Complex phylogeography and historical hybridization between sister taxa of freshwater sculpin (Cottus)

4 AUTHORS, INCLUDING:

Jason Baumsteiger
University of California, Davis

20 PUBLICATIONS **136** CITATIONS

SEE PROFILE

Andrew P. Kinziger
Humboldt State University

27 PUBLICATIONS **222** CITATIONS

SEE PROFILE

Andres Aguilar
California State University, Los Angeles

29 PUBLICATIONS **438** CITATIONS

SEE PROFILE

Complex phylogeography and historical hybridization between sister taxa of freshwater sculpin (*Cottus*)

JASON BAUMSTEIGER,* ANDREW P. KINZIGER,† STEWART B. REID‡ and ANDRES AGUILAR*[1]

*School of Natural Sciences & Sierra Nevada Research Institute, University of California Merced, 5200 N. Lake Rd., Merced, CA 95343, USA, †Department of Fisheries Biology, Humboldt State University, 1 Harpst Street, Arcata, CA 95521, USA, ‡Western Fishes, 2045 East Main Street, Ashland, OR 97520, USA

Abstract

Species ranges that span different geographic landscapes frequently contain cryptic species- or population-level structure. Identifying these possible diversification factors can often be accomplished under a comparative phylogeographic framework. However, comparisons suffer if previous studies are limited to a particular group or habitat type. In California, a complex landscape has led to several phylogeographic breaks, primarily in terrestrial species. However, two sister taxa of freshwater fish, riffle sculpin (*Cottus gulosus*) and Pit sculpin (*Cottus pitensis*), display ranges based on morphological identifications that do not coincide with these breaks. Using a comprehensive sampling and nuclear, mitochondrial and microsatellite markers, we hypothesized that proposed species ranges are erroneous based on potential hybridization/gene flow between species. Results identified a phylogeographic signature consistent with this hypothesis, with breaks at the Coast Range Mountains and Sacramento/San Joaquin River confluence. Coastal locations of *C. gulosus* represent a unique lineage, and 'true' *C. gulosus* were limited to the San Joaquin basin, both regions under strong anthropogenic influence and potential conservation targets. *C. pitensis* limits extended historically throughout the Sacramento/Pit River basin but currently are restricted to the Pit River. Interestingly, locations in the Sacramento River contained low levels of ancestral hybridization and gene flow from *C. gulosus* but now appear to be a distinct population. The remaining population structure was strongly correlated with Sierra Nevada presence (high) or absence (low). This study stresses the importance of testing phylogeographic breaks across multiple taxa/habitats before conservation decisions are made, but also the potential impact of different geographic landscapes on evolutionary diversification.

Keywords: biogeography, California, diversification, introgression

Received 13 December 2013; revision received 11 April 2014; accepted 13 April 2014

Introduction

Complex geographic landscapes often lead to high rates of diversification and similarly complex population structure due to the creation of variable ecological conditions and evolutionary selective pressures (Grinnell 1924; Peterson *et al.* 1999; Thompson 1999). Yet, these landscapes are as dynamic as the ecological and evolutionary parameters they create, constantly being modified over time and space, leading to not just creation but occasional blurring of species boundaries through historical/contemporary hybridization and introgression (Harrison 1993; Swenson & Howard 2005; Bryson *et al.* 2010). Recent anthropogenic modifications have only created additional complexity in the biogeographic landscape (Turner 1989; Wu 2013). Therefore, disentangling the impact of these processes is essential to understand and conserve existing flora and fauna.

Species identifications and their distributions have traditionally been through morphological traits (Dayrat

Correspondence: Jason Baumsteiger, Fax: 209-228-4060;
E-mail: jb4799@gmail.com
[1]Present address: Department of Biological Sciences, California State University, 5151 State University Drive, Los Angeles, CA 90032-8530, USA

2005). However, not all evolutionary changes manifest as observable traits (Avise 1989). Physiological adaptations, stochastic changes through genetic drift, pre- and postzygotic barriers and past introgressive events can all lead to 'cryptic' species- and population-level differences (Bickford *et al.* 2007). By approaching species diversification, range limits and population structure with genetic analyses, important evolutionary information is made available. Thus, a phylogeographic approach is warranted, with direct correlations drawn between the physical geography and genetic information over time (Avise *et al.* 1987; Avise 2000).

As the field of phylogeography continues to improve, novel approaches are needed to ensure that patterns observed between geography and genetic analyses are causative, not merely corroborative (Hickerson *et al.* 2010). One approach is to use a comparative phylogeographic framework to validate the impact of a particular geographic landscape on multiple, often unrelated species (Taberlet *et al.* 1998; Arbogast & Kenagy 2001; Soltis *et al.* 2006). However, it is unclear whether observed phylogeographic patterns are consistent across organisms from terrestrial and aquatic habitats. Continued reappraisal of proposed phylogeographic breaks across habitats is needed for a comparative framework to be effective. Once identified, proposed breaks could have strong conservation implications, allowing managers to better understand and conserve regional boundaries and important evolutionary hot spots (da Silva & Patton 1998).

In California, a complex geographic history has combined with extensive contemporary processes to shape species ranges and population structure (Calsbeek *et al.* 2003; Shaffer *et al.* 2004; Burns & Barhoum 2006; Starrett & Hedin 2006). Volcanism (Pease 1965), glaciation (Mulch *et al.* 2008), marine incursions (Dupré *et al.* 1991) and mountain formation (Howard 1979) within the last five million years (Dupré *et al.* 1991), along with more recent anthropogenic changes (Leu *et al.* 2008), have all been shown to impact species in the region. Numerous studies have investigated the link between these processes and the distribution of terrestrial organisms (Wake 1997; Lapointe & Rissler 2005; Spinks *et al.* 2010) as terrestrial organisms often migrate according to specific landscape variables (Manel *et al.* 2003; Storfer *et al.* 2010). Yet, few (King *et al.* 1996) have looked at species in freshwater environments, where organisms occur in discrete riverine or lacustrine networks (Poissant *et al.* 2005; Hughes *et al.* 2009).

The endemic freshwater fauna along the western drainage of the Sierra Nevada Mountains (Hershler 1995; Light *et al.* 2002; Moyle 2002) is a product of a dynamic geographic landscape (MacDonald & Gay 1968; Mount 1995) and an extensive network of dams and water diversions essential to water resource management (Moyle & Williams 2005). Two freshwater species of fish, riffle sculpin (*Cottus gulosus*) and Pit sculpin (*C. pitensis*), do not show ranges or population structure consistent with previously identified phylogeographic breaks in this region (Moyle 2002). These sister taxa (see Baumsteiger *et al.* 2012) currently exhibit a parapatric distribution based on morphological characteristics (Page & Burr 2011). *Cottus gulosus* has a disjunct distribution, occurring in small coastal streams in central California and the foothill and high-elevation headwaters of most tributaries to the Sacramento and San Joaquin Rivers. In contrast, *Cottus pitensis* is confined to the Pit River, a tributary to the upper Sacramento River (Moyle 2002; Page & Burr 2011). Phylogenetic analyses show that both species are part of an endemic western North American clade of the genus *Cottus* (Kinziger *et al.* 2005). This clade has recently been suspected of containing a number of cryptic species (Baumsteiger *et al.* 2012) and potential hybrids (Moyle 2002). Freshwater sculpin, as a whole, appear to readily hybridize, and in some cases, this hybridization has led to hybrid speciation (Hunt *et al.* 1997; Kinziger & Raesly 2001; Kontula *et al.* 2003; Nolte *et al.* 2005).

We used a multitiered genetic approach, incorporating nuclear and mitochondrial sequence markers and microsatellites, to assess the phylogeographic patterns in *C. gulosus* and *C. pitensis*. Using a comprehensive sampling scheme, phylogenetic analysis, gene flow measurements and several fine-scale population structure analyses, we tested the following hypotheses: (i) *C. gulosus* isolated along the coast of central California is a cryptic species; (ii) phylogeographic breaks associated with both species of sculpin coincide with other phylogeographic studies in the region and not currently identified species limits; (iii) regional phylogeographic breaks are similar between terrestrial and aquatic species; (iv) hybrids will be found at the parapatric boundary between species; and (v) population structure will be associated with the presence or absence of Sierra Nevada Mountain geology. Our findings have direct conservation implications, not just for each sculpin species, but regionally as important phylogeographic boundaries are further established and new species identified. Given the importance of these riverine systems to the anthropogenic water needs of the state, results could have impacts on future water management decisions.

Materials and methods

Sampling

Using backpack electroshocking and seine nets, one to 50 specimens were collected at nine locations for

C. pitensis and 24 locations for *C. gulosus*, including five disjunct locations from the central California coast (Table 1 and Fig. 1). Caudal fin clips and whole individuals were collected and stored in ethanol prior to DNA extraction, with whole individuals vouchered into the Humboldt State University Fish Collection. DNA was extracted from fin clips using the DNeasy tissue extraction kit (Qiagen Inc.) and quantified on a Nano-Drop 2000 spectrophotometer (Thermo Scientific), producing sufficient quantities (25–150 ng/µL) of pure DNA for amplification and sequencing.

Sequencing marker development

Three nuclear sequence markers (Locus 508, 517 and 520) and one mitochondrial marker (cytochrome b – *cytb*)

were assayed (Baumsteiger *et al.* 2012) (Appendix S1, Supporting information). Gene-specific alignments were performed with MUSCLE (Edgar 2004) and all sequences deposited in GenBank (accession nos: JX484692–JX484693, KJ509201–KJ509583). Individual loci (nuclear and mitochondrial) were tested for adherence to a molecular clock using the likelihood scores function (LSCORES) in PAUP* (Swofford 2002). Additionally, we examined each nuclear locus for recombination using GARD (Genetic Algorithm for Recombination Detection), found on the online server DATAMONKEY (Pond *et al.* 2006).

