

Riverscape genetics identifies speckled dace (*Rhinichthys osculus*) cryptic diversity in the Klamath–Trinity Basin

Jesse C. Wiesenfeld^{1,2}  · Damon H. Goodman³ · Andrew P. Kinziger¹

Received: 26 December 2016 / Accepted: 5 November 2017
© Springer Science+Business Media B.V., part of Springer Nature 2017

Abstract Cataloging biodiversity is of great importance given that habitat destruction has dramatically increased extinction rates. While the presence of cryptic species poses challenges for biodiversity assessment, molecular analysis has proven useful in uncovering this hidden diversity. Using nuclear microsatellite markers and mitochondrial DNA we investigated the genetic structure of Klamath speckled dace (*Rhinichthys osculus klamathensis*), a subspecies endemic to the Klamath–Trinity basin. Analysis of 25 sample sites within the basin uncovered cryptic diversity including three distinct genetic groups: (1) a group that is widely distributed throughout the Klamath River mainstem and its tributaries, (2) a group distributed in the Trinity River, the largest tributary to the Klamath River, and (3) a group identified above a 10 m waterfall in Jenny Creek, a small tributary to the Klamath River. All groups were resolved as divergent in nuclear microsatellite analysis and exhibited levels of divergence in mitochondrial DNA that were comparable to those observed among recognized *Rhinichthys* species. No physical barriers currently separate the Klamath and Trinity groups and the precise mechanism that generated and maintains the groups as distinct despite contact and hybridization

is unknown. The present study highlights the importance of incorporating molecular analysis into biodiversity research to uncover cryptic diversity. We recommend that future biodiversity inventories recognize three genetically distinct groups of speckled dace in the Klamath–Trinity Basin.

Keywords Riverscape genetics · Cryptic species · Speckled dace · *Rhinichthys osculus*

Introduction

Biodiversity is being lost at an astonishing rate due to forces associated with human population growth, habitat alteration and global climate change (McNeely et al. 1990; Chapin III et al. 2000). The situation is particularly troublesome for freshwater ecosystems, as freshwater biodiversity has declined at a faster rate compared to terrestrial and marine systems (Ricciardi and Rasmussen 1999; Jenkins 2003). Freshwater ecosystems contain 0.01% of the world's water by volume; yet these systems are inhabited by 6% of all described species (Dudgeon et al. 2006). Fishes are the most-studied indicators of biodiversity decline in freshwater ecosystems (Moyle et al. 2011) but our knowledge and understanding of freshwater fish diversity, patterns of endemism, and genetic variation is limited, with considerable amounts of diversity yet to be described (Myers et al. 2000; Abell et al. 2008).

With limited resources for conservation, it is vital that ecologists are able to identify the appropriate level of conservation units for protection. Species have long been considered the foundation of biodiversity (McNeely et al. 1990), and a standard metric to monitor environmental conditions (Noss 1990). However, with the expansion of genetic data in conservation, and the uncertainty in species designation,

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10592-017-1027-6>) contains supplementary material, which is available to authorized users.

✉ Jesse C. Wiesenfeld
jesse.wiesenfeld@fishsciences.net

¹ Department of Fisheries Biology, Humboldt State University, One Harpst Street, Arcata, CA 95521, USA

² Cramer Fish Sciences, 3300 Industrial Blvd #100 West, Sacramento, CA 95691, USA

³ US Fish and Wildlife Service, 1655 Heindon Road, Arcata, CA 95521, USA

conservation managers have begun to incorporate diversity below the taxonomic species level (i.e. subspecies, evolutionary significant units (ESUs), distinct population segments). Delimitation of fauna into definable conservation units is especially difficult for organisms that are morphologically indistinguishable or cryptic (Bernardo 2011). Cryptic diversity is commonly misclassified by traditional approaches, where two or more units are mistakenly classified as one. With the increase in DNA-based studies, cryptic diversity is rapidly being discovered throughout taxa and bioregions (Bickford et al. 2007; Pfenninger and Schwenk 2007). Once cryptic diversity has been identified by molecular analysis, how or if it is delimited into a conservation unit, such as a “species” or an “ESU”, is controversial (Niemiller et al. 2013; Murphy et al. 2015). Though the identification and recognition of cryptic diversity further complicates taxonomy and conservation management, it provides a more realistic picture of biodiversity and extinction risks. An accurate accounting of cryptic diversity is essential for identification of biodiversity hotspots, species distributions, and reserve design (Bickford et al. 2007; Bernardo 2011; Piggott et al. 2011).

Speckled dace (*Rhinichthys osculus*), a small (typically less than 100 mm fork length) freshwater cyprinid that occurs throughout much of the western United States, displays considerable amounts of cryptic diversity that have yet to be categorized (e.g., Hoekzema and Sidlauskas 2014). Speckled dace inhabit a wide assortment of habitats (Moyle 2002) but only subtle morphological differences have been recorded between speckled dace populations exhibiting species level genetic divergence (Hoekzema and Sidlauskas 2014). Presently, speckled dace are considered to be a single, wide ranging species comprised of a number of subspecies. There are four subspecies listed under the Endangered Species Act and four subspecies considered to be of conservation concern in California (Moyle et al. 2015).

Several phylogenetic studies have undertaken the task of untangling the evolutionary history of speckled dace (Oakey et al. 2004; Pfrender et al. 2004; Smith and Dowling 2008; Ardren et al. 2010; Billman et al. 2010; Hoekzema and Sidlauskas 2014). Molecular assessments have shown that speckled dace represent a polyphyletic taxon containing a number of undescribed taxa (Ardren et al. 2010; Hoekzema and Sidlauskas 2014). In general, these studies have shown that deep genetic divergence occurs among speckled dace isolated in separate river basins. Diversification of speckled dace, like many freshwater fishes of the American west, can be attributed to complex geologic and climatic processes that caused extended periods of isolation between basins, interspersed with episodes of dispersal associated with drainage rearrangements (Minckley et al. 1986).

Understanding of speckled dace taxonomy and evolutionary history is complicated by the potential for hybridization.

Genetic studies investigating hybridization in speckled dace are limited, but evidence of hybridization between introduced speckled dace and relict dace (*Relictus solitarius*) in the Great Basin has been reported (Houston et al. 2012). Also, Smith (1973) suggested hybridization between speckled dace and two other species (longnose dace *Rhinichthys cataractae* and reidside shiners *Richardsonius balteatus*) based upon morphological analyses.

Our research investigates speckled dace in the Klamath–Trinity Basin in south central Oregon and northwestern California. Due to genetic distinctiveness and geographic isolation, speckled dace occurring in the Klamath–Trinity Basin are recognized as an endemic subspecies, Klamath speckled dace (*R. osculus klamathensis*) (Oakey et al. 2004; Pfrender et al. 2004). Klamath speckled dace are nearly continuously distributed throughout the entire Klamath–Trinity Basin and occur at high abundance in many areas in the basin (Moyle 2002). While speckled dace are common and native to the Klamath–Trinity Basin, they are absent from all adjacent coastal basins, with the exception of introduced populations (Moyle 2002).

Two studies have resolved evidence for multiple speckled dace lineages within the Klamath–Trinity Basin. First, Pfrender et al. (2004) uncovered evidence of two highly divergent mitochondrial DNA haplotypes in the Upper Klamath River and hypothesized that there might be two reproductively isolated forms of speckled dace co-occurring in the region. Second, Kinziger et al. (2011) found deep divergence in nuclear microsatellites and mitochondrial DNA between speckled dace occurring in the Trinity River system versus those from the Klamath River system. While these studies suggest the presence of multiple distinct genetic groups, the numbers and distributions of the previous sampling locations were insufficient to provide a clear understanding of speckled dace genetic structuring within the Klamath–Trinity Basin.

