast clade are thought to be the most recently emerged uplands in the region, available since approximately 2 Ma (Lock *et al.* 2006; fig. 9). More generally, eastern uplands of the north Coast Ranges are older than western habitats, with regional uplift hypothesized to be related to the northern migration of the Mendocino triple junction (Lock *et al.* 2006). A prediction of this paleogeographic model is that northeastern populations (i.e. Northern clade) should be the oldest in the region, which is supported by MRCA time estimates for this clade (Fig. 6).

Several biogeographic patterns in the Bay Area clade are similar to patterns found in other regional taxa, including low-vagility salamanders. The more densely sampled mitochondrial and 28S gene trees (Figs 2 and 3) indicate nearly congruent structuring into northern (locations 45–48; Fig. 1), central (locations 49–70; Fig. 1), and southern (locations 71–77; Fig. 1) Bay Area subclades. The strongly supported phylogenetic separation of northern populations coincides with the region south of the modern Russian River. This region includes the Wilson Grove Formation resulting from a shallow marine embayment (Fox 1983) and is possibly near the mouth of the paleo-Russian river, estimated at 6–2 Ma (Lock *et al.* 2006). This low-lying region also likely represents a modern-day barrier to gene flow, with few suitable habitats for Bay Area spiders. Genetic breaks in this geographic region have been found in multiple amphibian species (Rissler *et al.* 2006; Martinez-Solano *et al.* 2007). The central subclade includes populations found both east and west of the San Francisco Bay, and both north and south

Table 3 Divergence time estimates from BEAST analyses using alternative tree priors

	Birth death	Birth death incomplete	Yule	Coalescent
MRCA – Monterey subclade	1.22 Ma (95% HPD 0.71–1.86 Ma)	1.21 Ma (95% HPD 0.71–1.81 Ma)	1.30 Ma (95% HPD 0.70–1.84 Ma)	1.22 Ma (95% HPD 0.69–1.92 Ma)
MRCA – entire Valley clade	1.84 Ma (95% HPD 1.28–2.57 Ma)	1.83 Ma (95% HPD 1.29–2.52 Ma)	1.83 Ma (95% HPD 1.30–2.53 Ma)	1.84 Ma (95% HPD 1.23–2.58 Ma)

of the Golden Gate (locations 49–70; Fig. 1). Spiders from Angel Island in the San Francisco Bay (location 52) carry unique mitochondrial haplotypes (Fig. 3), suggesting that this population is a hilltop remnant of Bay submersion rather than over-water dispersal. Most island populations of *Batrachoseps* salamanders in the San Francisco Bay were also established on uplands that later became islands (Martinez-Solano & Lawson 2009). There is mitochondrial evidence within the central subclade for an influence of the Colma Strait (locations 53, 62), a mid- to late-Pleistocene marine passage that may have drained the San Francisco Bay prior to the opening of the Golden Gate (Elder 2001; Sloan 2006). A Colma Strait break is also found in the mitochondrial genomes of *Batrachoseps* salamanders (Martinez-Solano & Lawson 2009). Another biogeographic break shared between Bay Area spiders and *Batrachoseps* (Martinez-Solano *et al.* 2007) is found in the east Bay at Dublin Canyon, east of Castro Valley (locations 59 and 60, Figs 1 and 2). The southern boundary of the southern Bay subclade is coincident with the 'Pajaro Gap', possibly defined by the northern edge of the San Joaquin marine embayment. Many Californian taxa have southern distributional limits coincident with this well-known biogeographic barrier (see Wake 1997; Kuchta *et al.* 2009a).

In the Sierra Nevada, most primary clades are distributed in a north-to-south manner relative to other clades (Fig. 1), with populations within any single clade ranging from mid- to high elevations (600–3200 m). An exception involves Sierran populations of the Valley clade, with most populations found below 600 m in the foothills west of both the Monterey and Placer clades. This suggests potential elevational partitioning of genetic clades in this region, although further sampling at lower elevations for all clades is needed to rigorously test this hypothesis. Major rivers and river canyons are oriented approximately perpendicular to the Sierran montane axis and might act as potential distributional barriers (e.g. Rovito 2010). However, modern-day clade boundaries in *A. riversi* are not obviously coincident with these riverine barriers, as all primary clades include populations on opposite sides of large rivers.