Unique haplotypes for mitochondrial sequences were obtained with DNASP (Rozas *et al.* 2003; Librado & Rozas 2009). Nuclear sequence haplotypes were identified using PHASE (Stephens *et al.* 2001; Stephens & Scheet 2005) and later confirmed with DNASP. Nucleotide

Table 1 Sampling locations and numbers of individuals covering known ranges of each species

Pop no.	Sites	Collected as	N	Location, County	Latitude	Longitude
1	Pit_Goose*	*C. pitensis*	21	Lassen Cr., Modoc Co.	41.8296	−120.2983
2	Pit_SF*	*C. pitensis*	29	South Fork Pit R., Lassen Co.	41.2325	−120.3451
3	Pit_Ash*	*C. pitensis*	28	Ash Cr., Modoc Co.	41.1592	−120.8285
4	Pit_Fall[†]	*C. pitensis*	30	Bear Cr., Fall R., Shasta Co.	41.1902	−121.7356
5	Pit_1	*C. pitensis*	26	Pit R. #1, Shasta Co.	40.9919	−121.5076
6	Pit_Hat[†]	*C. pitensis*	10	Hat Cr., Shasta Co.	40.9809	−121.5712
7	Pit_Rock	*C. pitensis*	25	Rock Cr., Shasta Co.	41.0107	−121.7045
8	Pit_Clark[†]	*C. pitensis*	18	Clark Cr., Shasta Co.	40.9853	−121.7766
9	Pit_5	*C. pitensis*	30	Pit R. #5, Shasta Co.	41.0019	−121.9667
10	Sacramento_Dunsmuir[†]	*C. gulosus*	13	Big Springs Cr., Siskiyou Co.	41.3106	−122.3303
11	Sacramento_Cantara	*C. gulosus*	24	Sacramento R. (Cantara), Siskiyou Co.	41.2661	−122.3078
12	Sacramento_Sims	*C. gulosus*	24	Sacramento R. (Sims), Siskiyou Co.	41.0631	−122.3599
13	Sacramento_Clear[‡]	*C. gulosus*	31	Clear Cr., Shasta Co.	40.4990	−122.4150
14	Sacramento_Battle[‡]	*C. gulosus*	27	Battle Cr., Tehama Co.	40.3983	−122.1500
15	Sacramento_RBDD[‡]	*C. gulosus*	43	Red Bluff Diversion Dam, Shasta Co.	40.1550	−122.2050
16	Sacramento_RM205[‡]	*C. gulosus*	21	Sacramento R. (RM205.5), Colusa Co.	39.7910	−122.0350
17	Feather[‡]	*C. gulosus*	28	Feather R., Butte Co.	39.3630	−121.6010
18	NMF_American	*C. gulosus*	25	North/Middle F American R., Placer Co.	38.9168	−121.0370
19	SF_American	*C. gulosus*	24	South Fork American R., El Dorado Co.	38.8078	−120.9001
20	Lower_American	*C. gulosus*	16	American R. (Nimbus), Sacramento Co.	38.6343	−121.2243
21	Mokelumne	*C. gulosus*	25	Mokelumne R., San Joaquin Co.	38.3159	−120.7104
22	Stanislaus	*C. gulosus*	22	Stanislaus R., Stanislaus Co.	38.1430	−120.3760
23	Lower_Tuolumne	*C. gulosus*	27	Tuolumne R., Stanislaus Co.	37.6642	−120.4548
24	Upper_ Tuolumne	*C. gulosus*	22	Tuolumne R., Tuolumne Co.	37.8750	−119.9630
25	Merced	*C. gulosus*	30	Merced R., Mariposa Co.	37.6532	−119.7825
26	Lower_Kings	*C. gulosus*	89	Kings R., Kings Co.	36.7833	−119.4167
27	Upper_Kings	*C. gulosus*	30	Kings R., Fresno Co.	36.8590	−119.0958
28	Kaweah	*C. gulosus*	16	Kaweah R., Tulare Co.	36.4567	−118.8512
29	Russian[†]	*C. gulosus*	2	Russian R., Sonoma Co.	38.4844	−122.8244
30	Guadalupe[§]	*C. gulosus*	31	Guadalupe Cr., Santa Clara Co.	37.2250	−121.9050
31	Penitencia[§]	*C. gulosus*	30	Penitencia Cr., Santa Clara Co.	37.3980	−121.8000
32	Uvas[§]	*C. gulosus*	29	Uvas Cr., Santa Clara Co.	37.0120	−121.6270
33	Bird	*C. gulosus*	27	Bird Cr., San Benito Co.	36.7803	−121.4011

*Samples obtained with the assistance of Stewart Reid of Western Fishes.
[†]Samples obtained from Humboldt State University Fish Collection or personal collection of Andrew Kinziger.
[‡]Samples collected by U.S. Fish and Wildlife Service personnel.
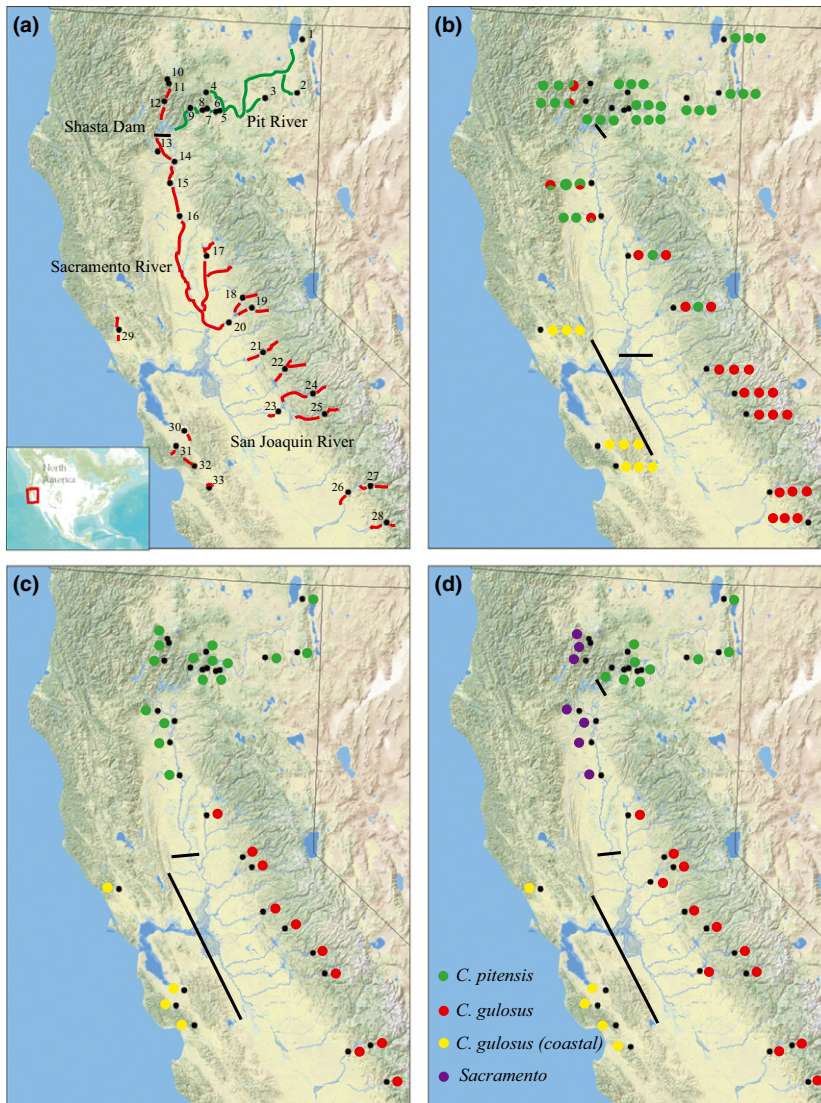[§]Samples collected by Jerry Smith and students, San Jose State University.

Fig. 1 Collection location (black dots) and distribution map for (a) contemporary species ranges, (b) nuclear haplotypes from each individual nuclear locus (508, 517 and 520, reading left to right) according to three individuals per location, (c) mitochondrial haplotypes and (d) putative populations according to clustering obtained in STRUCTURE for six microsatellite loci. Black bars indicate possible breaks in species/populations, and colours are indicative of species haplotypes: green (*Cottus pitensis*), red (*C. gulosus*), yellow (coastal *Cottus gulosus*) and purple (Sacramento basin *C. pitensis*).

substitution testing was conducted with JMODELTEST v0.1.1 (Posada 2008) using the corrected Akaike information criterion (AIC$_c$) of Posada & Buckley (2004). Haplotype networks were inferred by locus with TCS v1.21 (Clement *et al.* 2000). Numbers of segregating sites (*S*), number of haplotypes (*H*), haplotype diversity (*H*$_d$) and nucleotide diversity (π) were generated with DNASP (Table 2).

Phylogenetic and coalescence-based analysis

A comprehensive species tree approach employing information from three nuclear markers (508, 517 and 520) was conducted using the multispecies coalescence-based analysis in *BEAST v.1.6 (Drummond & Rambaut 2007) to estimate evolutionary relationships between individuals at different locations on a nuclear-DNA-only timescale. An uncorrelated lognormal molecular clock was employed using the default distribution for the rate parameter, following Baumsteiger *et al.* (2012). Runs consisted of 100 million iterations sampled every 10 000th iteration, resulting in 7500 trees after removing the first 25% as burn-in. TREEANNOTATOR (Drummond & Rambaut 2007) was used to obtain a consensus tree representing the most consistent topology over all trees. A traditional ≥ 0.95 posterior probability was used to define a well-supported branch or node.

Estimates of divergence times were generated based upon a standardized molecular clock substitution rate in *cytb*, as nuclear sequences exhibited too little variation and a lack of fossil evidence precluded external calibration. A conservative estimate of mutation rates was calculated at 0.9% per million years (0.5–1.3%) based on Bermingham *et al.* (1997) and Mueller (2006). Estimation of time of divergence for *C. gulosus* and *C. pitensis* was

Table 2 Descriptive statistics, by location, for three nuclear (508, 517 and 520), one mitochondrial (*cytb*) and six microsatellite (μsats) loci. Number of individuals (N), substitutions (S), nucleotide diversity (π), number of haplotypes (H), haplotype diversity (H_d), allele number corrected for rarefaction (A_R), expected heterozygosity (H_E), observed heterozygosity (H_O) and inbreeding coefficient (F_{IS})