In this study, we conducted a higher resolution assessment of speckled dace genetic structure in the Klamath–Trinity Basin using both nuclear microsatellites and mitochondrial DNA markers over a broader geographic range of sampling sites. Our objective was to identify cryptic genetic groups, their geographic boundaries, contact zones and levels of hybridization. Overall our analysis resolved highly complex patterns of genetic structuring, including (1) resolution of deep genetic divergence between speckled dace in the Trinity River system and the Klamath River system (hereafter referred to as the “Klamath” and “Trinity” groups; Fig. 1), (2) evidence for hybridization between the Klamath and Trinity groups near the confluence of the Klamath and Trinity rivers, specifically in Tish Tang Creek (TT) and the Salmon River (NFS and SFS, hereafter these three sites are referred to as Klamath X Trinity hybrids), (3) identification of a genetically distinctive group of speckled dace isolated by a waterfall, located in Jenny Creek (JEN), a tributary to

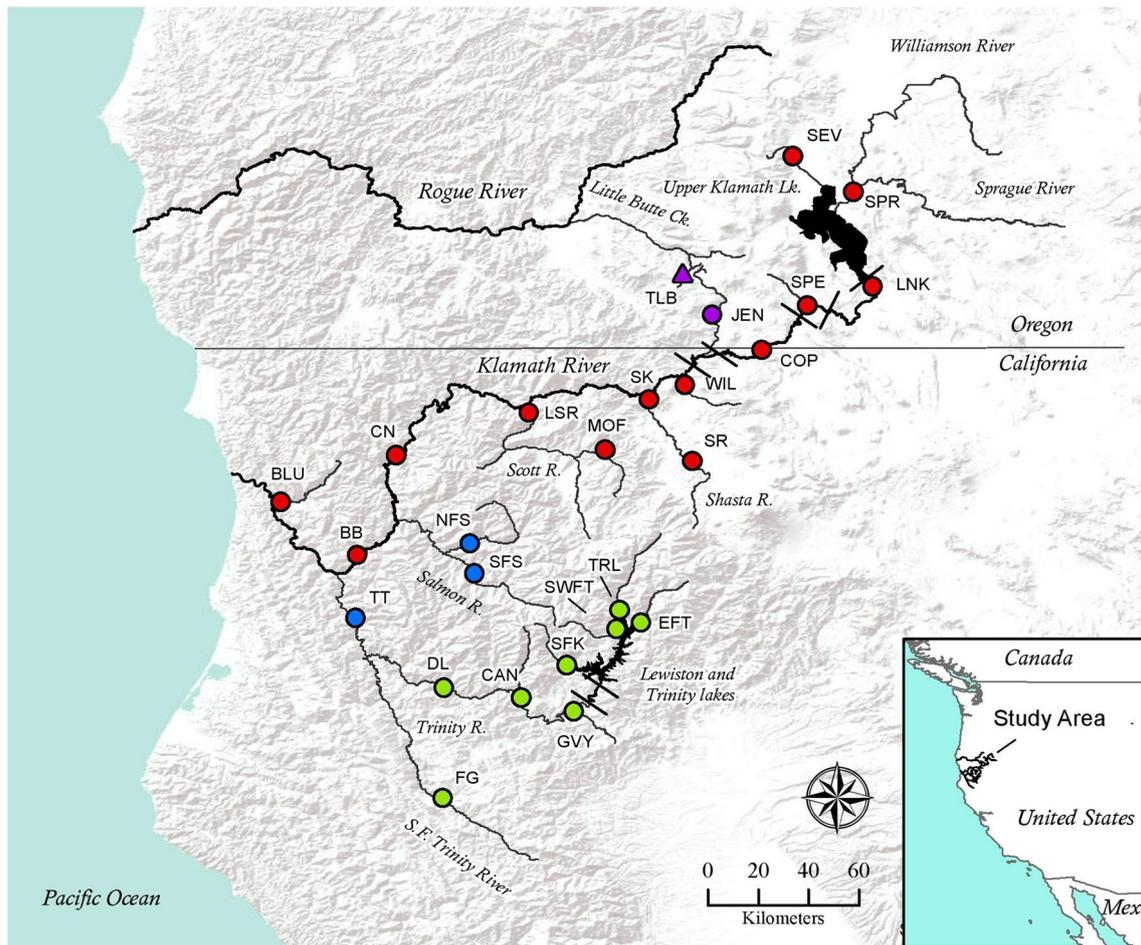


Fig. 1 Location of the 25 Klamath–Trinity Basin speckled dace collections in California and Oregon, USA. Collections from the Klamath River system (red), Trinity River system (green), Jenny Creek (purple), and Klamath X Trinity hybrids (blue). Circles repre-

sent collections for this study or from Kinziger et al. (2011) and the triangle represents the location (tributary to Little Butte Creek, TLB) of sequence data provided by Thomas Dowling. Site abbreviations as in Table 1. (Color figure online)

the Klamath River, and (4) an extremely rare mitochondrial haplotype (SEV01) restricted to the upper reaches of the Klamath River basin, that is highly divergent from all other speckled dace examined herein.

Materials and methods

Field collections

Speckled dace were collected at 25 sites throughout the Klamath–Trinity Basin (Fig. 1; Table 1). Collections of the Klamath group (N=13) spanned from the mouth of the Klamath River to above Klamath Lake. Collections of the Trinity group (N=9) included sample sites throughout the Trinity River system, including those above Trinity Lake. Collections of the Klamath X Trinity hybrids consisted of three sites near the confluence of the Klamath and Trinity

ivers, including Tish Tang Creek (TT) and the Salmon River (NFS and SFS). Jenny Creek was represented by a single collection. Specimens were collected using seine nets, or a backpack electrofisher and were supplemented by specimens from Kinziger et al. (2011) which were archived at the Humboldt State University (HSU) Fish Collection. Specimens were euthanized using an overdose of tricaine methanesulfonate and vouchered whole, or were anesthetized and then a caudal fin clip was collected before releasing the fish. Whole specimens/tissue were preserved in 95% ethanol and deposited into the HSU Fish Collection.

Microsatellite genotyping methods

Speckled dace were genotyped at nine microsatellite loci (Baerwald and May 2004; Turner et al. 2004; Girard and Angers 2006, Table S1). Whole genomic DNA was extracted from fin tissue using the Chelex DNA extraction method

Table 1 River system, collection site, site abbreviation (ID), latitude, longitude, collection date, and the Humboldt State University Fish Collection numbers (HSU ID) from Klamath–Trinity basin speckled dace

| River system | Collection site | ID | Latitude | Longitude | Date | HSU ID |
|--------------|------------------------------------|------------------|----------|------------|------------|-----------|
| Klamath | | | | | | |
| | Seven Mile Creek | SEV | 42.69421 | – 122.0727 | 9/2/2012 | 5079/5080 |
| | Sprague River | SPR | 42.56339 | – 121.8624 | 7/25/2012 | 5068 |
| | Link River | LNK | 42.2216 | – 121.7934 | 9/2/2012 | 5077/5078 |
| | Spencer Creek | SPE | 42.15257 | – 122.0278 | 6/23/2012 | 5066 |
| | Jenny Creek | JEN | 42.11791 | – 122.367 | 6/23/2012 | 5065 |
| | Klamath River (above Copco 2) | COP | 41.9915 | – 122.1902 | 7/26/2012 | 5002 |
| | Willow Creek | WIL | 41.86591 | – 122.4648 | 7/25/2012 | 5001 |
| | Shasta River (upper) | SR ^a | 41.591 | – 122.438 | 2/5/2010 | 4943 |
| | Shasta River (lower) | SK | 41.81309 | – 122.5923 | 7/26/2012 | 4999 |
| | Moffet Creek | MOF | 41.6337 | – 122.749 | 7/26/2012 | 5010 |
| | Lower Scott River | LSR | 41.76529 | – 123.0213 | 7/1/2012 | 5131 |
| | Klamath River (coon creek) | CN | 41.61363 | – 123.4958 | 8/31/2012 | 5076 |
| | Klamath River (big bar) | BB | 41.25204 | – 123.6348 | 8/31/2012 | 5073 |
| | Blue Creek | BLU | 41.44417 | – 123.9069 | 5/31/2012 | 4998 |
| | North Fork Salmon | NFS | 41.29301 | – 123.2301 | 7/25/2012 | 5036 |
| | South Fork Salmon | SFS | 41.18689 | – 123.2139 | 7/25/2012 | 5014 |
| Trinity | | | | | | |
| | Tish Tang Creek | TT | 41.02586 | – 123.6403 | 7/24/2012 | 5023/5083 |
| | Trinity River (Del Loma) | DL | 40.7749 | – 123.325 | 7/31/2013 | 5123 |
| | South Fork Trinity River | FG ^a | 40.377 | – 123.3256 | 10/25/2004 | 4940 |
| | Canyon Creek | CAN ^a | 40.738 | – 123.0495 | 8/10/2010 | 4945/4944 |
| | Grass Valley Creek | GVY | 40.68979 | – 122.8576 | 7/24/2012 | 5007 |
| | Stuart Fork Creek | SFK | 40.8558 | – 122.8851 | 7/24/2012 | 5030 |
| | Swift Creek | SWFT | 40.98642 | – 122.7089 | 7/25/2012 | 5012 |
| | Trinity River (above Trinity Lake) | TRL | 41.05328 | – 122.6965 | 7/25/2012 | 5004 |
| | East Fork Trinity | EFT | 41.0085 | – 122.6201 | 7/26/2012 | 5016 |

^aSamples analyzed in Kinziger et al. (2011)

(Walsh et al. 1991) and microsatellite loci were amplified via polymerase chain reaction. Amplifications were performed as either 10 or 12.7- μ l reactions using GoTaq Colorless Master Mix (Promega, Madison, WI) in a MJ Research (Waltham, MA) PTC-100 or an Applied Biosystems (Grand Island, NY) 2720 thermal cycler. The forward primer of each primer pair was labeled with a WellRED fluorescent dye (Sigma-Aldrich, St Louis, MO) for identification. Products were visualized and allele sizes determined with a Beckman-Coulter CEQ 8000 Genetic Analysis System (Brea, CA). Allele sizes were scored twice and any discrepancies were either resolved or the genotype was removed. Individuals missing more than two loci from their multi-locus genotype were removed from the dataset. Tests for conformance to Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium were conducted using GENEPOP V4.2 (Raymond and Rousset 1995). Loci were checked for null alleles, stutter peaks, and large allele dropout using MICRO-CHECKER v2.2.3 (Van Oosterhout et al. 2004).