Trans-valley biogeography

Western populations of the Monterey and Valley clades occur in the south Coast Ranges of California, including the Sierra de la Salinas, Santa Lucia, Gabilan and Diablo Ranges. Most populations occur in the Salinas, Santa Lucias and Gabilans, west of the San Andreas fault on the Salinian block of the Pacific plate. Here, we summarize the paleogeography and geological history of this region (see also Kuchta & Tan 2006). Prior to 8 Ma, uplands on the Pacific plate were much further south and mostly submerged, forming island arcs (Sims 1993; Hall 2002; plates 2–5). From approximately 8–2.2 Ma, the uplands near Monterey were close to their current latitudinal position and were apparently emergent (Hall 2002; plates 6–8). Multiple authors have suggested that the time period after 5 Ma included extensive south Coast Range tectonic uplift (Miller 1999; Argus & Gordon 2001), although some authors suggest much more recent uplift (e.g. 0.4 Ma, Page *et al.* 1998). Also during the 8 to 2.2 Ma time period, a large marine embayment extended from just north of Monterey south-eastward, cutting through the south Coast Ranges and filling most of the southern San Joaquin Valley (Dupre 1990; Hall 2002, plates 6–8; Bowersox 2005, fig. 8; Powell *et al.* 2007). The northern limit of this embayment was approximately at a latitude similar to modern-day Monterey Bay (Fig. 1; Davis & Coplen 1989).

Marine invertebrate fossils have been well studied in the San Joaquin Valley, and there is an obvious marine to freshwater fossil faunal turnover at 2.2 Ma (Bowersox 2005; Powell *et al.* 2007). Both the northern and southern arms of the Central Valley were emergent from 2.2 to approximately 0.8 Ma, with an outlet of Sierran paleoriver systems just north of Monterey. A temporal pulse in sediment deposition in the Monterey submarine fan is apparently correlated with large riverine sources during this time period (Normark 1999), and the Feather and Yuba rivers of northern California show particularly high incision rates from 5 Ma to the present (Wakabayashi & Sawyer 2001). The Monterey outlet was ultimately closed (perhaps because of continued south Coast Range uplift), creating a large freshwater lake (Lake Corcoran) in the Central Valley (Sarna-Wojcicki *et al.* 1985). This lake was short-lived, and ultimately drained to the ocean at Carquinez Strait about 0.6 Ma, creating the modern-day drainage of Sierran rivers to the San Francisco Bay (Sarna-Wojcicki *et al.* 1985).

We contend that western populations of Monterey and Valley clades cannot be older than 8 Ma, as Salinian block upland habitats were either unavailable or much further south at this time. Extant members of the *A. riversi* complex are northern in origin (Starrett & Hedin 2007) and have never been recorded as far south as habitats would have been at this time (Hall 2002). From 8 to 2.2 Ma, the San Joaquin marine embayment likely prevented northern access to south Coast Ranges (Fig. 6), if one accepts the argument that marine habitats would be difficult to cross for small terrestrial habitat-specialized animals. After 0.6 Ma, there is no obvious biogeographic connection between foothill habitats in the northern Sierras and uplands in the south Coast Ranges, because Sierran rivers were then flowing to the San Francisco Bay. As such, we hypothesize that

western populations of the Monterey and Valley clades originated from north-eastern ancestors sometime during the 2.2- to 0.6-Ma time interval. Not only were south Coast Range habitats uplifted and accessible to northern ancestors at this time, but large rivers (and associated riparian habitat corridors) in fact connected Sierran habitats to those in the vicinity of Monterey from 2.2–0.8 Ma. Older ages are possible if one allows for a permeable San Joaquin embayment barrier (some mygalomorphs spiders unrelated to antrodiaetids can cross marine barriers, Raven 1994), and acknowledges that this barrier did not extend very far north (Davis & Coplen 1989). However, under this scenario, there is no readily apparent biogeographic connection between the northern Sierran foothills and uplands in the south Coast Ranges. Also, more geographically proximate regions (e.g. Bay Area uplands) would seem to be more probable source areas if the marine embayment was indeed crossable.

Relaxed molecular clock analyses are consistent with the 2.2–0.8 Ma Sierran to Monterey temporal window hypothesis for both Monterey and Valley clades (Fig. 6), using multiple BEAST tree prior settings (Table 3). The actual mechanism of spider movement from Sierran to Monterey region habitats (or vice versa) is less clear, although two alternative hypotheses seem most plausible. The first hypothesis is that spiders were transported passively via river rafting events, a 'jump dispersal' hypothesis. An alternative hypothesis invokes active overland dispersal, perhaps along riparian riverine corridors. For the Monterey clade, all available evidence is consistent with a jump dispersal hypothesis. Phylogenetic analysis of four independent gene regions indicates a clear east to west directionality, with eastern populations paraphyletic with respect to a nested set of western populations (Figs 2–4, Supplemental Fig. 1). Western populations carry a subset of genetic variation found in the clade as a whole (Fig. 4), well illustrated by k values for all genes (Table 1), and Migrate-n results (Table 2). Levels of genetic divergence suggest a relatively recent genetic connection between the east and the west, with eastern and western populations (sites 78–80, 85, 86) sharing identical alleles for some genes (Fig. 4). Site 86 is from the Feather River canyon (Table S1), which historically drained to the Monterey Bay. The level of genetic similarity observed over such a large geographic distance is inconsistent with isolation by distance dynamics elsewhere in the *A. riversi* complex, where large genetic distances accrue over short geographic distances (Fig. 5). At the same time, western Monterey populations have evidently been in place long enough for some genetic differences to evolve for the rapidly evolving mitochondrial CO1 (Fig. 4, Table 1). All of these genetic signatures are consistent with a historical jump dispersal event and associated population bottleneck (Excoffier *et al.* 2009). A rafting hypothesis is also consistent with the natural history of these spiders, which often build burrows amongst the root systems of large trees where soils are shaded and stabilized (Coyle 1971; Vincent 1993; M. Hedin & J. Starrett, personal observation).