Sites	508 N	S	π	H	H_d	517 N	S	π	H	H_d	520 N	S	π	H	H_d	CB N	S	π	H	H_d	μsats N	A_R	H_E	H_O	F_{IS}
Pit_Goose	3	1	0.000	1	0.000	3	0	0.000	1	0.000	3	0	0.000	1	0.000	8	0	0.000	1	0.000	21	1.41	0.250	0.136	0.462*
Pit_SF	3	1	0.001	2	0.600	3	0	0.000	1	0.000	3	0	0.000	1	0.000	8	0	0.000	1	0.000	29	1.28	0.108	0.111	−0.029
Pit_Ash	3	1	0.000	1	0.000	3	0	0.000	1	0.000	3	0	0.000	1	0.000	5	0	0.000	1	0.000	28	1.28	0.100	0.088	0.119
Pit_Fall	1	1	0.000	1	0.000	1	0	0.000	1	0.000	1	0	0.000	1	0.000	8	0	0.000	1	0.000	30	1.03	0.011	0.011	0.000
Pit_1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	6	1	0.001	2	0.533	26	1.49	0.149	0.149	0.004
Pit_Hat	—	—	—	—	—	—	—	—	—	—	1	1	0.003	2	1.000	2	1	0.001	2	1.000	10	1.33	0.075	0.056	0.273
Pit_Rock	1	1	0.000	1	0.000	1	0	0.000	1	0.000	1	0	0.000	1	0.000	8	0	0.000	1	0.000	25	1.34	0.129	0.136	−0.060
Pit_Clark	1	1	0.000	1	0.000	1	0	0.000	1	0.000	3	1	0.001	2	0.533	7	3	0.001	4	0.714	18	1.41	0.138	0.146	−0.057
Pit_5	3	1	0.000	1	0.000	3	0	0.000	1	0.000	—	—	—	—	—	8	0	0.000	1	0.000	30	1.55	0.182	0.197	−0.078
Sac_Dunsmuir	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	2	1	0.001	2	1.000	13	2.33	0.434	0.323	0.256*
Sac_Cantara	3	1	0.001	2	0.533	3	1	0.003	2	0.600	3	0	0.000	1	0.000	7	0	0.000	1	0.000	24	1.27	0.121	0.091	0.251
Sac_Sims	3	1	0.001	2	0.533	3	1	0.003	2	0.533	3	0	0.000	1	0.000	8	0	0.000	1	0.000	24	1.44	0.202	0.223	−0.105
Sac_Clear	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	13	3	0.001	2	0.282	31	1.77	0.292	0.279	0.046
Sac_Battle	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	6	0	0.000	1	0.000	27	2.25	0.414	0.404	0.024
Sac_RBDD	3	3	0.002	3	0.600	3	0	0.000	1	0.000	3	2	0.003	2	0.533	8	3	0.001	4	0.750	43	2.25	0.400	0.303	0.245*
Sac_RM205	3	3	0.003	2	0.533	3	1	0.003	2	0.533	3	1	0.001	2	0.333	8	6	0.002	5	0.786	21	2.07	0.389	0.282	0.285*
Feather	3	0	0.000	1	0.000	3	1	0.003	2	0.600	3	0	0.000	1	0.000	9	6	0.002	6	0.833	28	6.88	0.467	0.475	−0.016
NMF_American	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	8	0	0.000	1	0.000	25	1.36	0.135	0.133	0.013
SF_American	3	0	0.000	1	0.000	3	0	0.000	1	0.000	3	0	0.000	1	0.000	7	2	0.001	3	0.524	24	1.14	0.047	0.028	0.418
Lower_American	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	16	3.33	0.456	0.428	0.064
Mokelumne	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	8	0	0.000	1	0.000	25	1.16	0.082	0.080	0.020
Stanislaus	3	0	0.000	1	0.000	3	0	0.000	1	0.000	3	0	0.000	1	0.000	8	0	0.000	1	0.000	22	1.29	0.108	0.129	−0.194
Lower_Tuolumne	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	27	2.67	0.341	0.302	0.115
Upper_Tuolumne	3	0	0.000	1	0.000	3	0	0.000	1	0.000	3	0	0.000	1	0.000	7	1	0.001	2	0.476	22	1.16	0.147	0.090	0.395
Merced	3	0	0.000	1	0.000	3	0	0.000	1	0.000	3	0	0.000	1	0.000	8	0	0.000	1	0.000	30	6.01	0.099	0.093	0.060
Lower_Kings	3	0	0.000	1	0.000	3	0	0.000	1	0.000	3	2	0.003	3	0.800	10	1	0.000	2	0.200	89	1.25	0.111	0.095	0.149
Upper_Kings	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	7	1	0.000	2	0.286	30	1.27	0.117	0.122	−0.049
Kaweah	3	0	0.000	1	0.000	3	0	0.000	1	0.000	3	0	0.000	1	0.000	8	3	0.001	4	0.643	16	1.36	0.115	0.049	0.673*
Russian	2	1	0.000	1	0.000	2	0	0.000	1	0.000	2	0	0.000	1	0.000	2	4	0.004	2	1.000	2	1.67	0.417	0.083	0.857*
Guadalupe	4	1	0.000	1	0.000	4	0	0.000	1	0.000	4	9	0.003	2	0.429	8	3	0.001	3	0.464	31	1.15	0.060	0.059	0.015
Penitencia	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	8	5	0.003	5	0.893	30	2.64	0.080	0.067	0.164
Uvas	2	1	0.000	1	0.000	2	0	0.000	1	0.000	2	0	0.000	1	0.000	7	0	0.000	1	0.000	29	1.76	0.328	0.278	0.136
Bird	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	27	1.31	0.108	0.019	0.830*
Total/Mean	56	14	0.007	9	0.752	56	7	0.005	4	0.206	56	17	0.007	8	0.678	217	87	0.018	44	0.895	873	2.18	0.201	0.166	0.169

*Outside HWE ($P < 0.01$).

obtained with BEAST, beginning with assumptions of a constant population size, strict molecular clock and uniform priors. The final Bayesian distribution was obtained using 100 million iterations sampled every 10 000th iteration. Different priors, including calculated upper and lower bounds for the molecular clock, were modified until effective sample sizes and other measures were within acceptable ranges. A burn-in of 25% was sufficient to allow for convergence of the MCMC chains. A closely related sympatric species, prickly sculpin (*Cottus asper*), was included as in Baumsteiger *et al.* (2012).

Gene flow between species

We used the isolation with migration approach implemented in the program IMA2 to examine historical vs. contemporary gene flow between *C. gulosus* and *C. pitensis* (Hey & Nielsen 2007). We used nuclear DNA sequences and assumed a species break consistent with the mitochondrial DNA. Three independent runs were performed of 250 000 burn-in steps followed by 2.5×10^6 steps sampled every 100 steps (25 000 total). Runs were highly similar and combined to estimate migration parameters. A generation time of 3 years was assumed based on life history data from other closely related cottids (Moyle 2002).

Microsatellite markers

Seven microsatellite loci were chosen for this study based on previous work in the closely related *C. asper* (Baumsteiger & Aguilar 2013 and Appendix S1, Supporting information). Markers were run for sampling locations with at least 10 individuals (Table 1). Descriptive statistics were generated using the software GDA (Lewis & Zaykin 2001) and ARLEQUIN v3.5 (Excoffier *et al.* 2005; Excoffier & Lischer 2010). Allelic richness was standardized to a sample size of 10 using HP-RARE (Kalinowski 2005). Correction for multiple tests used the Bonferroni method (Rice 1989). Pairwise estimates of population differentiation (F_{ST}, significance compared over 100 permutations) were estimated using GENETIX v4.05 (Belkhir *et al.* 1996-2004), and significant genetic differentiation among population groups was evaluated using an analysis of molecular variance (AMOVA) in ARLEQUIN.

A Bayesian clustering approach was employed in STRUCTURE (Pritchard *et al.* 2000) to assess population structure. An initial burn-in of 100 000 was followed by 100 000 runs using an admixture model, with allele frequencies assumed to be correlated, K (clusters) ranging from 1 to 25, and 10 iterations per K. The number of genetically distinct clusters was determined by Evanno *et al.* (2005) or by plotting log-likelihood values vs. K using STRUCTURE HARVESTER (Earl & vonHoldt 2012). To average over the number of iterations of a chosen K, CLUMP (Jakobsson & Rosenberg 2007) was run using the Greedy algorithm over 1000 replicates. Additional K values were tested to insure that the best K was obtained. Visual presentation of clustering was prepared in DISTRUCT (Rosenberg 2003).

Results

A total of 872 individuals were assayed, including 655 *C. gulosus* and 217 *C. pitensis* at 24 and nine locations, respectively (Fig. 1a and Table 1). For the nuclear and mitochondrial sequencing, a subset of individuals was assayed per location to reduce costs, with a single location chosen when two locations were close to one another (Table 2). At each site, one to four individuals were sequenced for each nDNA marker and about eight individuals (range 2–13) for mitochondrial DNA, minus Bird Creek and the lower Tuolumne and American Rivers (Table 2).

Sequence substitution model testing was consistent with the HKY model (Hasegawa *et al.* 1985) for locus 508, the K80 model (Kimura 1980) for loci 517 and 520, and the TrN + G (gamma) model (Tamura & Nei 1993) for *cytb*. No recombination was evident for any of the three nuclear loci. Descriptive statistics for the nDNA sequences exhibited low nucleotide and haplotype diversity, with most locations containing a single haplotype (Table 2). The highest nuclear sequence variation occurred in the Sacramento River, with three nuclear haplotypes at the Red Bluff diversion dam. Identical patterns were evident in mitochondrial haplotypes, with highest variation in the Sacramento River (Table 2). Pit River collection sites exhibited low nucleotide and haplotype diversity, with the exception of Clark Creek (Table 2).

nDNA sequences

Species tree analysis with BEAST using nuclear DNA resolved three primary clades (Fig. 2). A coastal *C. gulosus* clade was found to be basal to all other locations, although branch support was relatively low (posterior probability = 0.845). Pit River basin locations form a moderately supported (0.906) monophyletic clade, although no additional clades are supported within the basin itself. An inland *C. gulosus* clade representing locations south of the American River (Fig. 1a) was discovered with low support (0.814). Sacramento basin locations were intermediate between the *C. pitensis* and inland *C. gulosus*.