Microsatellite analysis

Nuclear diversity

Observed heterozygosity (H_o), Hardy–Weinberg expected heterozygosity (H_e), and allelic richness (A) were calculated in ARLQUIN 3.11 (Excoffier et al. 2005). Rarified allelic richness (A_R) and private allelic richness (A_p), both standardized to a sample size of 48 genes, were calculated using HP-Rare (Kalinowski 2005). Permutation tests (2000 replicates) for significant differences in A_R and H_e between the Klamath and Trinity groups were conducted using the software FSTAT v2.9.3 (Goudet 1995).

Population structure

Pairwise estimates of genetic differentiation (F_{ST}) between sites and tests of their significance were conducted using FSTAT v2.9.3 (Goudet 1995). Pairwise estimates of

standardized F_{ST} , (F'_{ST}) which is standardized to the largest possible value obtainable, were calculated in GENODIVE (Meirmans and Van Tienderen 2004). A graphical depiction of genetic divergence between sites was generated by constructing a neighbor-joining tree of Cavalli-Sforza and Edwards chord distances using the software PHYLIP v3.68 (Felsenstein 2005). Branch support was evaluated by a bootstrap analysis with 1000 replicates.

Two different genetic clustering approaches were employed to analyze the microsatellite dataset. The first approach utilized a Bayesian clustering algorithm that is implemented in the software STRUCTURE v 2.3.4 (Pritchard et al. 2000) to estimate the number of discrete genetic clusters (K) of individuals. An individual's assignment to each cluster was also calculated, called the admixture proportion (q), and can be used to estimate hybridization levels. After discarding the first 100,000 steps of the MCMC simulations as burn-in, 100,000 additional steps were performed. A total of 20 iterations were conducted at each level of assumed $K = 1 \dots 12$. To estimate the number of clusters in the data the ad hoc method of ΔK (Evanno et al. 2005) was calculated using STRUCTURE HARVESTER (Earl and vonHoldt 2012), where the largest change in K infers the number of clusters. The results from STRUCTURE were visualized using the software DISTRUCT V 1.1 (Rosenberg 2004).

The second genetic clustering method, Discriminant Analysis of Principal Components (DAPC) (Jombart et al. 2010), uses a multivariate approach to visualize population differentiation with no assumptions of population genetic models. The analysis attempts to maximize between group variation among predefined groups. The DAPC analysis was conducted in the program R (R core team 2016) using the package ADEGENET (Jombart 2008).

Isolation-by-distance

Tests for conformance to an isolation-by-distance (IBD) gene flow model were conducted by evaluating the relationship between river distances and genetic distances by conducting a Mantel test (10,000 randomizations) in the software IBDWS v.3.23 (Jensen et al. 2005). Pairwise river distances (KM) between collection sites were calculated in GIS ArcMap 10.1 and genetic distances consisted of pairwise F'_{ST} .

Mitochondrial DNA

A subset of 236 speckled dace from all 25 sites (average of nine individuals per site, range 2–23 individuals) was sequenced for a 530-bp fragment of the mitochondrial cytochrome *b* gene (*cyt b*). Amplification of DNA was conducted with primers LA and HA (Dowling and Naylor 1997)

using the following thermal cycling routine: 35 cycles of 94 °C for 60 s, 48 °C for 60 s, and 72 °C for 120 s. The primer LA and three primers designed for this study were used for sequencing (Table S2). PCR products were purified and sequenced at High-Throughput Genomics Center (University of Washington, Department of Genome Sciences) using an Applied Biosystem 3730 xl sequencer. Chromatograms were visually inspected and manual corrections were made to the sequences. Sequences were aligned in MEGA6 (Tamura et al. 2013) using MUSCLE (Edgar 2004).

Mitochondrial DNA diversity and population structure

DNASP v5.10 (Librado and Rozas 2009) was used to estimate the number of haplotypes (H), the number of variable sites (S), nucleotide diversity (π), haplotype diversity (H_d), and average percent sequence divergence (D_{xy}). Uncorrected p -distances were calculated as the average number of nucleotide differences between site pairs using MEGA6 (Tamura et al. 2013).

A maximum-likelihood (ML) tree was generated using the software MEGA6, with branch support estimated via 1000 bootstrap pseudoreplicates. Publically available sequences were included to evaluate monophyly of speckled dace originating from the Klamath–Trinity River drainage and provide comparisons to other drainages. (Table S3). Additionally, we included a speckled dace haplotype from Dead Indian Creek (42.25194, – 122.4516) a tributary of Little Butte Creek in the Rogue River drainage, Oregon (Dowling et al. in press) (hereafter referred to as (TLB) and additional speckled dace haplotypes. *Rhinichthys atratulus* served as an out-group to root the tree (Dowling et al. 2002). The additional sequences were obtained from GenBank and Dryad Digital Repository (<https://doi.org/10.5061/dryad.ht554>; Table S3). The best model of sequence evolution was identified using jModelTest v2.1.4 (Posada 2008) based upon Akaike's Information Criterion corrected for small sample size AIC (Burnham and Anderson 2002). To visualize relationships among the mtDNA haplotypes, a 95% maximum parsimony haplotype network was constructed with TCS 1.21 (Clement et al. 2000).

Results

Microsatellites

A total of 1075 individuals were assayed in the microsatellite data with an average of 43 fish per location (range 30–48). Preliminary tests indicated that the locus CypG 13 consistently departed from Hardy-Wienberg expectations and therefore was removed. The final eight microsatellite loci were highly polymorphic, containing a total

of 258 alleles (mean 32.3 alleles/locus; range 5–78). Out of a total of 200 tests (8 loci and 25 populations) for conformance to Hardy–Weinberg expectations (HWE), nine were significant after Bonferroni correction (critical value = 0.00025, Rice 1989). A less conservative multiple test correction, the B-Y FDR method, found twenty-seven tests that were significant (critical value = 0.008506, Narum 2006). Departures from HWE were likely due to null alleles, as suggested by MICRO-CHECKER, but there was no evidence of stuttering, large allele dropout, or linkage disequilibrium in the loci. No single locus or site consistently departed from expectations, eliminating locus and site-specific factors as causes for the deviations.

Genetic diversity and population structure

Genetic diversity was higher in the Klamath group than in the Trinity group (Table 2). Mean expected heterozygosity (H_e) was 0.73 (range 0.69–0.74) in the Klamath and 0.52 in the Trinity (range 0.49–0.56). Mean rarified allelic richness (A_R) was 13.10 (11.82–14.26) in the Klamath and 9.08 (8.03–9.76) in the Trinity. Expected heterozygosity and rarified allelic richness were significantly lower in the Trinity in comparison to the Klamath ($P=0.001$). Among the three sites containing hybrids between the Trinity and Klamath groups, one had relatively high diversity (TT, $A_R = 12.55$, $H_e = 0.67$) whereas the other two locations in the Salmon River exhibited levels of diversity that were intermediate to the Klamath and Trinity groups (NFS: $A_R = 10.56$, $H_e = 0.62$; SFS: $A_R = 9.19$, $H_e = 0.59$). The genetically distinctive speckled dace from JEN had low values of diversity

Table 2 Genetic diversity at eight microsatellite loci for Klamath–Trinity Basin speckled dace, including genetic group, site, site abbreviation (ID), sample size (N), observed heterozygosity (H_o), expected heterozygosity (H_e), allelic richness (A), rarified allelic richness (A_R), and rarified number of private alleles (A_p)