Eastern Monterey populations in the northern Sierra Nevada occur in coniferous forest, where spiders build relatively tall silken turrets in *Pinus* sp. and *Pseudotsuga menziesii* coniferous forest litter (M. Hedin & J. Starrett, personal observation). The three known western Monterey populations are found in Monterey pine (*Pinus radiata*) forest, with spiders building relatively tall silken turrets in coniferous forest litter. It appears that there has not been an ecological shift coincident with range expansion, but instead, ecological conservatism may have constrained the habitat colonization event to a specific habitat type. We also note that western Monterey populations are not likely to be more geographically widespread than current collections indicate – nearby chaparral habitats are occupied by Valley clade spiders (Fig. 1), and we have not found populations further south in the Santa Lucia Range despite extensive search efforts. This region includes coniferous forest (dominated by *Sequoia sempervirens*), but mostly lacks Monterey pine (see www.calflora.org/). Relatively recent colonization, in combination with dispersal limitation, may likewise explain this very small geographic distribution.

Although the mean age estimate for the MRCA of the Valley clade is consistent with the 2.2–0.8 Ma Sierran to Monterey time window (Fig. 6), available genetic data do not clearly support a particular movement hypothesis. The lack of reciprocal monophyly of western Valley populations (Figs 2–4, Fig. S1), and the evidence that western Valley populations are as genetically diverse as the clade as a whole (Fig. 4, Tables 1 and 2), conflicts with a 'Monterey-like' genetic bottleneck hypothesis. Genetic patterns in the Valley clade are potentially consistent with a jump dispersal event involving a more genetically diverse source population, or perhaps multiple dispersal events. The star-like shape of the mitochondrial network (Fig. 4) is likewise potentially consistent with spatial and demographic expansion (Excoffier *et al.* 2009), with mutations on long-terminal branches. However, Fu and Li tests (Fu & Li 1993) for all genes for both Valley and western Valley sequences are nonsignificant (Table 1). Although isolation by distance plots seems inconsistent with overland dispersal, we are also hesitant to reject an overland dispersal with vicariance model, where gene tree mixing results from the retention of ancestral genetic polymorphism.

Ecologically, eastern Valley populations have been found in relatively xeric oak woodland habitats, includ-

ing the Sutter Buttes, a relictual oak woodland island in the otherwise flat Sacramento Valley (Schoenherr 1992; Fig. 1). Sierran Valley clade spiders often build burrows with minimal silken turrets. Western Valley populations are also found in relatively xeric situations (including *Arctostaphylus*-dominated chaparral), also building burrows with minimal silken turrets. We note that the western Valley populations are relatively geographically widespread in comparison with the microendemic western Monterey populations (Fig. 1). This larger geographic distribution might reflect a greater extent of suitable habitat in south Coast Ranges, but could also reflect multiple colonization events or higher rates of overland dispersal. Finally, we point out that western Valley populations are geographically closer to eastern Valley populations in comparison with the widely separated eastern and western Monterey clade populations (Fig. 1), again raising the possibility of multiple colonization events. Ultimately, denser geographic sampling, additional genetic markers and spatially explicit coalescent analyses may help clarify the origins of eastern Valley populations.

Trans-valley patterns found in other taxa

Several phylogeny-based biogeographic studies have revealed putatively conspecific populations on opposite sides of the Central Valley, consistent with trans-valley dispersal or vicariance. In the dispersal-limited salamanders *Ensatina eschscholtzii xanthopicta* (Wake 1997) and *Batrachoseps attenuatus* (Martinez-Solano *et al.* 2007), a trans-valley leak is centered at latitudes near or north of the San Francisco Bay. In these taxa, phylogenetic evidence supports a west (Coast Range) to east (Sierran) directionality, and low molecular divergence is consistent with dispersal in oak woodlands in the mid- to late Pleistocene (Wake 1997; Martinez-Solano *et al.* 2007). Kuchta *et al.* (2009) argued that the directionality in *Ensatina* may be from east to west, but agreed on the mid to late-Pleistocene timeframe. Mitochondrial evidence for a similarly recent mid-Central Valley dispersal event is seen in *Neotoma* woodrats, in this case with east to west directionality (Matocq 2002). Several phylogeographic studies include south Coast Range populations that show relatedness to conspecifics in the Sierra Nevadas. These include trapdoor spiders (*Aliatypus janus*; Satler *et al.* 2011), kingsnakes (Rodriguez-Robles *et al.* 1999) and *Emys* turtles (Spinks *et al.* 2010). Although timing and directionality is not certain in most cases, phylogenetic evidence suggests that these south Coast Range populations share ancestry with Sierran populations at similar or more southern latitudes, not with populations from the northern Sierran foothills.