Using the mtDNA species break associated with the mouth of the Feather River (Fig. 1c and Table S2, Supporting information), nuclear haplotypes were
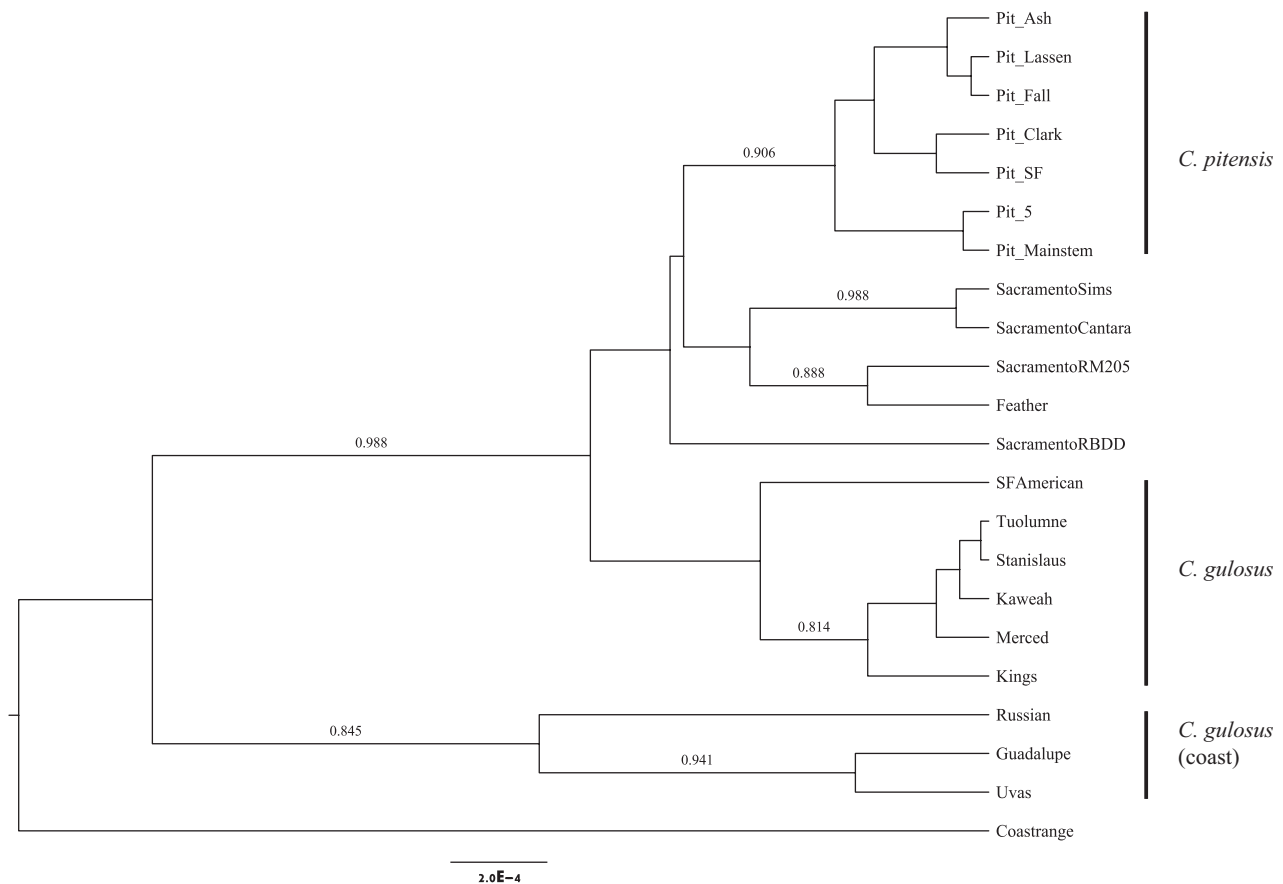
Fig. 2 Bayesian phylogenetic species tree developed from three nuclear loci. A single coastrange sculpin (*C. aleuticus*) served as an out-group (Baumsteiger *et al.* 2012). Three clades are evident: coastal *Cottus gulosus*, *C. gulosus* and *Cottus pitensis*. Individuals from locations within the Sacramento River are intermediate between *C. gulosus* and *C. pitensis* clades. Branch support indicated by posterior probabilities.

assigned to prospective species. Haplotype networks generated from unique nDNA sequences revealed similar patterns to the species tree, although species-specific haplotypes belonging to *C. pitensis* and *C. gulosus* were found in some locations (Fig. 1b and Fig. S1, Supporting information). Locations from the Sacramento River basin, Feather River and American River all contain haplotypes from both species (Fig. S1 and Table S1, Supporting information).

Three independent IMA2 runs converged on the same parameter estimates. We focused exclusively on the estimate of asymmetric migration rates to assess migration rates between the two species. Results indicate that low levels of historical migration occurred from inland *C. gulosus* to *C. pitensis* ($2Nm_{Riffle} > 2Nm_{Pit} = 0.3231$; 95% HPD = 0.001–0.0884) while migration in the alternate direction was not different from zero ($2Nm_{Pit} > 2Nm_{Riffle} = 0.0018$; 95% HPD = 0–0.0595). Based on likelihood ratio tests, the population-level migration rates from inland *C. gulosus* to *C. pitensis* were statistically significant (LLR = 4.32; $P < 0.05$) and those

from *C. pitensis* to inland *C. gulosus* were not significant (LLR = 0.0 P = n.s.).

mtDNA sequences

Analysis of mtDNA resolved the same three major clades identified using nDNA (Fig. 1c). Individual species or lineages contained sufficient differences between sequences to warrant individual haplotype networks (> 15 changes between clades) (Fig. 4). In contrast to nDNA, locations north of the Feather River contained *C. pitensis* mtDNA haplotypes, while locations south contained *C. gulosus* haplotypes (Fig. 1c). Coastal *C. gulosus* showed similar levels of mtDNA diversity to the other lineages, with nine haplotypes distributed over three locations. Each inland *C. gulosus* location contained at least one distinct haplotype, whereas most *C. gulosus* locations were dominated by a single haplotype (Table S2 and Fig. 3).

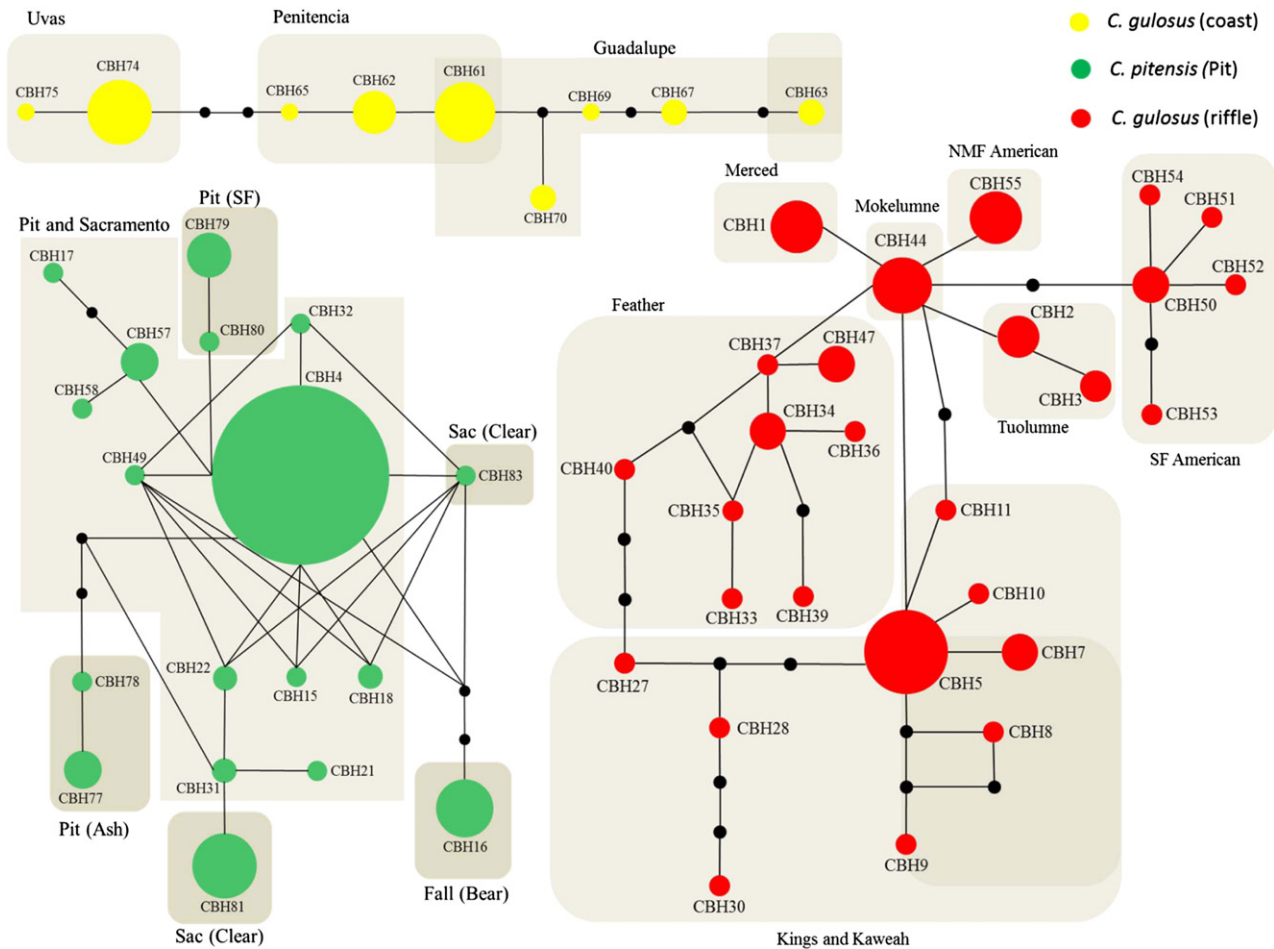Using a generalized mitochondrial clock and assuming a coalescence based on constant population size,

Fig. 3 mtDNA haplotype networks from unique *cytb* sequences for each of three putative species (*Cottus gulosus* – riffle; *Cottus pitensis* – Pit; and *C. gulosus* – coast). Circle size correlates with number of individuals, with the circle in the legend approximately equal to one individual. Grey outlines represent locations containing a particular haplotype. Grey unlabelled box in coastal *C. gulosus* represents haplotype also found in Penitencia Creek. The grey box imbedded in the Kings River group is the Kaweah individuals. Black dots represent inferred missing haplotypes. A complete list of haplotypes (CBH), by location, can be found in Table S2.

BEAST estimated divergence at 3.29 Ma (1.46–6.32 Ma; 95% C.I.) between inland *C. gulosus* and *C. pitensis* (Fig. 4). The remaining phylogenetic tree was similar to species tree and haplotype predictions generated using nuclear sequence data.

Microsatellites

The average number of individuals genotyped per location was 26 (range: 10–89, Table 2). Allelic richness was low after correcting for sample size (rarefaction = 10 individuals), averaging 2.18. Some exceptions were noted in the Feather River and Merced River (Table 2). Most locations conform to Hardy–Weinberg expectations ($P > 0.05$), as seen in reported F_{IS} values, with the exception of Bird Creek, Pit_Goose, Sac_Dunsmuir and the Kaweah River (Table 2). Observed heterozygosity

values tended to be low, with no value higher than 0.475.