| Genetic group | Site | ID | N | H_o | H_e | A | A_R | A_p |
|-------------------|--------------------------|------|----|-------|-------|------|-------|-------|
| Klamath | Seven Mile Creek | SEV | 44 | 0.65 | 0.74 | 15.0 | 12.5 | 0.15 |
| | Sprague River | SPR | 46 | 0.67 | 0.74 | 17.9 | 14.3 | 0.32 |
| | Link River | LNK | 46 | 0.7 | 0.74 | 14.1 | 11.8 | 0.13 |
| | Spencer Creek | SPE | 43 | 0.65 | 0.74 | 15.4 | 12.8 | 0.05 |
| | Copco 2 | COP | 47 | 0.66 | 0.73 | 16.5 | 13.2 | 0.17 |
| | Willow Creek | WIL | 44 | 0.72 | 0.72 | 15.4 | 12.4 | 0.16 |
| | Shasta River (upper) | SR | 30 | 0.67 | 0.69 | 13.5 | 12.3 | 0.22 |
| | Shasta River (lower) | SK | 48 | 0.72 | 0.74 | 17.6 | 13.6 | 0.23 |
| | Moffet Creek | MOF | 47 | 0.7 | 0.72 | 17 | 13.1 | 0.09 |
| | Lower Scott River | LSR | 47 | 0.67 | 0.74 | 17.4 | 13.8 | 0.05 |
| | Coon Creek | CN | 44 | 0.66 | 0.74 | 17.8 | 14.1 | 0.05 |
| | Big Bar | BB | 46 | 0.63 | 0.74 | 15.6 | 12.8 | 0.28 |
| | Blue Creek | BLU | 47 | 0.64 | 0.74 | 16.9 | 13.4 | 0.13 |
| | Mean | | 45 | 0.67 | 0.73 | 16.2 | 13.1 | 0.16 |
| Jenny Creek | Jenny Creek | JEN | 48 | 0.41 | 0.45 | 9.4 | 7.58 | 0.4 |
| Trinity | Del Loma | DL | 43 | 0.47 | 0.52 | 11.9 | 9.76 | 0.01 |
| | South Fork Trinity River | FG | 48 | 0.51 | 0.55 | 10.3 | 8.69 | 0.12 |
| | Canyon Creek | CAN | 31 | 0.47 | 0.49 | 8.5 | 8.03 | 0.05 |
| | Grass Valley Creek | GVY | 42 | 0.47 | 0.54 | 10.5 | 9.07 | 0.04 |
| | Stuart Fork Creek | SFK | 43 | 0.46 | 0.52 | 10.9 | 8.99 | 0.07 |
| | Swift Creek | SWFT | 46 | 0.45 | 0.5 | 10.3 | 8.85 | 0 |
| | Above Trinity Lake | TRL | 37 | 0.45 | 0.5 | 11.3 | 9.71 | 0.13 |
| | East Fork Trinity | EFT | 40 | 0.5 | 0.56 | 11.4 | 9.52 | 0.13 |
| | Mean | | 41 | 0.47 | 0.52 | 10.6 | 9.08 | 0.07 |
| Klamath X Trinity | Tish Tang Creek | TT | 40 | 0.6 | 0.67 | 14.5 | 12.6 | 0.04 |
| | North fork Salmon | NFS | 34 | 0.54 | 0.62 | 11.9 | 10.6 | 0.15 |
| | South Fork Salmon | SFS | 44 | 0.53 | 0.59 | 11.6 | 9.2 | 0.05 |
| | Mean | | 39 | 0.56 | 0.63 | 12.7 | 10.8 | 0.08 |

($Ar=7.58$, $H_e = 0.45$), compared to geographically proximate locations.

A total of 257 of the 300 pairwise F_{ST} values (range 0.00–0.345) were significant ($P \leq 0.000167$, Table S4). Significant F_{ST} was generally resolved in pairwise comparisons between Klamath and Trinity groups whereas comparisons among sites within each group were not generally significant. The mean standardized F'_{ST} for comparisons between Klamath and Trinity was 0.488, whereas the mean level of divergence among sites within the Klamath (0.051) and within the Trinity (0.029) were much lower (Table 3). The three Klamath X Trinity hybrid sites were more similar to the Trinity (mean standardized F'_{ST} 0.11) than the Klamath (0.406). JEN was resolved as divergent from all collection sites (range 0.539–0.625).

The unrooted neighbor-joining tree revealed two distinct clusters of speckled dace, including one containing all sites from the Klamath and a second including all sites from the Trinity. The Klamath X Trinity hybrids (TT, NFS and SFS) were resolved as most similar to the Trinity group (Fig. 2). Jenny Creek (JEN) was resolved as very divergent from all other populations and located just outside of Klamath cluster.

Bayesian cluster analysis using STRUCTURE suggested that the data was best described by two genetic clusters ($K=2$) using the ad hoc statistic ΔK (Figs. S1 and S2). All 20 replicate runs at $K=2$ resulted in an identical clustering pattern (Figs. 3 and S1). In the individual assignment plot the two distinct clusters were consistent with recognition of Klamath and Trinity genetic groups. The distribution of individual admixture proportions (q) indicated most individuals assigned to either pure Trinity ($q > 0.9$) or pure Klamath ($q < 0.1$) with 10% of individuals ($n = 109$) being assigned as hybrids ($0.1 < q < 0.9$) (Fig. 4a). However, hybrids were not assigned with high confidence as none of the individuals with intermediate q values had 90% probability intervals that were entirely contained with 0.1 and 0.9, except for three individuals from the Salmon (NFS and SFS) (Fig. 4b–d). Tish Tang (TT) contained the highest numbers of hybrids, with 48% of assayed individuals assigned as

hybrids (Fig. 4b). The Salmon River sites (NFS and SFS), showed lower levels of hybridization with 23% of individuals assigned as hybrids (Fig. 4c, d).

The Bayesian cluster analysis assuming $K=3$ distinct groups resulted two different solutions across the 20 independent STRUCTURE runs (Fig. 3). The most common solution resolved JEN as a third distinct cluster (19 of 20 runs) and the other solution resolved the Salmon River sites as distinct (1 of 20 runs). At $K=4$ there was evidence for multimodality and some solutions assignments were symmetric to all sites, suggesting that the number of clusters represented by the data was being overestimated (not shown). The multivariate DAPC analysis resolved similar results to the STRUCTURE analysis (Figs. S3, S4).

Isolation-by-distance

Tests for isolation-by-distance (IBD) were conducted separately for the Klamath and Trinity groups due to the large divergence resolved between them. The Klamath X Trinity hybrids were grouped in the Trinity IBD analysis as they were more closely aligned with the Trinity group in the Bayesian cluster analysis. The relationship between pairwise genetic distances (F'_{ST}) and river distance was significant for both the Klamath and Trinity groups ($p < 0.05$) and the intercept was essentially zero in both cases, which is consistent with a IBD model of gene flow (Fig. S4). However, geographic distance only explained 15% of the variation in genetic differentiation between sites in the Klamath ($R^2 = 0.145$, $P = 0.0115$), whereas 58% of the variation in genetic differentiation was explained for the Trinity ($R^2 = 0.581$, $P = 0.0005$). The slope of Trinity IBD relationship ($6.350e-04$) was nearly twice that of the Klamath IBD relationship ($3.550e-04$). Trinity IBD analysis performed without the Salmon sites (NFS and SFS) resolved the same results (data not shown). Jenny Creek (JEN) was excluded for tests of IBD because of its deep level of divergence from all sites.

Mitochondrial results

For mtDNA, the average number of fish sequenced per location was nine and ranged from 2 to 23 (Table 4). A total of 79 unique haplotypes were defined by 88 variable nucleotide positions (Genbank accession numbers MF066096–MF066172). Mean nucleotide diversity was 1.91% in the mtDNA dataset. Within collection sites, the number of haplotypes ranged from 2 to 9 and nucleotide diversity ranged from 0.0 to 3.1%. Jmodel test selected the Kimura two-parameter model (K80) of sequence evolution as the best fit to describe the mtDNA sequence data.

Similar to the pattern observed in the microsatellite data, the Klamath displayed higher levels of mtDNA

Table 3 Mean pairwise standardized genetic differentiation (F'_{ST}) at eight microsatellite loci among the main genetic groups of Klamath–Trinity Basin speckled dace

| | Klamath | Trinity | Klamath X Trinity | Jenny Creek (JEN) |
|-------------------|--------------|--------------|-------------------|-------------------|
| Klamath | <i>0.051</i> | | | |
| Trinity | 0.488 | <i>0.029</i> | | |
| Klamath X Trinity | 0.406 | 0.113 | <i>0.042</i> | |
| Jenny Creek (JEN) | 0.539 | 0.592 | 0.625 | N/A |