A small number of studies show hints of the phylogeographic pattern that we have discovered in *A. riversi* but are single gene studies with sparse sampling in the geographic regions of primary interest. In *Rana boylii* frogs, a mitochondrial 'central coast' clade includes animals from the south Coast Ranges and the northern Sierran foothills (Yuba county, Lind *et al.* 2011). In *Promyrmekiaphila* trapdoor spiders, populations from the eastern edge of the Diablo Range are related by mitochondrial evidence to animals from the northern Sacramento Valley, rather than to geographically adjacent populations (Stockman & Bond 2007). Overall, despite a long history of biogeographic study on a wide variety of Californian taxa, no prior studies reveal trans-valley patterns that are convincingly similar to those observed in the *A. riversi* complex.

There are many possible reasons for the apparent uniqueness of the trans-valley biogeographic pattern found in the *A. riversi* complex. First, we note that *among taxa that have been studied*, mygalomorph spiders are extremely low dispersal organisms, implying limited 'post-event' gene exchange that may erase biogeographic signal. It is possible that similar biogeographic events have occurred in other taxa, but that subsequent distributional movements and/or gene flow have concealed this history. Second, we acknowledge that for the Monterey clade, the hypothesized jump dispersal event must be rare in occurrence and that *A. riversi* has a unique silken burrow biology facilitating survival of such a dispersal event. Third, we hypothesize that the south Coast region must have represented vacant habitat for *A. riversi*, since clade syntopy is unknown in this complex. For other taxa, it is possible that regional populations already occupied the south Coast Ranges, making colonization difficult because of competition with congeners or conspecifics (see Waters 2011). Finally, we emphasize that although California biogeography might be characterized as 'well known', in fact much remains to be learned. The majority of California endemic radiations remain to be studied, particularly in diverse ground-dwelling arthropod taxa (e.g. millipedes, harvestmen, ground-dwelling hexapods and other spiders) for which very few phylogeny-based studies exist. With studies of these taxa, we anticipate the discovery of many new Californian biogeographic patterns.

Primary clades as cryptic species

The level of genealogical congruence observed in the *A. riversi* complex is only expected if primary genetic clades represent independently evolving groups (i.e. species). Our data includes examples where a relatively fast-evolving gene region (e.g. CO1) fragments members of a purported single

clade, and cases where a slowly evolving gene region collapses members of two purported clades. However, the data include no examples of specimens 'misplaced' in genetic clades. This pattern exists despite several instances of clade parapatry with the potential for genetic exchange. The closest contact that we have discovered involves Northern Sierran (locations 116–120) and Placer clades (111), with spider colonies separated by less than one kilometre along the south bank of the South Fork American River (Fig. 1). Analysis of several specimens per site for both CO1 and 28S genes reveals no evidence for genetic exchange in this region (data not shown), despite an apparent lack of extrinsic barriers. Phylogenetic evidence suggests that these clades are not sister taxa (Figs 2, 3 and 6; Fig. S1). Other cases of possible nonsister secondary contact involve Northern and Northcoast clades (locations 19 and 25), and Valley vs. Monterey clades (locations 85 and 89; 79 and 102). Cases of potential contact involving apparent sister clades include Northcoast and Bay Area (e.g. locations 42 and 49), and Northern Sierran vs. Southern Sierran clades (locations 125 and 130). All of the above-highlighted cases require denser geographic sampling to close small geographic gaps, but are consistent with interactions involving reproductively isolated units.

All prior studies of *A. riversi* have suggested that this taxon may represent a species complex rather than a single species. Coyle (1968) noted morphological divergence in both genitalic and somatic characters between coastal and Sierran populations of *A. riversi* and suggested that these might represent two incipient species. Ramirez & Chi (2004) tested Coyle's two species hypothesis through analysis of allozyme variation, with spiders collected from Monterey, Sutter Buttes, the Bay Area and the central Sierra Nevada (8 total locations). These authors suggested the possibility of multiple cryptic species. Starrett & Hedin (2007), with a smaller geographic sample than reported here, also argued for a species complex with five cryptic species. In the current study, eight primary genetic clades are hypothesized to represent cryptic species. This study differs from prior studies in having greater geographic sampling density and number of gene regions utilized, and illustrates how species discovery in dispersal-limited taxa is an iterative process. Because small areas can house unique lineages (e.g. southern Sierran, south-eastern Northcoast, etc.), refined geographic sampling is crucially important for species discovery. Jockusch *et al.* (2001) discuss a similar situation in Californian *Batrachoseps* salamanders, where species discovery continues despite decades of research by regional herpetologists.