Bayesian clustering analyses with STRUCTURE suggested a conservative estimate of $K = 5$ using the *ad hoc* approach of Evanno *et al.* (2005), whereas plots of the log-likelihood values vs. K suggested K was 8 or 9 (Fig. S2, Supporting information). When $K = 5$, five distinct groups are identified, including three previously identified clades (coastal *C. Gulosus,* inland *C. gulosus* and *C. pitensis*) and groups restricted to the Sacramento River and Kings River (Fig. 1d and Fig. 5). When $K = 8$ or 9 is used, subclusters within the Sacramento River were evident above (10–12) and below (13–16) Shasta Reservoir (as numbered in Table 1). Additionally, inland *C. gulosus* locations contained four subclusters: (i) Feather and lower American; (ii) Stanislaus, upper Tuolumne and Merced; (iii) SF and NMF American,
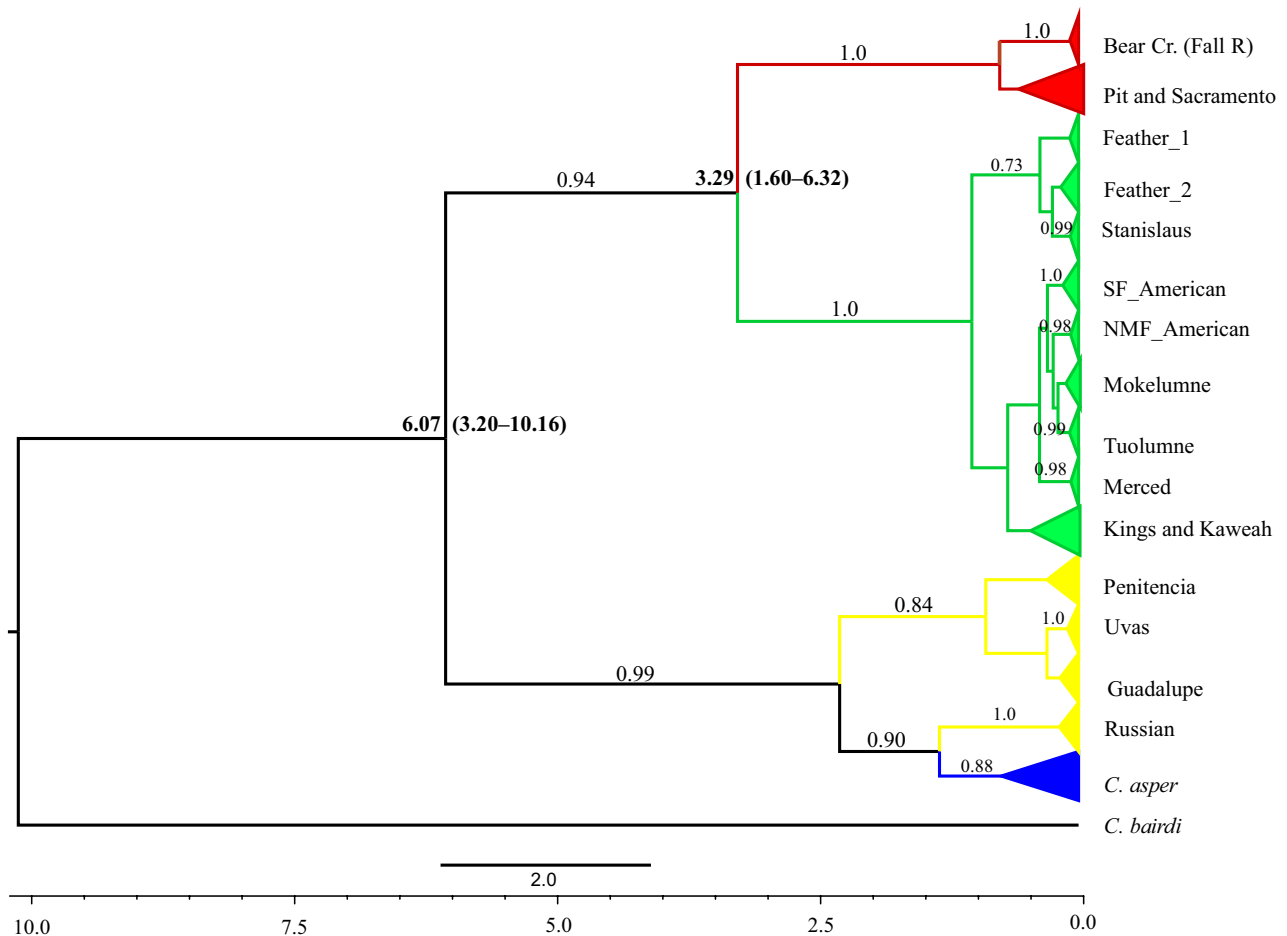
Fig. 4 Mitochondrial coalescent tree estimated in BEAST, showing age of diversification based on a 0.9% mutation rate. Parentheses represent upper and lower bounds of estimate. Numbers above branches are posterior probabilities based on 37 500 trees. Colours match Fig. 1 distinctions, with the exception of prickly sculpin (*Cottus asper*), shown in blue.

Mokelumne, lower Tuolumne and Kaweah; and (iv) upper and lower Kings. Finally, when $K = 9$, additional clustering is found in *C. pitensis*, separating locations 1–3 from 4 to 9 (Fig. 5).

Unrooted neighbour-joining trees using Cavalli-Sforza Edwards chord distance, Nei's distance and Reynolds distance found similar clustering to STRUCTURE (Fig. S3, Supporting information). Four major groups are observable: (i) Pit River, (ii) Sacramento group, (iii) San Joaquin group, and (iv) Coastal group. With the exception of Sacramento_Dunsmuir, a general trend of increasing branch lengths is observed from northernmost locations to southernmost locations in the chord distance tree.

In pairwise comparisons between location and population differentiation (F_{ST}), most locations significantly diverged in allele frequencies from one another in the inland *C. gulosus* ($P < 0.001$), whereas only some differences were significant in the Sacramento cluster and even fewer in the *C. pitensis* cluster (with the exception of Bear Creek) (Table S3). Isolation by distance (IBD)

was tested in *C. pitensis*, Sacramento and inland *C. gulosus*, with coastal *C. gulosus* excluded from this analysis (Table S3 and Fig. S4, Supporting information). Comparing F_{ST} vs. distance (km) using a Mantel test revealed no isolation by distance for *C. pitensis* ($r = 0.37$, $P = 0.1$), but significant IBD for Sacramento ($r = 0.466$, $P = 0.026$) and inland *C. gulosus* ($r = 0.469$, $P < 0.001$) clusters. *Cottus pitensis* was found to exhibit significant IBD when samples from Bear Creek, a Fall River tributary, were removed ($r = 0.76$, $P < 0.001$). A corrected F_{ST} ($F_{ST}/1-F_{ST}$) gave similar results in all tests.

Discussion

California's complex geographic history has shaped multiple within- and between-species divergences for *C. pitensis* and *C. gulosus*, leading to recurring and novel phylogeographic breaks, cryptic lineages and population-level differences. We found continued genetic support for a novel lineage of *C. gulosus*
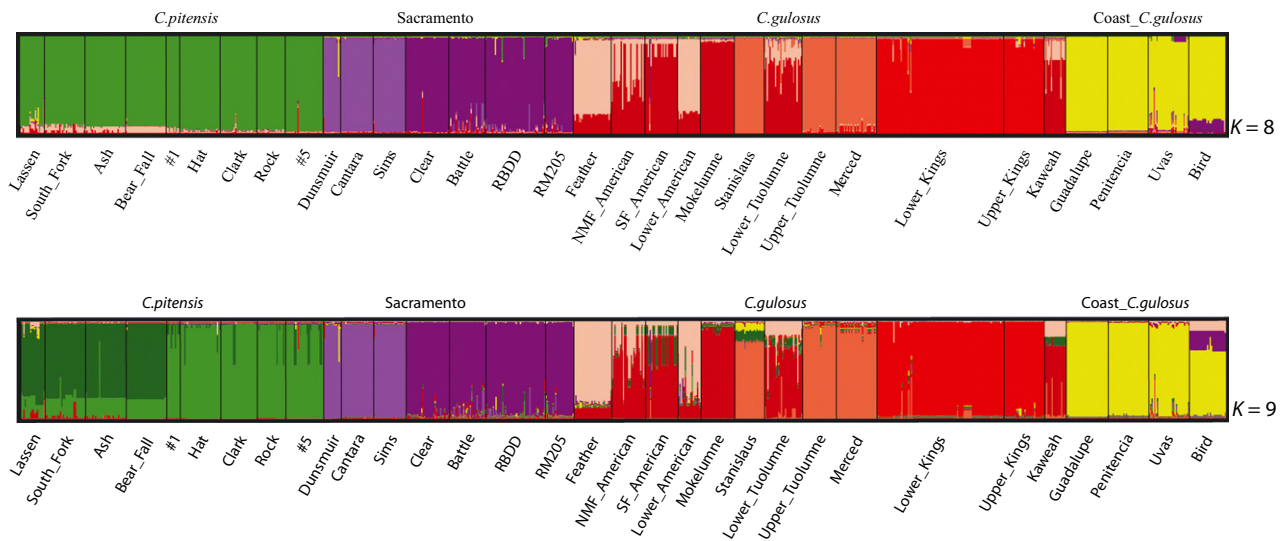
Fig. 5 Clustering analysis based on microsatellite allele frequencies, as determined in STRUCTURE. Primary colours follow distinctions from nDNA and mtDNA sequences, whereas clusters (K) are indicated by shade. Individuals appear as slight vertical bars, whereas locations are delimited by thick black lines.

restricted to a few streams along the central California coast (Baumsteiger *et al.* 2012). Isolation of this group occurred prior to the divergence between *C. pitensis* and inland *C. gulosus* and is consistent with phylogeographic patterns seen in other species. Second, we identified incongruence between species ranges defined by morphological characters and molecular genetic analyses. Our analyses suggest that *C. pitensis* occurs throughout the Sacramento River and that *C. gulosus* occurs in the San Joaquin River. This Sacramento/San Joaquin break has been observed in previous studies of terrestrial and aquatic species. Third, the putative hybridization between species appears historical in nature and unexpectedly extended throughout the Sacramento River. Population analyses confirmed the Sacramento River to be a unique population as well, introducing the idea of hybrid speciation in this region. Lastly, population structure was strongly associated with the presence or absence of Sierra Nevada Mountain geology, and possibly representative of ecological differences present between regions. Overall, results demonstrate the effectiveness of a comparative phylogeographic framework for identifying important geographic impacts on species diversification and population structure as well as future conservation decisions related to sculpin and other endemic species in this highly variable geographic region.

Comparative phylogeography

The currently recognized distributions of these sculpin species proved to be an inaccurate measure of species ranges. Instead, species limits appear to be correlated with previously identified phylogeographic breaks associated with the Coast Range Mountains and Sacramento/San Joaquin River basins. Studies of multiple amphibian and reptile species (Rissler *et al.* 2006), a bird (Sgariglia & Burns 2003) and endemic California plants (Calsbeek *et al.* 2003) have all shown the Coast Range Mountains to be an important phylogeographic break for terrestrial species. But this study is one of the first to ascribe this break to an endemic freshwater fish. Similarly, multiple species of plants (Hickman 1993; Calsbeek *et al.* 2003), amphibians (Fisher & Shaffer 2002), reptiles (Feldman & Spicer 2006), mammals (Maldonado *et al.* 2001; Matocq *et al.* 2012) and a fish (Aguilar & Jones 2009) share the confluence of the Sacramento and San Joaquin Rivers as a phylogeographic break. Although differing ecological or climatic conditions brought on by this break are suspected to be the actual cause of species diversification (see Lapointe & Rissler 2005), our findings continue to bolster support for these breaks in the evolution of species in California.