Numbers in italics along the diagonal are mean within group F'_{ST}

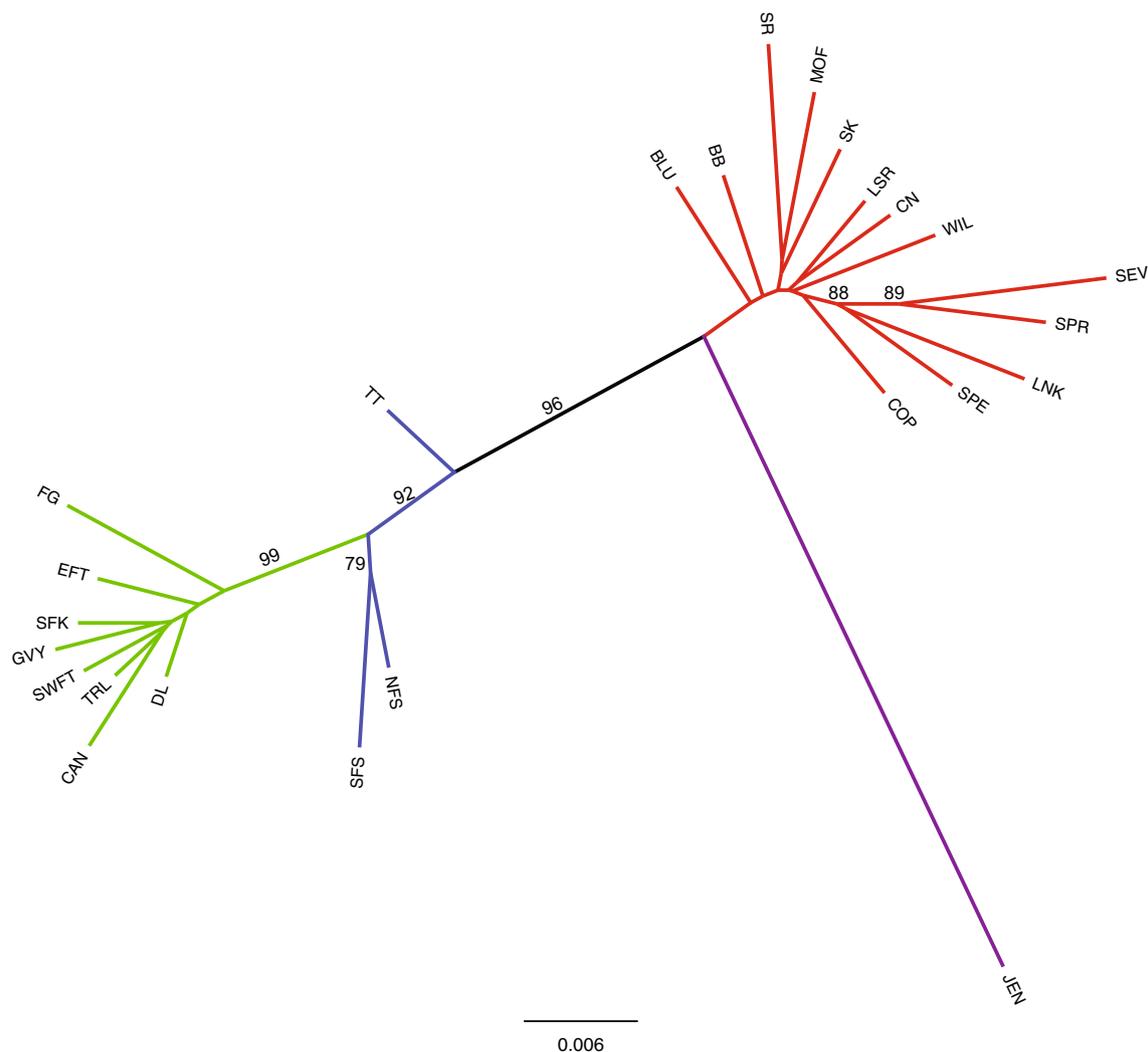


Fig. 2 Unrooted neighbor-joining tree based on the Cavalli-Sforza and Edwards's chord distances calculated using eight microsatellite loci assayed in Klamath-Trinity Basin speckled dace. The numbers on the branches are bootstrap values from 1000 replicates. Only

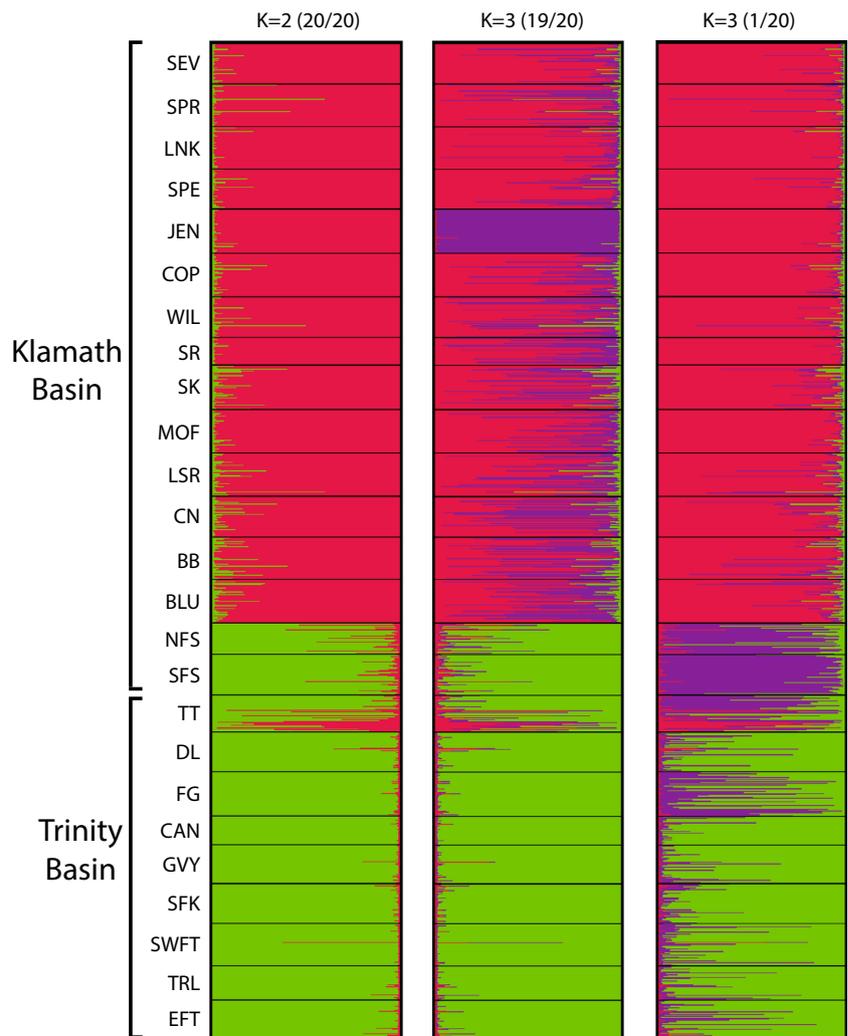
values above 75 are shown. Trinity branches are green, Klamath branches are red, Jenny Creek (JEN) is purple and the Klamath X Trinity hybrids (TT, NFS, and SFS) are blue. (Color figure online)

diversity than the Trinity (Table 4). The Klamath group contained 47 haplotypes ($N = 103$) with a mean haplotype diversity of 0.88, while Trinity contained 21 haplotypes ($N = 84$) with a mean haplotype diversity of 0.54. Mean nucleotide diversity was also lower in the Trinity (0.10%) compared to Klamath (1.0%). The Trinity X Klamath hybrids contained 13 haplotypes ($N = 36$) with a mean haplotype diversity of 0.57 and a mean nucleotide diversity of 0.7%. Among the Klamath X Trinity hybrids, the Salmon River populations (NFS and SFS) had reduced levels of mtDNA diversity, while TT had elevated levels of diversity. Jenny Creek (JEN) contained four haplotypes ($N = 12$) with a haplotype diversity of 0.561 and a mean nucleotide diversity of 0.2%.

Mitochondrial structure

The Klamath–Trinity Basin speckled dace were resolved as nearly monophyletic [bootstrap (BS) 77], and exhibited a sister group relationship with nearby basins in California and Oregon (Sacramento, Pit River, and Goose Lake Basin, Fig. 5). There were two exceptions to monophyly for Klamath–Trinity speckled dace: (1) a divergent haplotype (sequence divergence $> 5.8\%$) identified in a single individual found in Seven Mile Creek, a tributary to Klamath Lake, which was resolved in a basal position in our tree (hereafter referred to as SEV01), and (2) a haplotype from the tributary of Little Butte Creek (TLB), Rogue River Basin, that aligned with Jenny Creek (JEN) haplotypes.

Fig. 3 Results from Bayesian cluster analysis based upon eight microsatellite loci for Klamath–Trinity Basin speckled dace. Each individual is displayed as a thin horizontal line divided into sections, whose length is equal to the probability of membership to a cluster (q) while populations are differentiated by thick black lines. Two genetic clusters ($K=2$) was the most probable number of clusters based on ad hoc statistic ΔK (Fig. S1) and all 20 iterations showed the same genetic clustering. Assuming three clusters ($K=3$) displayed evidence of multimodality, 19/20 iterations had Jenny Creek as a distinct cluster while the remaining iteration found the Salmon River populations (NFS and SFS) distinct



Excluding the (SEV01) haplotype, the Klamath–Trinity Basin contained three primary clades: (1) Trinity (BS = 75), (2) Klamath (BS = 90), and (3) Jenny (BS = 70) (Fig. 5). The Trinity was composed exclusively of individuals that originated from the Trinity River. The Klamath included the majority of the individuals collected from Klamath River and 10 of the 15 speckled dace from the hybrid location in the lower Trinity River, TT, and the hybrids from NFS, and SFS. The Jenny group was comprised all 12 individuals examined from JEN, five individuals from the upper Klamath River and the two individuals examined from a tributary to Little Butte Creek (TLB). The Trinity was resolved as sister to Jenny (BS = 89) and Klamath clade was sister to this group. The close relationship between the Trinity clade and the Jenny clade was unexpected given the distances involved and the geographic proximity of the Klamath and Jenny groups. The Klamath contained considerably more within-group structuring than the Trinity, including a well-supported (BS = 88) branch containing haplotypes from above Upper Klamath Lake in Oregon.

In the 95% maximum parsimony network Klamath–Trinity Basin speckled dace were divided into two unconnected primary networks (excluding the single divergent haplotype SEV01, Fig. 6). The same three clades resolved by the ML tree were evident: (1) Trinity, (2) Klamath, and (3) Jenny. The SEV01 haplotype (not shown) was highly divergent from all other groups (5.87–6.16%). The Klamath displayed considerable structure in the haplotype network and possessed the largest within-group percent sequence diversity (1.22%, Table 5). The Trinity had the lowest within group percent sequence diversity (0.150%) and contained one primary haplotype at high frequency in all populations. Klamath and Trinity displayed the largest sequence divergence from one another (2.96%), followed by Klamath versus the Jenny group (2.64%), and lowest was Klamath versus Klamath X Trinity hybrids (1.21%). Among the Klamath X Trinity hybrids NFS and SFS contained only Klamath haplotypes whereas TT contained a mixture of Klamath and Trinity haplotypes and a sequence diversity of 1.06%.