Conclusions

Many prior studies have shown the biogeographic impact of the San Joaquin marine embayment on Californian taxa and have also shown that south Coast Range populations originate from further south in California, including the southern Sierra Nevada. To the best of our knowledge, no prior studies have discussed or revealed a northern Sierran to south Coast Range biogeographic connection. In addition to this unique pattern of north-eastern ancestry, we hypothesize that distinct south Coast Range populations have crossed the Central Valley at approximately the same time, but have perhaps done so via different evolutionary mechanisms. Burrowing mygalomorph spiders are roughly analogous to immoveable objects (e.g. rocks) in an evolving landscape. These objects themselves move very little, but are sometimes carried passively by external forces. With sufficient genetic and geographic sampling density, lineages of burrowing mygalomorphs have nearly unlimited potential to reveal biogeographic patterns at multiple spatial and temporal scales.

Acknowledgements

This study was supported by National Science Foundation REU supplement grant to M. Hedin and J. Starrett (DEB 0322650), and DDIG grant to C. Hayashi and J. Starrett (DEB 0910365). RJ Adams, Jason Bond, Dave Carlson, Shahan Derkarabetian, Bob Keith, Robin Keith, Joel Ledford, Maureen McCormack, Jordan Satler, Patrick Stadille and Steven Thomas helped to collect spiders. Dean Leavitt has been particularly helpful in specimen collection and in the discussion of ideas presented in this article. Harvey Kelsey provided insights into the geological history of northern California. Axel Schönhofer helped with initial BEAST analyses, Peter Beerli helped with Migrate analyses, and Amanda Kuelbs assisted in the laboratory. The Starrett family graciously housed and fed many folks on multiple field expeditions. Comments of Rosie Gillespie, Miquel Arnedo and two anonymous reviewers served to strengthen the manuscript.

References

Argus DF, Gordon RG (2001) Present tectonic motion across the Coast Ranges and San Andreas fault system in central California. *Geological Society of America Bulletin*, **113**, 1580–1592.

Arnedo MA, Ferrandez M (2007) Mitochondrial markers reveal deep population subdivision in the European protected spider *Macrothele calpeiana* (Walckenaer, 1805) (Araneae, Hexathelidae). *Conservation Genetics*, **8**, 1147–1162.

Beerli P (2012) Migrate documentation, Version 3.2.1. Available from http://popgen.sc.fsu.edu/Migrate/Download.html.

Beerli P, Palczewski M (2010) Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics*, **185**, 313–326.

Bidegaray-Batista L, Arnedo MA (2011) Gone with the plate: the opening of the western Mediterranean basin drove the

diversification of ground-dweller spiders. *BMC Evolutionary Biology*, **11**, 317.

Bond JE (2004) Systematics of the Californian euctenizine spider genus *Apomastus* (Araneae: Mygalomorphae: Cyrtaucheniidae): the relationship between molecular and morphological taxonomy. *Invertebrate Systematics*, **18**, 361–376.

Bond JE (2012) Systematics and taxonomic revision of the trapdoor spider genus *Aptostichus* Simon (Araneae: Mygalomorphae: Euctenizidae). *Zookeys*, in press.

Bond JE, Stockman AK (2008) An integrative method for delimiting cohesion species: finding the population-species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. *Systematic Biology*, **57**, 628–646.

Bond JE, Hedin MC, Ramirez MG, Opell BD (2001) Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. *Molecular Ecology*, **10**, 899–910.

Bond JE, Beamer DA, Lamb T, Hedin M (2006) Combining genetic and geospatial analyses to infer population extinction in mygalomorph spiders endemic to the Los Angeles region. *Animal Conservation*, **9**, 145–157.

Bowersox JR (2005) Reassessment of extinction patterns of Pliocene molluscs from California and environmental forcing of extinction in the San Joaquin Basin. *Paleogeography, Paleoclimatology, Paleoecology*, **221**, 55–82.

Calsbeek R, Thompson JN, Richardson JE (2003) Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. *Molecular Ecology*, **12**, 1021–1029.

Chatzimanolis S, Caterino MS (2007) Toward a better understanding of the "transverse range break": lineage diversification in southern California. *Evolution*, **61**, 2127–2141.

Clement M, Posada D, Crandall K (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1660.

Cooper SJB, Harvey MS, Saint KM, Main BY (2011) Deep phylogeographic structuring of populations of the trapdoor spider *Moggridgea tingle* (Migidae) from southwestern Australia: evidence for long-term refugia within refugia. *Molecular Ecology*, **20**, 3219–3236.