Novel lineage of coastal C. gulosus

The nDNA species tree, mtDNA haplotype network and microsatellite population analyses all resolved a unique lineage of *Cottus* restricted to a few streams along the central coast of California, consistent with the findings of Baumsteiger *et al.* (2012). The divergence estimate of this group from our mtDNA coalescence tree (3.2–10 Ma) coincides with uplifting in the Coast Range Mountains during the late Miocene/early Pliocene (4–10 Ma), separating habitats in the Great Central Valley from the coast (Howard 1979). Despite strong nDNA

support both here and in Baumsteiger *et al.* (2012), we did not find reciprocal monophyly for this lineage using mtDNA. Instead, potential mitochondrial introgression may have occurred between this lineage and the sympatric species *C. asper*, as originally reported in Baumsteiger *et al.* (2012). An ongoing study seeks to explore this problem further by ascertaining whether morphological differences exist between *C. asper*, inland *C. gulosus* and this novel coastal *C. gulosus* lineage, but currently, it is unknown what role *C. asper* has in the lineage evolution of coastal *C. gulosus*.

Divergence of inland *C. gulosus* and *C. pitensis*

Unlike currently defined species distributions, which confine *C. pitensis* to the Pit River basin and *C. gulosus* to the Sacramento and San Joaquin River basins (Moyle 2002; Page & Burr 2011), we believe that the original divergence occurred somewhere around the confluence of the Sacramento and San Joaquin Rivers. This break led to isolation of *C. pitensis* in the Sacramento basin (including the Pit River) and *C. gulosus* in the San Joaquin basin. Higher mtDNA variation in multiple Sacramento River locations, no shared nuclear haplotypes with inland *C. gulosus* and no evidence of gene flow from Pit to Sacramento Rivers all support an ancestral *C. pitensis* population within the Sacramento River. Divergence at this break does not appear to be linked to a physical barrier to movement, either historically or currently, as contact between groups is possible through fluvial connections (Fig. 1). Instead, dissimilar geographic landscapes between basins probably accounted for ecological or climatic differences that led to diversification (Minckley *et al.* 1986; Durrell 1987; Wagner *et al.* 1997; Lapointe & Rissler 2005; Beesley 2007). The Sacramento River basin underwent extensive volcanic and tectonic activity during much of the Plio-Pleistocene (Pease 1965), consistent with the predicted divergence time for *C. gulusos* and *C. pitensis* (approximately 1.4–6 Ma). These changes, combined with lower elevational peaks in the northern Sierra Nevada vs. much higher peaks in the southern portion of the range, resulted in distinctive climatic conditions for this region during the late Pleistocene (675–500 ka) (Mulch *et al.* 2008). Ecological and climatic changes brought on by the geographic differences have also been attributed to diversification in mountain kingsnakes (Myers *et al.* 2013), woodrats (Matocq *et al.* 2012) and jumping mice (Moritz *et al.* 2008; Malaney *et al.* 2013).

The San Joaquin River basin was historically distinctive from the Sacramento River basin, having been influenced by an inland sea over 16 Ma (Howard 1979; Dupré 1990). During the Pliocene (1.8–5.3 Ma), substantial uplifting in the Sierra Nevada and Coast Range Mountains

caused connections to the sea to be lost, allowing the San Joaquin basin to slowly become shallow freshwater Lake Turlock around the mid-Pleistocene (approximately 1.5 Ma) (Frink & Kues 1954; Howard 1979; Dupré *et al.* 1991; Figueroa & Knott 2010). Concurrently, much of the Sierra Nevada was experiencing intermittent glaciation, which caused changes in river flow and direction (Dupré *et al.* 1991; Gillespie *et al.* 2004). Both factors probably created highly variable and constantly changing environmental conditions for ancestral sculpin in this region, potentially driving diversification.

The presence of Lake Turlock may also have accounted for the low levels of hybridization seen in our nDNA. If ancestral *C. gulosus* were able to migrate north through the lake or were driven out of the San Joaquin basin by its formation (as a strictly riverine species), invasion of the Sacramento River may have occurred. IMA2 results support this conclusion, showing evidence of nDNA gene flow from inland *C. gulosus* into *C. pitensis* when the Sacramento/San Joaquin River phylogeographic break is used. Thus, events leading to the formation of Lake Turlock, along with changes in the Sacramento River basin, could easily have isolated ancestral *C. gulosus*/*C. pitensis* into their prospective regions, leading to speciation.

Hybridization in the Sacramento River basin

Our analyses of the nuclear sequence data suggest that hybridization occurred between *C. pitensis* and inland *C. gulosus* in the Sacramento River, as indicated by the polyphyletic relationships of Sacramento region individuals in the nuclear species tree, shared species nuclear haplotypes and an intermediate geographic location between either taxon. Additionally, IMA2 analysis indicated low levels of migration from inland *C. gulosus* into *C. pitensis* in the Sacramento River. In contrast, analysis of mitochondrial and microsatellite data did not provide evidence for contemporary hybridization between these taxa. Combined, the discordance among molecular marker types implies that hybridization was historical between *C. pitensis* and inland *C. gulosus* in the Sacramento River. A review of the morphologically defined species ranges indicates uncertainties in the distribution of each species in the Sacramento River, which also lends support for the hybridization hypothesis (Bailey & Bond 1963; Moyle 2002).

An alternative hypothesis for polyphyly of nuclear sequence data is incomplete lineage sorting (ILS), where newly formed species have yet to evolve distinctiveness in certain genes (Carstens & Knowles 2007 and citations within). It is reasonable that lineage sorting has occurred in more quickly evolving genes such as mitochondrial DNA and microsatellites but has yet to occur

in more slowly evolving nuclear genes. Another possibility is a combination of hybridization and ILS gave rise to results seen here, as a larger population size within the Sacramento River may delay lineage sorting compared to other regions. Ultimately, future analyses using a more comprehensive genomic approach will be necessary to confirm the cause of discordance across molecular marker types.

Unique contemporary genetic structuring of locations within the Sacramento River basin, as confirmed by microsatellite data, presents an unusual scenario: Are individuals in this region representative of a hybrid taxon? Widely accepted among plant taxonomists but more heavily debated among animal taxa, speciation by way of hybridization has been found in nature (Dowling & Secor 1997; Seehausen 2004). The premise revolves around hybrids being better adapted for a novel environment, one currently unoccupied by either parental species. Freshwater sculpin have two suspected cases of this occurring. The first was in Lake Baikal, where Kontula *et al.* (2003) and Hunt *et al.* (1997) potentially found two sculpin species hybridized prior to invading a new deep-water pelagic habitat, later undergoing extensive adaptive radiations leading to speciation. The second was by Nolte *et al.* (2005) who found that two phylogenetic lineages of *Cottus gobio* were able to undergo hybridization and invade and colonize a novel warm downstream section of the Rhine River in Germany (also see Stemshorn *et al.* 2011). In both cases, hybrids were fitter for a particular habitat and establishing population- or species-level structure over a relatively short evolutionary period (1–4 Ma). However, our genetic data do not support speciation at present in the Sacramento River, merely distinct population structure.

Population-level differentiation

The phylogeographic break and associated ecological and climatic factors observed between the Sacramento/San Joaquin River basins are apparent in population structure analyses of microsatellites, with one exception. The Feather and American Rivers, both tributaries to the lower Sacramento River, were found to contain inland *C. gulosus*, not *C. pitensis*. These rivers occur in the northernmost geologic extent of the Sierra Nevada. Thus, both Feather and American River basins are geologically more similar to the San Joaquin basin where inland *C. gulosus* occur than volcanic regions where *C. pitensis* occurs. If invading inland *C. gulosus* were better adapted for these unique ecological conditions (e.g. strongly seasonal flows, Moyle 2002), natural selection may have driven populations towards inland *C. gulosus* genotypes, including mtDNA, creating

modern-day distributions. However, additional studies are needed to confirm whether proposed geologic differences equate to ecological or climatic discrepancies in habitat between taxa.

Sacramento River and Pit River basins. *Cottus pitensis* populations in the Pit River exhibited little among population structuring as indicated by the presence of a single common haplotype found throughout much of the range, few significant tests for genetic differentiation among populations and lack of IBD signal. An exception to this pattern was Bear Creek, which is isolated from the Pit River proper by lentic habitats and exhibited substantial differentiation. When removed, a strong signal of IBD is resolved, supporting a migration–drift equilibrium. The Pit River appears to contain a unique population of *C. pitensis* in relation to the Sacramento River, with evidence for a break at the mouth of the Pit River. This population-level break is consistent with current morphological species limits (Moyle 2002; Page & Burr 2011) and could explain contemporary species range predictions.

One additional population-level phylogeographic break was discovered to coincide with Pit River Falls. Analysis of levels of clustering ($K = 9$) in STRUCTURE and the unrooted microsatellite trees show that systems upstream of Pit River Falls (Bear, Ash, South Fork and Lassen Creek) cluster separately from all remaining locations downstream. Pit River Falls lies just downstream of the mouth of Fall River, a formidable barrier to upstream sculpin migration and species such as tule perch (*Hysterocarpus traski*) (Moyle 2002).

Although physically connected to populations of inland *C. gulosus* and *C. pitensis*, locations within the Sacramento River basin were resolved as distinctive, showing no contemporary gene flow with populations of either species. Analyses of microsatellites support this finding, whereas nDNA and mtDNA do not, suggesting that isolation may be recent (Holocene). And although we see statistical evidence for IBD, model fit (*r*) values are less than those for the Pit River, implying that the Sacramento River locations exhibit greater variation, a result also evident by the higher number of significant tests for pairwise genetic differentiation. It is not surprising that isolation would lead to a distinct population in the Sacramento River, given its geologic and ecological differences from the Pit River or southern Sierra Nevada rivers (Howard 1979; Durrell 1987; Dupré *et al.* 1991; Knowles & Cayan 2002). *Cottus pitensis* is generally confined to cold-water tributary systems and absent from the mainstem habitats of the basin, especially in the lower reaches (Moyle 2002). Collectively, these differences could drive the contemporary population structure seen in the Sacramento River.

San Joaquin River basin. The San Joaquin River basin is a complex hydrologic region due to its dynamic geology. Since the Pleistocene, factors have influenced individual inland *C. gulosus*, creating an extensive mtDNA haplotype network and fine-scale population structuring resolved by our microsatellite trees, STRUCTURE and AMOVA. Quantifying which factors led to contemporary population structure is currently unavailable, but probably these effects were combinatorial. In addition to a slow, southerly draining freshwater lake and rapid uplifting in the southern Sierra Nevada (Dupré 1990), much of the region was undergoing intermittent glaciation (mid- to late Pleistocene) (Dupré *et al.* 1991; Gillespie *et al.* 2004). Advance and retreat of glaciers throughout this time period (up to Tioga/Hilgard approximately 50–10 ka – Gillespie *et al.* 2004) may have caused continual separation and reconnection of populations of ancestral inland *C. gulosus*. These factors could explain the relational clustering of different sampling locations, as seen in the microsatellite trees and STRUCTURE analyses. Newly available habitat left by retreating glacial episodes or invasion by generalist *C. asper* from coastal populations may have driven ancestral inland *C. gulosus* upstream, accounting for current population distributions.