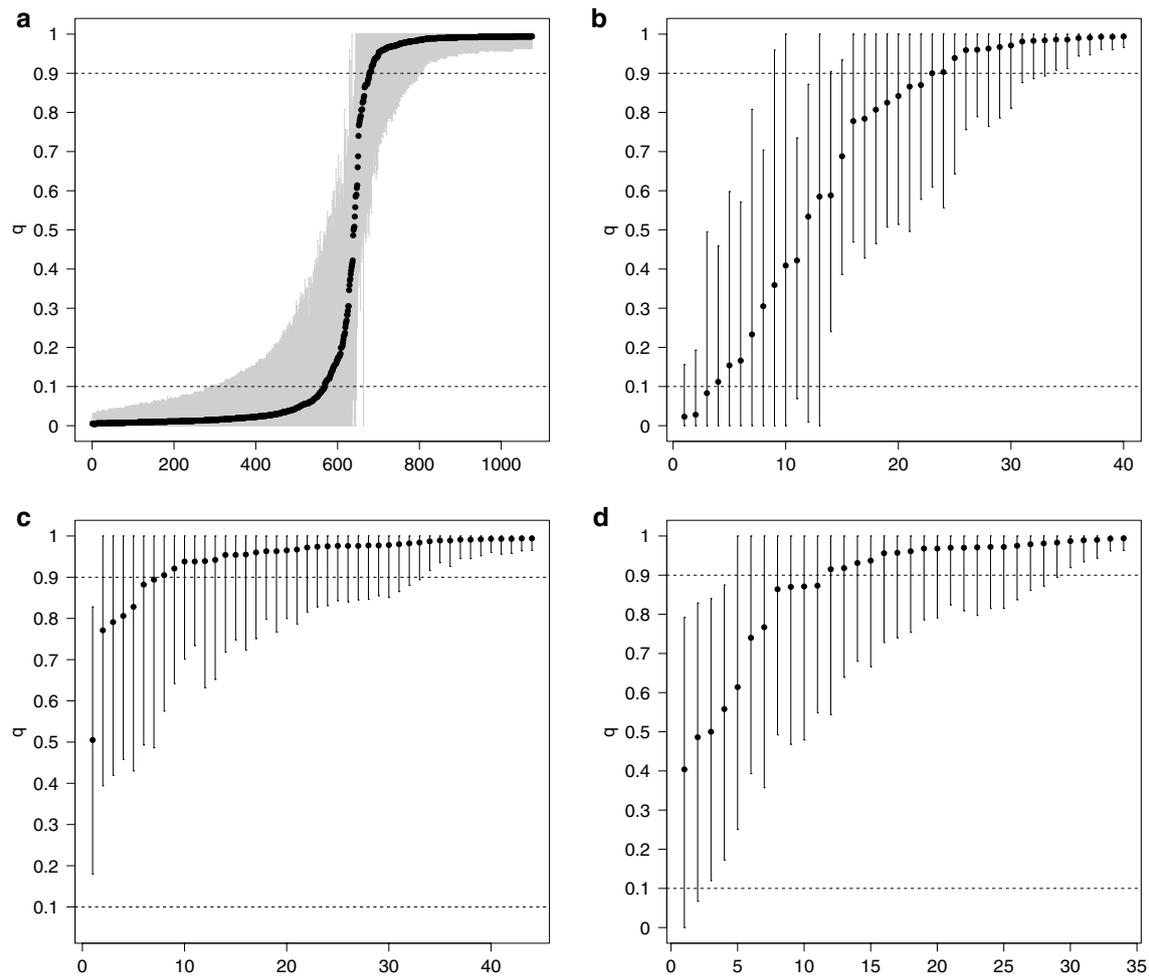


Fig. 4 Individual admixture proportions (q) estimated by Bayesian cluster analysis ($K=2$) of eight microsatellite loci assayed in Klamath–Trinity Basin speckled dace. Shown are point estimates of q and 90% probability intervals for **a** all individuals in the data set,

b Tish Tang Creek (TT), **c** North Fork Salmon (NFS), and **d** South Fork Salmon (SFS). Individual q values close to zero and one signify Klamath and Trinity clusters, respectively. Individual q values are sorted from low to high for presentation

Discussion

Collections from throughout the Klamath–Trinity basin combined with analysis of both mitochondrial DNA and microsatellite loci revealed complex within-basin patterns including evidence for multiple cryptic groups within *R. o. klamathensis*. Overall, our analysis indicates the presence of three genetically distinct groups of speckled dace within the Klamath–Trinity Basin: (1) Klamath, (2) Trinity, and (3) Jenny, each named according to its geographic distribution. The most notable result of this study was the extent of genetic divergence among the speckled dace lineages occurring in the Klamath–Trinity Basin. The level of genetic divergence in mtDNA cyt b (1.38–6.16%) between groups found in the Klamath–Trinity Basin is comparable to divergences found between recognized *Rhinichthys* species (McPhail and Taylor 2009).

Klamath and Trinity groups

The Klamath group is distributed throughout the Klamath River and its tributaries and the Trinity group is distributed in the Trinity River and its tributaries (Fig. 1). The area near the confluence of the Klamath and Trinity rivers and the Salmon River appear to be zones of contact and hybridization between the Klamath and Trinity groups (see below).

The precise biogeographic processes in the Klamath–Trinity Basin that have impacted speckled dace populations during the Pliocene–Pleistocene are unclear because the substantial number of geomorphic rearrangements that have occurred in this geologically active region over the past few million years (Minckley et al. 1986; Aalto 2006). Furthermore, repeated glacial advance, retreat and associated erosion make reconstruction of historical river drainage patterns uncertain (Anderson 2008). The evidence

Table 4 Mitochondrial DNA sequence diversity for Klamath–Trinity speckled dace including genetic group, site abbreviation (ID), number of sequences (N), number of variable sites in the sequences (S), sequence diversity (π), unique haplotypes (H), and haplotype diversity (Hd)

| Genetic Group | ID | N | S | π | H | Hd |
|-------------------|------------------|----|------|-------|-------|-------|
| Klamath | SEV ^a | 5 | 17 | 0.018 | 3 | 0.800 |
| | SPR | 10 | 13 | 0.007 | 7 | 0.911 |
| | LNK | 7 | 25 | 0.018 | 7 | 1 |
| | SPE | 9 | 23 | 0.012 | 7 | 0.944 |
| | COP | 8 | 11 | 0.007 | 6 | 0.893 |
| | WIL | 8 | 11 | 0.009 | 6 | 0.893 |
| | SR | 16 | 16 | 0.008 | 9 | 0.817 |
| | SK | 8 | 9 | 0.005 | 7 | 0.964 |
| | MOF | 8 | 9 | 0.008 | 3 | 0.607 |
| | LSR | 7 | 13 | 0.009 | 7 | 1 |
| | CN | 7 | 9 | 0.005 | 4 | 0.714 |
| | BB | 2 | 7 | 0.013 | 2 | 1 |
| | BLU | 8 | 11 | 0.009 | 6 | 0.929 |
| | Mean | | 8 | 13.4 | 0.010 | 5.7 |
| Jenny | JEN | 12 | 4 | 0.002 | 4 | 0.561 |
| Trinity | DL | 8 | 4 | 0.002 | 5 | 0.786 |
| | FG | 15 | 7 | 0.003 | 7 | 0.857 |
| | CAN | 23 | 5 | 0.001 | 6 | 0.518 |
| | GVY | 7 | 4 | 0.002 | 4 | 0.714 |
| | SFK | 8 | 1 | 0.000 | 2 | 0.25 |
| | SWFT | 8 | 3 | 0.001 | 3 | 0.464 |
| | TRL | 7 | 1 | 0.001 | 2 | 0.286 |
| | EFT | 8 | 2 | 0.001 | 3 | 0.464 |
| | Mean | | 10.5 | 3.38 | 0.001 | 4.0 |
| Klamath X Trinity | TT | 15 | 26 | 0.018 | 9 | 0.905 |
| | NFS | 11 | 4 | 0.002 | 4 | 0.6 |
| | SFS | 10 | 1 | 0.000 | 2 | 0.2 |
| | Mean | | 12 | 10.33 | 0.007 | 5 |

^aThe divergent SEV01 haplotype was removed for diversity calculations

that remains suggests that the northwest Klamath Mountains are the result of recent uplift during Pliocene–Pleistocene and before that, the area was a low-lying peneplain (Aalto 2006). Before this uplift, the Trinity and Klamath Rivers presumably had different drainage patterns than they do today, though the nature of their flow and outlet are obscured (Anderson 2008).