Coyle FA (1968) The mygalomorph spider genus *Atypoides* (Araneae: Antrodiaetidae). *Psyche*, **75**, 157–193.

Coyle FA (1971) Systematics and natural history of the mygalomorph spider genus *Antrodiaetus* and related genera (Araneae: Antrodiaetidae). *Bulletin Museum Comparative Zoology Harvard*, **141**, 269–402.

Cranston KA, Rannala B (2007) Summarizing a posterior distribution of trees using agreement subtrees. *Systematic Biology*, **56**, 578–590.

Davis GH, Coplen TB (1989) Late Cenozoic paleohydrogeology of the western San Joaquin Valley, California, as related to structural movements in the Central Coast Ranges. *Geological Society of America Special Paper*, **234**, 1–40.

Davis EB, Koo MS, Conroy C, Patton JL, Moritz C (2008) The California hotspots project: identifying regions of rapid diversification of mammals. *Molecular Ecology*, **17**, 120–138.

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

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology*, **4**, e88.

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology & Evolution*, **29**, 1969–1973.

Dupre WR (1990) Quaternary geology of the Monterey Bay region, California. In: *Geology and Tectonics of the central California Coastal Region, San Francisco to Monterey* (eds Garrison RE, Greene HG, Hicks KR, Weber GE, Wright TL). pp. 185–191. Pacific Section of the American Association of Petroleum Geologists, Bakersfield, California.

Elder WP (2001) Geology of the Golden Gate Headlands. In: *Geology and Natural History of the San Francisco Bay Area: A Field Trip Guidebook* (eds Stoffer PW, Gordon LC). pp. 61–86. US Geological Survey Bulletin 2188.

Eskov K, Zonshtein S (1990) First Mesozoic mygalomorph spiders from the Lower Cretaceous of Siberia and Mongolia, with notes on the system and evolution of the infraorder Mygalomorphae (Chelicerata:Araneae). *Neues Jahrbuch fur Geologie und Paläontologie, Abhandlungen*, **178**, 325–368.

Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. *Annual Review of Ecology, Evolution & Systematics*, **40**, 481–501.

Feldman CR, Spicer GS (2006) Comparative phylogeography of woodland reptiles in California: repeated patterns of cladogenesis and population expansion. *Molecular Ecology*, **15**, 2201–2222.

Fox Jr KF (1983) Tectonic setting of late Miocene, Pliocene, and Pleistocene rocks in part of the Coast Ranges north of San Francisco, California. *Geological Survey Professional Paper*, **1239**, 1–33.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

Griswold CE, Ledford J (2001) A monograph of the migid trap door spiders of Madagascar and review of the world genera (Araneae, Mygalomorphae, Migidae). *Occasional Papers of the California Academy of Sciences*, **151**, 1–120.

Hall Jr CA (2002) Nearshore marine paleoclimate regions, increasing zoogeographic provinciality, molluscan extinctions, and paleoshorelines, California: late Oligocene (27 Ma) to late Pliocene (2.5 Ma). *Geological Society of America Special Paper*, **357**, 1–489.

Hedin M, Carlson D (2011) A new trapdoor spider species from the southern Coast Ranges of California (Mygalomorphae, Antrodiaetidae, *Aliatypus coylei*, nov sp), including consideration of mitochondrial phylogeographic structuring. *Zootaxa*, **2963**, 55–68.

Hendrixson BE, Bond JE (2005) Testing species boundaries in the *Antrodiaetus unicolor* complex (Araneae: Mygalomorphae: Antrodiaetidae): "paraphyly" and cryptic diversity. *Molecular Phylogenetics and Evolution*, **36**, 405–416.

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology & Evolution*, **23**, 254–267.

Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. v 3.16. *BMC Genetics*, **6**, 13. Available from http://ibdws.sdsu.edu.

Jockusch EL, Wake DB (2002) Falling apart and merging: diversification of slender salamanders (Plethodontidae: *Batracho-*

seps) in the American West. *Biological Journal of the Linnean Society*, **76**, 361–391.

Jockusch EL, Yanev KP, Wake DB (2001) Molecular phylogenetic analysis of slender salamanders, genus *Batrachoseps* (Amphibia: Plethodontidae), from central coastal California with descriptions of four new species. *Herpetological Monographs*, **15**, 54–99.

Kass RE, Raftery AE (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.

Kraft NB, Baldwin BG, Ackerly DD (2010) Range size, taxon age and hotspots of neoendemism in the California flora. *Diversity & Distributions*, **16**, 403–413.

Kuchta SR, Tan A (2006) Lineage diversification on an evolving landscape: phylogeography of the California newt, *Taricha torosa* (Caudata: Salamandridae). *Biological Journal of the Linnean Society*, **89**, 213–239.