Conservation implications

Contemporary populations of *C. gulosus* face threats from climate change impacts on Sierra Nevada snowpack, the primary source of water for these systems (Lettenmaier & Gan 1990; Knowles & Cayan 2002; Miller *et al.* 2007; Maurer 2007), and anthropogenic factors such as water diversions, dams and herbicides/pesticides from nearby farms (Fisher & Shaffer 2002). All major rivers within the San Joaquin River basin are dammed somewhere in the Sierra Nevada foothills (Moyle 2002). These factors have all but isolated inland *C. gulosus* populations to their natal streams, similar to what Brown *et al.* (1992) found in California roach (*Lavinia symmetricus*). Thus, the restrictions of each river along with mounting climatic changes make these systems immediate conservation priorities that must be addressed if endemic aquatic species are to be protected both now and in the future.

Climate change and anthropogenic modifications such as dams and water diversions are not unique to California, but global problems for freshwater fish (Ficke *et al.* 2007). Recent large-scale changes in China (Fu *et al.* 2003), Japan (Han *et al.* 2008), the Mediterranean basin (Smith & Darwall 2006) and Australia (Morrongiello *et al.* 2011) have already begun to modify pre-existing species distributions and population structure. Few of these changes have as many climatic and anthropogenic factors as found in our study region (Lund 2010). For example, the confluence of the two river systems in this study is a major source of water for agriculture throughout California and represents up to 40% of the water resources for highly populated southern California (Norgaard *et al.* 2009). To what effect continued changes will have on contemporary and future species distributions or population structure, especially freshwater species, is unknown but could serve as an important model for the conservation of other systems.

The complex phylogeographic structure observed in this study is somewhat unusual for freshwater fishes. General patterns suggest that basin boundaries will be distinct barriers to dispersal and that within-basin gene flow will be relatively low (Hughes 2007; Unmack *et al.* 2013). Yet in this study, we find evidence for historical dispersal between Sacramento and San Joaquin River basins and different patterns of population structure within each basin. Additionally, most comparative phylogeographic studies of freshwater fish, including freshwater sculpin (Hänfling *et al.* 2002), in North America and Europe are associated with different glacial refugia, sea level changes and post-Pleistocene habitat recolonization (Dodson *et al.* 1995; Bernatchez & Wilson 1998; Waters *et al.* 2000; Reyjol *et al.* 2007). Less common but no less important are regional phylogeographic breaks, such as those identified in this study. Diversification in regional systems is often ongoing and highly susceptible to anthropogenic modification. Thus, the identification and understanding of regional phylogeographic breaks is essential to conservation in these regions and makes this study relevant on a much larger scale.

References

Aguilar A, Jones WJ (2009) Nuclear and mitochondrial diversification in two native California minnows: insights into taxonomic identity and regional phylogeography. *Molecular Phylogenetics and Evolution*, **51**, 373–381.

Arbogast BS, Kenagy GJ (2001) Comparative phylogeography as an integrative approach to historical biogeography. *Journal of Biogeography*, **28**, 819–825.

Avise JC (1989) A role for molecular genetics in the recognition and conservation of endangered species. *Trends in Ecology and Evolution*, **4**, 279–281.

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

Avise JC, Arnold J, Ball RM *et al.* (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics*, **18**, 489–522.

Bailey RM, Bond CE (1963) Four new species of freshwater sculpins, genus *Cottus*, from western North America. *Occasional Papers of the Museum of Zoology University of Michigan*, **634**, 1–27.

Baumsteiger J, Aguilar A (2013) Nine original microsatellite loci in prickly sculpin (*Cottus asper*) and their applicability to other closely related *Cottus* species. *Conservation Genetics Resources*, **5**, 279–282.

Baumsteiger J, Kinziger AP, Aguilar A (2012) Life history and biogeographic diversification of an endemic Western North American freshwater fish clade using a comparative species tree approach. *Molecular Phylogenetics and Evolution*, **65**, 940–952.

Baumsteiger J, Kinziger AP, Reid SB, Aguilar A (2014) Data from: complex phylogeography and historical hybridization between sister taxa of freshwater sculpin (*Cottus*). Dryad Digital Repository. http://dx. doi:10.5061/dryad.1077n.

Beesley D (2007) *Crow's Range: An Environmental History of the Sierra Nevada*. University of Nevada Press, Reno, NV.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (1996-2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France).

Bermingham E, McCafferty SS, Martin AP (1997) Fish biogeography and molecular clocks: perspectives from the Panamanian Isthmus. In: *Molecular Systematics of Fishes* (eds Kocher TD, Stepien CA), pp. 113–128. Academic Press, San Diego, California.

Bernatchez L, Wilson CC (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology*, **7**, 431–452.

Bickford D, Lohman DJ, Sodhi NS *et al.* (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution*, **22**, 148–155.

Brown LR, Moyle PB, Bennett WA, Quelvog BD (1992) Implications of morphological variation among populations of California roach *Lavinia symmetricus* (Cyprinidae) for conservation policy. *Biological Conservation*, **62**, 1–10.

Bryson RW Jr, De Oca ANM, Jaeger JR, Riddle BR (2010) Elucidation of cryptic diversity in a widespread Nearctic treefrog reveals episodes of mitochondrial gene capture as frogs diversified across a dynamic landscape. *Evolution*, **64**, 2315–2330.

Burns KJ, Barhoum DN (2006) Population-level history of the wrentit (*Chamaea fasciata*): implications for comparative phylogeography in the California Floristic Province. *Molecular Phylogenetics and Evolution*, **38**, 117–129.

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.

Carstens BC, Knowles LL (2007) Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from Melanoplus grasshoppers. *Systematic Biology*, **56**, 400–411.

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

Dayrat B (2005) Towards integrative taxonomy. *Biological Journal of the Linnean Society*, **85**, 407–415.

Dodson JJ, Colombani F, Ng PKL (1995) Phylogeographic structure in mitochondrial DNA of a South-east Asian freshwater fish, Hemibagrus nemurus (Siluroidei; Bagridae) and Pleistocene sea-level changes on the Sunda shelf. *Molecular Ecology*, **4**, 331–346.

Dowling TE, Secor CL (1997) The role of hybridization and introgression in the diversification of animals. *Annual Review of Ecology and Systematics*, **28**, 593–619.

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

Dupré WR (1990) Quaternary geology of the Monterey Bay region, California. In: *Geology and Tectonics of the Central California Coastal Region, San Francisco to Monterey*. Vol. **67** (eds Garrison RE, Greene HG, Hicks KR, Weber GE, Wright TL), pp. 185–191. Pacific Section of the American Association of Petroleum Geologists, Tulsa, OK.

Dupré WR, Morrison RB, Clifton HE *et al.* (1991) Quaternary geology of the Pacific margin. In: *Quaternary Nonglacial Geology; Conterminous US* (ed. Morrison RB), pp. 141–214. The Geological Society of America, Boulder, Colorado.

Durrell C (1987) *Geologic History of the Feather River County, California*. UC Press, Berkeley, California.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**, 1792–1797.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, **1**, 47–50.

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

Ficke AD, Myrick CA, Hansen LJ (2007) Potential impacts of global climate change on freshwater fisheries. *Reviews in Fish Biology and Fisheries*, **17**, 581–613.

Figueroa AM, Knott JR (2010) Tectonic geomorphology of the southern Sierra Nevada Mountains (California): evidence for uplift and basin formation. *Geomorphology*, **123**, 34–45.

Fisher RN, Shaffer HB (2002) The decline of amphibians in California's Great Central Valley. *Conservation Biology*, **10**, 1387–1397.

Frink JW, Kues HA (1954) Corcoran clay-A Pleistocene Lacustrine deposit in San Joaquin Valley, California. *Bulletin of the American Association of Petroleum Geologists*, **38**, 2357–2371.

Fu C, Wu J, Chen J, Wu Q, Lei G (2003) Freshwater fish biodiversity in the Yangtze River basin of China: patterns, threats and conservation. *Biodiversity & Conservation*, **12**, 1649–1685.

Gillespie AR, Porter SC, Atwater BF (2004) *The Quaternary Period in the United States*, vol 1. Elsevier Science, Amsterdam, The Netherlands.

Grinnell J (1924) Geography and evolution. *Ecology*, **5**, 225–229.

Han M, Fukushima M, Kameyama S, Fukushima T, Matsushita B (2008) How do dams affect freshwater fish distributions in Japan? Statistical analysis of native and nonnative species with various life histories. *Ecological Research*, **23**, 735–743.

Hänfling B, Hellemans B, Volckaert FAM, Carvalho GR (2002) Late glacial history of the cold-adapted freshwater fish *Cottus gobio*, revealed by microsatellites. *Molecular Ecology*, **11**, 1717–1729.

Harrison RG (1993) *Hybrid Zones and the Evolutionary Process*. Oxford University Press, New York City, New York.

Hasegawa M, Kishino H, Yano T (1985) Dating of human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Hershler R (1995) New freshwater snails of the genus *Pyrgulopsis* (Rissooidea: Hydrobiidae) from California. *Veliger*, **38**, 343–373.

Hey J, Nielsen R (2007) Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 2785–2790.

Hickerson MJ, Carstens BC, Cavender-Bares J *et al.* (2010) Phylogeography's past, present, and future: 10 years after. Avise, 2000. *Molecular Phylogenetics and Evolution*, **54**, 291–301.

Hickman JC (ed.) (1993) *The Jepson Manual: Higher Plants of California*. UC Press, Berkeley, California.

Howard AD (1979) *Geologic History of Middle California*, vol. 43. UC Press, Berkeley, California.

Hughes JM (2007) Constraints on recovery: using molecular methods to study connectivity of aquatic biota in rivers and streams. *Freshwater Biology*, **52**, 616–631.

Hughes JM, Schmidt DJ, Finn DS (2009) Genes in streams: using DNA to understand the movement of freshwater fauna and their riverine habitat. *BioScience*, **59**, 573–583.

Hunt DM, Fitzgibbon J, Slobodyanyuk SJ, Bowmaker JK, Dulai KS (1997) Molecular evolution of the cottoid fish endemic to Lake Baikal deduced from nuclear DNA evidence. *Molecular Phylogenetics and Evolution*, **8**, 415–422.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Kalinowski ST (2005) HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. *Molecular Ecology Notes*, **5**, 187–189.

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111–120.

King JL, Simovich MA, Brusca RC (1996) Species richness, endemism and ecology of crustacean assemblages in northern California vernal pools. *Hydrobiologia*, **328**, 85–116.

Kinziger AP, Raesly RL (2001) A narrow hybrid zone between two *Cottus* species in Wills Creek, Potomac Drainage. *Heredity*, **92**, 309–314.