We hypothesize that the Klamath and Trinity groups diverged in allopatry and are now in secondary contact for several reasons. First, the estimated divergence time between the Klamath and Trinity groups ranges from 1.18 to 2.47 million years ago based upon Smith and Dowling’s (2008) estimate of cyt b mutation rate of 1.2–2.5% per million years for speckled dace. This estimated Pleistocene divergence time for Klamath and Trinity groups is consistent with the estimated time of uplift for the northwest

Klamath Mountains (Aalto 2006). Second, the confinement of the Klamath and Trinity groups to their respective river systems, and the sharp northward bend of the Trinity River where it connects with the Klamath River, suggests that these systems were once separate. Finally, the extent of divergence (~3%) between the Trinity and Klamath groups is similar to levels observed between speckled dace isolated in different river systems (Pfrender et al. 2004; Ardren et al. 2010). All potential within-basin barriers, such as Burnt Ranch Gorge on the Trinity River and Ishi Pishi Falls on the Klamath River, are surmountable to fishes and unlikely serve as long-term migration barriers to speckled dace, which have appreciable dispersal capacity for their size (Pearsons et al. 1992; Brown and Moyle 1997).

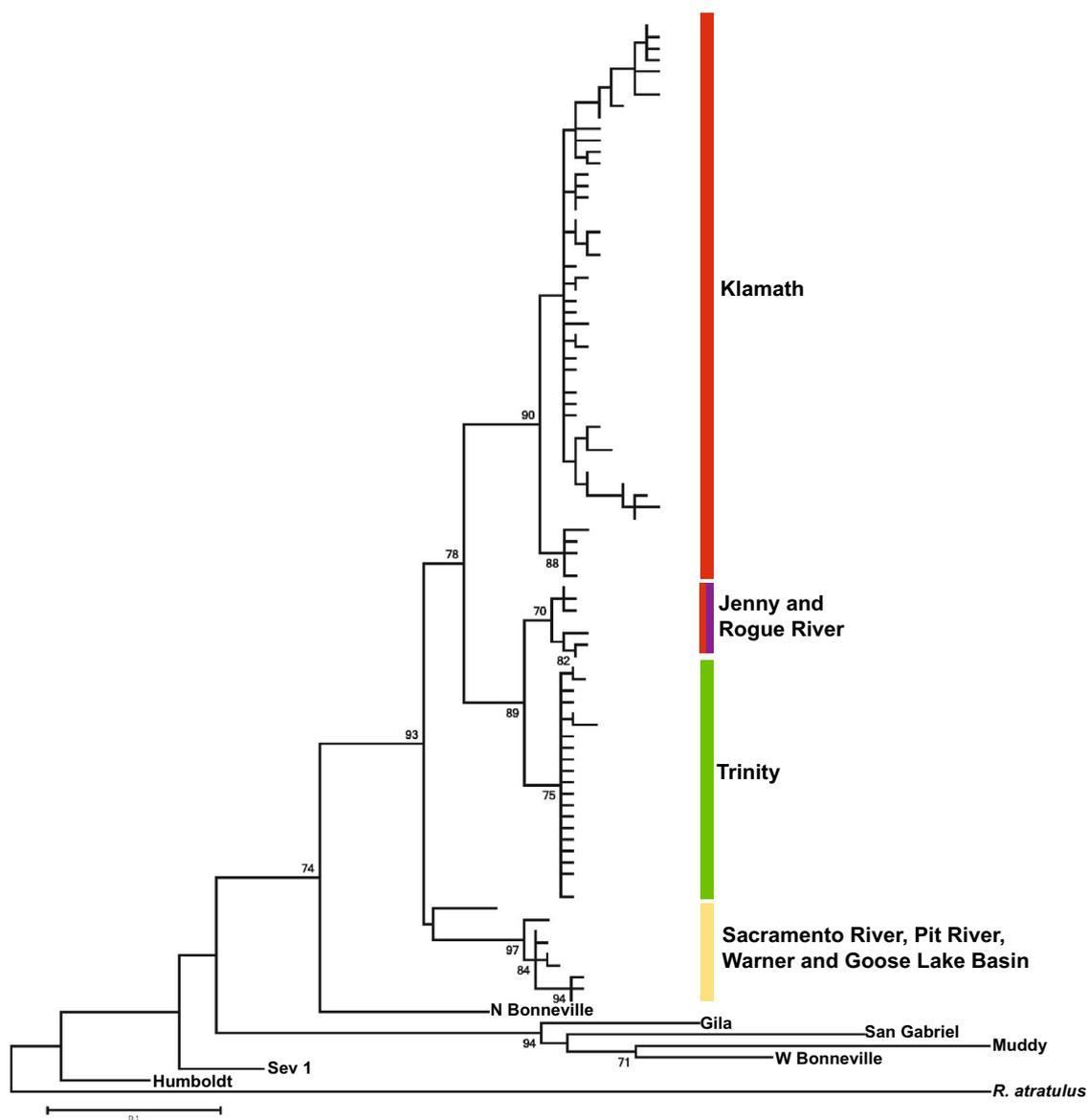


Fig. 5 Maximum Likelihood tree generated from unique mitochondrial cytochrome *b* haplotypes for Klamath–Trinity speckled dace and publicly available speckled dace haplotypes. A single *R. atratulus*

sequence served as an out-group. Support for the tree was established via 1000 bootstrap replicates and only values above 70 are shown

Klamath X Trinity hybridization

Most individuals examined from the Klamath and Trinity rivers were resolved as non-admixed as indicated by the bimodal distribution of individual admixture coefficients (Fig. 4). However, our analysis suggested the existence of contact and hybridization between Klamath and Trinity groups in the geographic region near the confluence of the Klamath and Trinity Rivers, including Tish Tang Creek (TT), and the Salmon River (NFS and SFS).

Hybridization in Tish Tang Creek was suggested in the analysis of microsatellite loci by the intermediate positioning

of this collection between the Klamath and Trinity groups in the tree-based analysis (Fig. 2). Further in Bayesian cluster analysis a large number (48%) of individuals from TT had intermediate admixture coefficients ($0.1 < q < 0.9$), although none of the hybrid individuals had probability intervals that could be used to make confident assignments of hybridization (Fig. 4). In the mtDNA analysis, Tish Tang Creek was the only location studied that contained a mix of haplotypes originating from the Klamath and Trinity groups, a pattern also suggestive of hybridization (Fig. 6). Hybridization in this region is in accord with geography, as pure Trinity and Klamath appear to be confined to their respective river

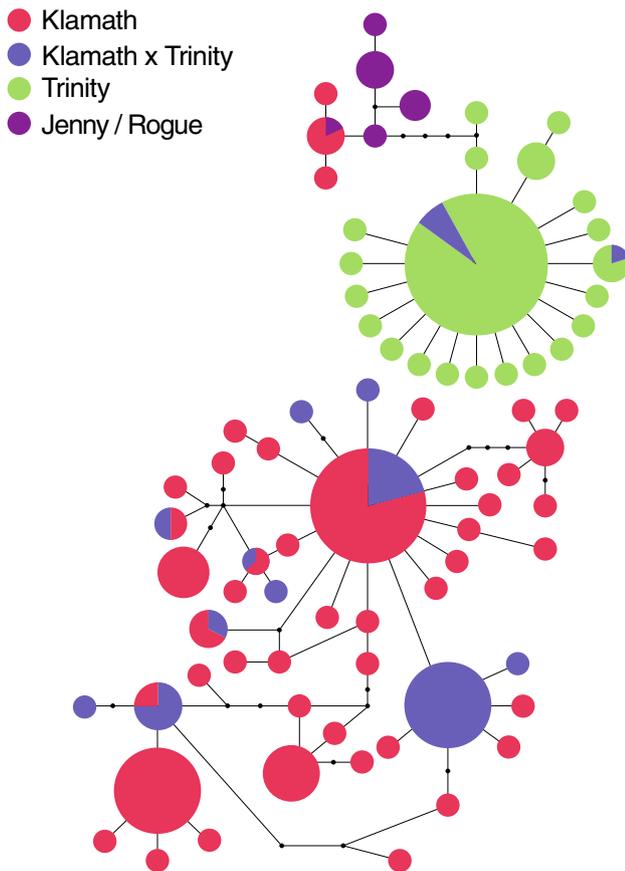


Fig. 6 Two unconnected ninety-five percent parsimony haplotype networks based on mitochondrial cytochrome *b* sequences for Klamath–Trinity Basin speckled dace. Circles represent unique haplotypes and the diameter of the circle corresponds to relative haplotype frequency. Small black dots represent inferred haplotypes. The circles are color coded to display the proportion of individuals from the Klamath (red), Trinity (green), Jenny/Rogue (purple) or Klamath X Trinity (blue) with the haplotype. (Color figure online)

Table 5 Average percent sequence divergence of mitochondrial cyt *b* sequences (D_{xy}) in Klamath–Trinity Basin speckled dace.

| | Klamath | Trinity | Klamth X Trin-ity | Jenny/Rogue | SEV01 |
|-------------------|-------------|--------------|-------------------|-------------|------------|
| Klamath | <i>1.22</i> | | | | |
| Trinity | 2.96 | <i>0.150</i> | | | |
| Klamath X Trinity | 1.21 | 2.55 | <i>1.08</i> | | |
| Jenny/Rogue | 2.64 | 1.38 | 2.53 | <i>0.31</i> | |
| SEV01 | 6.15 | 5.87 | 6.05 | 6.16 | <i>N/A</i> |

Sequence divergence among the main genetic groups are below the diagonal and within group nucleotide diversity in percent (π) are along the diagonal in italics. The SEV01 haplotype was only detected in one individual

systems and presumably contact each other near the confluence of the two rivers.