Kuchta SR, Parks DS, Mueller RL, Wake DB (2009a) Closing the ring: historical biogeography of the salamander ring species *Ensatina eschscholtzii*. *Journal of Biogeography*, **36**, 982–995.

Kuchta SR, Parks DS, Wake DB (2009b) Pronounced phylogeographic structure on a small spatial scale: Geomorphological evolution and lineage history in the salamander ring species *Ensatina eschscholtzii* in central coastal California. *Molecular Phylogenetics and Evolution*, **50**, 240–255.

Lapointe FJ, Rissler LJ (2005) Congruence, consensus, and the comparative phylogeography of codistributed species in California. *American Naturalist*, **166**, 290–299.

Leavitt DH, Bezy RL, Crandall KA, Sites Jr JW (2007) Multilocus DNA sequence data reveal a history of deep cryptic vicariance and habitat-driven convergence in the desert night lizard *Xantusia vigilis* species complex (Squamata: Xantusiidae). *Molecular Ecology*, **16**, 4455–4481.

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451–1452.

Lind AJ, Spinks PQ, Fellers GM, Shaffer HB (2011) Rangewide phylogeography and landscape genetics of the western U.S. endemic frog *Rana boylii* (Ranidae): implications for the conservation of frogs and rivers. *Conservation Genetics*, **12**, 269–284.

Lock J, Kelsey H, Furlong K, Woolace A (2006) Late Neogene and Quaternary landscape evolution of the northern California coast ranges: evidence for Mendocino triple junction tectonics. *GSA Bulletin*, **118**, 1232–1246.

Maddison WP, Maddison DR (2009) Mesquite: a modular system for Evolutionary Analysis. Version 2.6 Available from (http://mesquiteproject).

Martinez-Solano I, Lawson R (2009) Escape to Alcatraz: evolutionary history of slender salamanders (*Batrachoseps*) on the islands of San Francisco Bay. *BMC Evolutionary Biology*, **9**, 38.

Martinez-Solano I, Jockusch EL, Wake DB (2007) Extreme population subdivision throughout a continuous range: phylogeography of *Batrachoseps attenuatus* (Caudata: Plethodontidae) in western North America. *Molecular Ecology*, **16**, 4335–4355.

Matocq MD (2002) Phylogeographical structure and regional history of the dusky-footed woodrat, *Neotoma fuscipes*. *Molecular Ecology*, **11**, 229–242.

van der Meijden A, Chiari Y, Mucedda M, Carranza S, Corti C, Veith M (2009) Phylogenetic relationships of Sardinian cave salamanders, genus *Hydromantes*, based on mitochondrial and nuclear DNA sequence data. *Molecular Phylogenetics and Evolution*, **51**, 399–404.

Miller DD (1999) Sequence stratigraphy and controls on deposition of the upper Cenozoic Tulare Formation, San Joaquin Valley, California. PhD Thesis, Stanford University, Stanford, California. 1–179.

Miller RE, Buckley TR, Manos PS (2002) An examination of the monophyly of morning glory taxa using Bayesian phylogenetic inference. *Systematic Biology*, **51**, 740–753.

Moritz C, Schneider CJ, Wake DB (1992) Evolutionary relationships within the *Ensatina eschscholtzii* complex confirm the ring species interpretation. *Systematic Biology*, **41**, 273–291.

Normark WR (1999) Late Pleistocene channel-levee development on Monterey submarine fan, central California. *GeoMarine Letters*, **18**, 179–188.

Page BM, Thompson GA, Coleman RG (1998) Late Cenozoic tectonics of the central and southern Coast Ranges of California. *GSA Bulletin*, **110**, 846–876.

Papadopoulou A, Anastasiou I, Vogler AP (2010) Revisiting the insect mitochondrial molecular clock: the mid-Aegean trench calibration. *Molecular Biology & Evolution*, **27**, 1659–1672.

Parham JF, Papenfuss TJ (2009) High genetic diversity among fossorial lizard populations (*Anniella pulchra*) in a rapidly developing landscape (Central California). *Conservation Genetics*, **10**, 169–176.

Patton JL, Smith MF (1990) Evolutionary dynamics of the pocket gopher *Thomomys bottae*, with emphasis on California populations. *University of California Publications in Zoology*, **123**, vii–161.

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

Powell III CL, Barron JA, Sarna-Wojcicki AM *et al.* (2007) Age, stratigraphy, and correlations of the late Neogene Purisima Formation, central California Coast Ranges. *USGS Professional Paper*, **1740**, 1–32.

Rambaut A (2009) FigTree v1.3.1: Tree Figure Drawing Tool. Available from http://tree.bio.ed.ac.uk/software/figtree/.

Rambaut A, Drummond AJ (2007) Tracer v1.5, Available from http://beast.bio.ed.ac.uk/.