Kinziger AP, Wood RM, Neely DA (2005) Molecular systematics of the genus *Cottus* (Scorpaeniformes: Cottidae). *Copeia*, **2**, 303–311.

Knowles N, Cayan DR (2002) Potential effects of global warming on the Sacramento/San Joaquin watershed and the San Francisco estuary. *Geophysical Research Letters*, **29**, 1891.

Kontula T, Kirilchik SV, Väinölä R (2003) Endemic diversification of the monophyletic cottoid fish species flock in Lake Baikal explored with mtDNA sequencing. *Molecular Phylogenetics and Evolution*, **27**, 143–155.

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

Lettenmaier DP, Gan TY (1990) Hydrologic sensitivities of the Sacramento-San Joaquin River basin, California, to global warming. *Water Resources Research*, **26**, 69–86.

Leu M, Hanser SE, Knick ST (2008) The human footprint in the west: a large-scale analysis of anthropogenic impacts. *Ecological Applications*, **18**, 1119–1139.

Lewis PO, Zaykin D (2001) Genetic data analysis: computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the internet from http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php

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

Light T, Erman DC, Myrick C, Clarke J (2002) Decline of the Shasta crayfish (*Pacifastacus fortis* Faxon) of northeastern California. *Conservation Biology*, **9**, 1567–1577.

Lund JR (2010) *Comparing Futures for the Sacramento-San Joaquin Delta*. UC Press, Berkeley, California.

MacDonald GA, Gay TE (1968) Geology of the Cascade Range [repr.]. California Division of. Mines and Geology, Mineral. *Information Service*, **21**, 108–112.

Malaney JL, Conroy CJ, Moffitt LA, Spoonhunter HD, Patton JL, Cook JA (2013) Phylogeography of the western jumping mouse (*Zapus princeps*) detects deep and persistent allopatry with expansion. *Journal of Mammalogy*, **94**, 1016–1029.

Maldonado JE, Vila C, Wayne RK (2001) Tripartite genetic subdivisions in the ornate shrew (*Sorex ornatus*). *Molecular Ecology*, **10**, 127–147.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology and Evolution*, **18**, 189–197.

Matocq MD, Kelly PA, Phillips SE, Maldonado JE (2012) Reconstructing the evolutionary history of an endangered subspecies across the changing landscape of the Great Central Valley of California. *Molecular Ecology*, **21**, 5918–5933.

Maurer EP (2007) Uncertainty in hydrologic impacts of climate change in the Sierra Nevada, California, under two emissions scenarios. *Climatic Change*, **82**, 309–325.

Miller NL, Bashford KE, Strem E (2007) Potential impacts of climate change on California hydrology. *Journal of the American Water Resources Association*, **39**, 771–784.

Minckley WL, Hendrickson DA, Bond CE (1986) Geography of western North American freshwater fishes: description and relationship to intracontinental tectonism. In: *The Zoogeography of North American Freshwater Fishes*(eds Hocutt C. H. & Wiley E. O.), pp. 519–614. John Wiley & Sons, New York.

Moritz C, Patton JL, Conroy CJ, Parra JL, White GC, Beissinger SR (2008) Impact of a century of climate change on small-mammal communities in Yosemite National Park, USA. *Science*, **322**, 261–264.

Morrongiello JR, Beatty SJ, Bennett JC et al. (2011) Climate change and its implications for Australia's freshwater fish. *Marine and Freshwater Research*, **62**, 1082–1098.

Mount JF (1995) *California Rivers and Streams: The Conflict between Fluvial Process and Land Use*. UC Press, Berkeley, California.

Moyle PB (2002) *Inland Fishes of California*. UC Press, Berkeley, California.

Moyle PB, Williams JE (2005) Biodiversity loss in the temperate zone: decline of the native fish fauna of California. *Conservation Biology*, **4**, 275–284.

Mueller RL (2006) Evolutionary rates, divergence dates, and the performance of mitochondrial genes in Bayesian phylogenetic analysis. *Systematic Biology*, **55**, 289–300.

Mulch A, Sarna-Wojcicki AM, Perkins ME, Chamberlain CP (2008) A Miocene to Pleistocene climate and elevation record of the Sierra Nevada (California). *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 6819–6824.

Myers EA, Rodríguez-Robles JA, DeNardo DF et al. (2013) Multilocus phylogeographic assessment of the California Mountain Kingsnake (*Lampropeltis zonata*) suggests alternative patterns of diversification for the California Floristic Province. *Molecular Ecology*, **22**, 5418–5429.

Nolte AW, Freyhof J, Stemshorn KC, Tautz D (2005) An invasive lineage of sculpins, *Cottus* sp. (Pisces, Teleostei) in the Rhine with new habitat adaptations has originated from hybridization between old phylogeographic groups. *Proceedings of the Royal Society B*, **272**, 2379–2387.

Norgaard RB, Kallis G, Kiparsky M (2009) Collectively engaging complex socio-ecological systems: re-envisioning science, governance, and the California Delta. *Environmental Science and Policy*, **12**, 644–652.

Page LM, Burr BM (2011) *A Field Guide to Freshwater Fishes: North America North of Mexico*. Houghton Mifflin Harcourt, New York City, New York.

Pease RM (1965) *Modoc County: A Geographic Time Continuum on the California Volcanic Tableland*. UC Press, Berkeley, California.

Peterson AT, Soberón J, Sánchez-Cordero V (1999) Conservatism of ecological niches in evolutionary time. *Science*, **285**, 1265–1267.

Poissant J, Knight TW, Ferguson MM (2005) Nonequilibrium conditions following landscape rearrangement: the relative contribution of past and current hydrological landscapes on the genetic structure of a stream-dwelling fish. *Molecular Ecology*, **14**, 1321–1331.

Pond SLK, Posada D, Gravenor MB, Woelk CH, Frost SD (2006) Automated phylogenetic detection of recombination using a genetic algorithm. *Molecular Biology and Evolution*, **23**, 1891–1901.

Posada D (2008) jModelTest: phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology*, **53**, 793–808.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Reyjol Y, Hugueny B, Pont D et al. (2007) Patterns in species richness and endemism of European freshwater fish. *Global Ecology and Biogeography*, **16**, 65–75.

Rice WR (1989) Analyzing tables of statistical test. *Evolution*, **43**, 223–225.

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. *The American Naturalist*, **167**, 655–666.

Rosenberg NA (2003) DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, **4**, 137–138.

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

Seehausen O (2004) Hybridization and adaptive radiation. *Trends in Ecology and Evolution*, **19**, 198–207.

Sgariglia EA, Burns KJ (2003) Phylogeography of the California thrasher (*Toxostoma redivivum*) based on nested-clade analysis of mitochondrial-DNA variation. *The Auk*, **120**, 346–361.

Shaffer HB, Pauly GB, Oliver JC, Trenham PC (2004) The molecular phylogenetics of endangerment: cryptic variation and historical phylogeography of the California tiger salamander, *Ambystoma californiense*. *Molecular Ecology*, **13**, 3033–3049.

da Silva MN, Patton JL (1998) Molecular phylogeography and the evolution and conservation of Amazonian mammals. *Molecular Ecology*, **7**, 475–486.

Smith KG, Darwall WR (2006) *The Status and Distribution of Freshwater Fish Endemic to the Mediterranean Basin*, vol. **1**. IUCN, Switzerland.

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of un-glaciated Eastern North America. *Molecular Ecology*, **15**, 4261–4293.

Spinks PQ, Thomson RC, Shaffer BH (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 (2006) Multilocus genealogies reveal multiple cryptic species and biogeographical complexity in the California turret spider *Antrodiaetus riversi* (Mygalomorphae, Antrodiaetidae). *Molecular Ecology*, **16**, 583–604.

Stemshorn KC, Reed FA, Nolte AW, Tautz D (2011) Rapid formation of distinct hybrid lineages after secondary contact of two fish species (*Cottus* sp.). *Molecular Ecology*, **20**, 1475–1491.

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

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

Storfer A, Murphy MA, Spear SF, Holderegger R, Waits LP (2010) Landscape genetics: where are we now? *Molecular Ecology*, **19**, 3496–3514.

Swenson NG, Howard DJ (2005) Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. *The American Naturalist*, **166**, 581–591.

Swofford DL (2002) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*, vol. 4. Sinauer Associates, Sunderland, Massachusetts.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, **7**, 453.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.

Thompson JN (1999) The evolution of species interactions. *Science*, **284**, 2116–2118.

Turner MG (1989) Landscape ecology: the effect of pattern on process. *Annual Review of Ecology and Systematics*, **20**, 171–197.

Unmack PJ, Hammer MP, Adams M, Johnson JB, Dowling TE (2013) The role of continental shelf width in determining freshwater phylogeographic patterns in south-eastern Australian pygmy perches (Teleostei: Percichthyidae). *Molecular Ecology*, **22**, 1683–1699.

Wagner HM, Hanson CB, Gustafson EP, Gobalet KW, Whistler DP (1997) Biogeography of Pliocene and Pleistocene vertebrate faunas of northeastern California and their temporal significance to the development of the Modoc Plateau and the Klamath Mountains Oregon. *San Bernardino County Museum Association Quarterly*, **44**, 13–21.

Wake DB (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.

Waters JM, Dijkstra LH, Wallis GP (2000) Biogeography of a southern hemisphere freshwater fish: how important is marine dispersal? *Molecular Ecology*, **9**, 1815–1821.

Wu JJ (2013) *Ecological Systems*. Springer, New York City, New York.

Data accessibility

All sequence data has been uploaded to GenBank (JX484692-JX484693; KJ509201-KJ509583). Microsatellite genotypes, sequence alignments by marker, species tree, and usat tree files are uploaded on the Dryad Digital Repository (http://dx.doi:10.5061/dryad.1077n).

Supporting information

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

Appendix S1 Materials and methods.

Fig. S1 Haplotype networks for each of three nuclear markers.

Fig. S2 Plot of log likelihood values vs. clusters (K) showing the $K = 5$ identified by the Evanno *et al.* (2005) method in STRUCTUREHARVESTER.

Fig. S3 Unrooted neighbor-joining tree showing Reynolds (A), Nei (B), and Cavalli-Sforza Edwards (C) distance trees obtained from six microsatellite loci (locus 318 removed).

Fig. S4 Plots of Isolation-by-Distance showing genetic distance (F_{ST}) vs. geographic distance (in km) within identified clades.

Table S1 Haplotypes for three nDNA markers, showing the number of individuals from each location with that haplotype.

Table S2 Shown is the number of individuals containing a specific mtDNA haplotype at a location.

Table S3 Genetic distance (F_{ST} or Weir and Cockerham's θ) (above diagonal) and geographic distance (in km) (below diagonal) found between locations; measurements used for pairwise comparison of isolation by distance.

Table S4 Analysis of molecular variance (AMOVA) for two different potential groupings of locations based on STRUCTURE K outputs.