Ramirez M, Chi B (2004) Cryptic speciation, genetic diversity and gene flow in the California turret spider *Antrodiaetus riversi* (Araneae: Antrodiaetidae). *Biological Journal of the Linnean Society*, **82**, 27–37.

Raven RJ (1994) Mygalomorph spiders of the Barychelidae in Australia and the western Pacific. *Memoirs of the Queensland Museum*, **35**, 291–706.

Rissler LJ, Hijmans RJ, Graham CH, Moritz C, Wake DB (2006) Phylogeographic lineages and species comparisons in conservation analyses: A case study of California herpetofauna. *American Naturalist*, **167**, 655–666.

Rodriguez-Robles JA, Denardo DF, Staub RE (1999) Phylogeography of the California mountain kingsnake, *Lampropeltis zonata* (Colubridae). *Molecular Ecology*, **8**, 1923–1934.

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

Sarna-Wojcicki AM, Meyer CE, Bowman HR *et al.* (1985) Correlation of the Rockland Ash Bed, a 400,000-year-old stratigraphic marker in northern California and western Nevada, and implications of middle Pleistocene paleoge-

ography of central California. *Quaternary Research*, **23**, 236–257.
Satler J, Starrett J, Hayashi C, Hedin M (2011) Inferring species trees from gene trees in a radiation of California trapdoor spiders (Araneae, Antrodiaetidae, *Aliatypus*). *PLoS ONE*, **6**, e25355.
Schoenherr AA (1992) *A Natural History of California*. University of California Press, Berkeley, California.
Sims JD (1993) Chronology of displacement on the San Andreas fault in central California: evidence from reversed positions of exotic rock bodies near Parkfield, California. In: *The San Andreas fault system: Displacement, Palinspastic reconstruction, and Geologic Evolution* (eds Powell RE, Weldon RJ, Matti JC), pp. 231–256. Geological Society of America, Boulder, Colorado.
Sloan D (2006) *Geology of the San Francisco Bay Region*. University of California Press, Berkeley, California.
Spinks PQ, Thomson RC, Shaffer HB (2010) Nuclear gene phylogeography reveals the historical legacy of an ancient inland sea on lineages of the western pond turtle, Emys marmorata in California. *Molecular Ecology*, **19**, 542–556.
Starrett J, Hedin M (2007) Multilocus genealogies reveal multiple cryptic species and biogeographic complexity in the California turret spider *Antrodiaetus riversi* (Mygalomorphae, Antrodiaetidae). *Molecular Ecology*, **16**, 583–604.
Stockman AK, Bond JE (2007) Delimiting cohesion species: extreme population structure and the role of ecological interchangeability. *Molecular Ecology*, **16**, 3374–3392.
Swofford DL (2002) *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4*. Sinauer Associates, Sunderland, Massachusetts.
Vandergast AG, Bohonak AJ, Hathaway SA, Boys J, Fisher RN (2008) Are hotspots of evolutionary potential adequately protected in southern California? *Biological Conservation*, **141**, 1648–1664.
Vincent LS (1993) The natural history of the California turret spider *Atypoides riversi* (Araneae, Antrodiaetidae): demographics, growth rates, survivorship, and longevity. *Journal of Arachnology*, **21**, 29–39.
Wakabayashi J, Sawyer TL (2001) Stream incision, tectonics, uplift, and evolution of topography of the Sierra Nevada, California. *Journal of Geology*, **109**, 539–562.
Wake D (1997) Incipient species formation in salamanders of the *Ensatina* complex. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 7761–7767.
Wake DB (2009) What salamanders have taught us about evolution. *Annual Review of Ecology, Evolution & Systematics*, **40**, 333–352.
Waters JM (2011) Competitive exclusion: phylogeography's "elephant in the room". *Molecular Ecology*, **20**, 4388–4394.

M.H. and J.S. conducted fieldwork, performed lab work, and analyzed data. All authors helped to write and edit the manuscript.

Data accessibility

DNA sequences: GenBank numbers JX951761-JX952143

Google Earth kmz file: DRYAD data entry doi:10.5061/dryad.121tq.

BEAST xml file: DRYAD data entry doi:10.5061/dryad.121tq.

Sample location information: online supporting information.

Supporting information

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

Data S1 Gene data collection, sequence alignment & phylogenetic analyses.

Fig. S1 (a) EF1G Bayesian consensus phylogram (b) Anonymous locus Bayesian consensus phylogram. Strongly supported bipartitions (posterior probability > 0.95) indicated by thickened branches. Support values for other relevant bipartitions indicated by exact posterior probability values.

Fig. S2 (a) Posterior distribution of theta and *M* for Monterey clade (replicate run one); (b) Posterior distribution of theta and *M* for Valley clade (replicate run one).

Table S1 Clade membership, location information, and GenBank accession numbers.