Hybridization between the Klamath and Trinity groups in the Salmon River (NFS and SFS) was indicated by discordance between nuclear and mtDNA markers (Avisé 2004). All Salmon River speckled dace had Klamath mtDNA haplotypes ($n = 21$), whereas nuclear microsatellite analysis aligned the Salmon River populations with the Trinity group (Figs. 2, 3). The individual admixture coefficients from Bayesian cluster analysis indicated the presence of several individuals in the Salmon River with probability intervals contained entirely within $0.1 < q < 0.9$, (Fig. 4c, d) indicating the existence of hybridization between the Klamath and Trinity lineages within the Salmon River.

The geographic mechanism responsible for creating contact and hybridization in the Salmon River is unclear, as sites we examined near the confluence of the Klamath and Salmon Rivers (BB and CN) appeared to consist of pure Klamath individuals, suggesting the absence of contemporary contact between the Trinity and Klamath groups through the mouth of the Salmon River (Fig. 1). However, hybrid zones can be dynamic in space and time (Buggs 2007), and it is conceivable that Trinity and Klamath groups may have historically contacted one another through the lower Salmon River but they no longer do so. Another possibility is that the contact between the lineages may have resulted from headwater capture. There is a well-established river capture wherein a Trinity River tributary was diverted into the South Fork Salmon River during the Pleistocene (Hershey 1900; Sharp 1960; Fig. 1). Finer-scale sampling around the Salmon River and the confluence of the Trinity and Klamath rivers combined with application of molecular markers with improved resolution for hybrid identification are needed to more clearly define the geographic extent, frequency, and patterns of hybridization.

The mechanism(s) responsible for limiting hybridization to a relatively small geographic region (e.g., near confluence of the Klamath and Trinity rivers and the Salmon River) and maintaining the Klamath and Trinity groups from outside of this area as pure is uncertain. While hybrid zones can develop along environmental gradients (Harrison 1990), there are no readily apparent environmental gradients across the Trinity and Klamath rivers. Alternatively, interactions with predators have been shown to be important in constraining the distribution of speckled dace (Baltz et al. 1982; Harvey et al. 2004) as have water velocity and spawning substrate (Smith and Dowling 2008; Peden and Hughes 1981).

Jenny group

The analysis of both mtDNA and microsatellites indicated that the Jenny group was genetically distinct in comparison to the Klamath and Trinity groups (see also Pfrender

et al. 2004). The Jenny group is distributed in two disjunct river basins, Jenny Creek, a tributary to the upper Klamath River, and a tributary of Little Butte Creek (TLB) of the Rogue River Basin (Fig. 1). We hypothesize that the Jenny group originally diverged within Jenny Creek and was recently introduced from Jenny Creek into TLB via artificial or natural means. The original genetic divergence presumably resulted from isolation caused by a 10-m waterfall created by Pleistocene lava flow in the lower portions of Jenny Creek (Hohler 1981). A dwarf form of the Klamath smallscale sucker (*Catostomus rimiculus*) is also isolated to Jenny Creek (Hohler 1981).

Our analysis indicated (1) the presence of JEN mtDNA haplotypes in the Upper Klamath River, and (2) nuclear admixture between the Jenny and Klamath groups in the Upper Klamath River (Fig. 3). These patterns may be the result of one-way downstream leakage of the Jenny group into the upper Klamath River. Alternatively, the Jenny Creek group may once have been present in the upper Klamath River but experienced significant drift within Jenny Creek due to isolation and small population numbers. Jenny Creek's isolation and lack of outside contact could also explain the low levels of genetic diversity we observed (Tables 2, 5).

The genetic similarity in mtDNA between JEN and TLB supports a recent human-induced transfer between these rivers, as hypothesized by Bond (1994). Artificial means of transfer of Jenny Creek speckled dace into TLB may have resulted from water connections created by the US Bureau of Reclamation. However, a natural introduction via head-water capture is possible, as TLB and Jenny Creek systems are geographically close (3 m vertically and 0.4 km horizontally). Available evidence suggests the Jenny group is restricted to Jenny Creek and TLB, but does not occur elsewhere in the Rogue River basin as Pfrender et al. (2004) found that speckled dace from Little Butte Creek were more similar to those from other coastal Oregon streams than Jenny Creek. Jenny Creek speckled dace appear to represent a unique and isolated form that has gone undetected by traditional taxonomic surveys. We recommend incorporating collections from nearby coastal basins (Coos, Umpqua) and utilizing additional genetic markers to clarify Jenny Creek genetic relationships and geographic distribution.

SEV01 mitochondrial DNA haplotype

We detected a single mitochondrial DNA haplotype (SEV01) from the upper Klamath River that was highly differentiated from all other Klamath–Trinity Basin haplotypes (Fig. 5). Pfrender et al. (2004) identified the same haplotype from upper Klamath River but in our analysis we only resolved this haplotype in 1.67% of individuals examined (1/60) in the Upper Klamath whereas Pfrender et al. (2004) resolved

this haplotype at a higher frequency, 12.73% (7/55). In contrast to the mtDNA results, the individual with the SEV01 haplotype was resolved as a member of the Klamath group by STRUCTURE analysis of nuclear microsatellites. This discordance between markers suggests a hybridization event between the Klamath group and the SEV01 lineage. Pfrender et al. (2004) proposed that the SEV01 lineage and the main Klamath lineage may represent two reproductively isolated forms co-occurring in the Upper Klamath River but our analysis does not support this hypothesis. While further study is needed to understand the origin of SEV01 and its relationship with the main Klamath lineage, the low frequency of the haplotype may hamper research.

Phylogenetic relationships

Phylogenetic analysis of mitochondrial DNA generally resolved Klamath–Trinity Basin speckled dace as cohesive group excluding the SEV01 haplotype and the Jenny group. The close relationship of Klamath–Trinity Basin speckled dace to the Sacramento, Pit, Goose Lake and Warner Basin likely reflect the historic connections these drainages had with the hypothesized westward flowing Proto-Snake River (Oakey et al. 2004; Arden et al. 2010). The close association the Oregon Great basin drainages (Goose and Warner) share with the Pit and Sacramento Basins is explained by the historic Goose Lake overflows that spilled into the Pit River (Baldwin 1981).

Phylogenetic relationships among the Klamath, Trinity, and Jenny groups were discordant between mitochondrial and nuclear markers, suggestive of gene-tree and species-tree discordance. In the analysis of mitochondrial DNA, Jenny Creek was resolved as sister to the Trinity group but in analysis of nuclear microsatellites, Jenny Creek appeared more closely related to the Klamath group (Fig. 6). Incomplete lineage sorting of mitochondrial DNA or a hybridization event between Jenny Creek speckled dace and other forms are potential explanations for these patterns (Ballard and Whitlock 2004; Carstens and Knowles 2007; Waters et al. 2010). These findings highlight that using single genes, or even two genes, may fail to resolve true species relationships and that multiple independent genes are needed to resolve true species relationships (Maddison and Knowles 2006; Heled and Drummond 2010).

Within-group genetic structure

Our analysis indicated the Klamath group contained significantly higher levels of mitochondrial DNA and nuclear microsatellite diversity than the Trinity group. This finding is in accord with a recent survey of genetic diversity in speckled dace across 10 northern California sites which indicated that Klamath group speckled dace had higher genetic

diversity than the other sites studied (Kinziger et al. 2011). The higher level of genetic diversity within the Klamath group is consistent with biogeographic patterns, which support the upper Klamath River as a zoogeographic province for fishes (Moyle 2002).

Our field sampling specifically targeted sites isolated by impassable dams but these barriers produced no appreciable within-group genetic structuring in our study (Fig. 1). Based on the high level of genetic diversity detected in this study, the effective population sizes are most likely too large for the dams to have impacted isolated populations via drift. However, the Upper Klamath River sites (SEV, SPR, LNK, and SPE) did exhibit some minor sub-structuring (Fig. 2). The Klamath and Trinity group both exhibited conformance an IBD model of gene flow, but the Klamath group displayed a weak IBD relationship with less within river system genetic differentiation compared to the Trinity, despite the greater distances in the Klamath (Fig. S5). The Trinity group stronger IBD relationship could be explained by the steep gradients and high elevations, which have created physical barriers to gene flow (Castric et al. 2001).

Conclusion

The discovery of previously unknown cryptic diversity of speckled dace in the Klamath–Trinity Basin illustrates the usefulness of molecular studies for cataloging the planet's biodiversity. We recommend additional morphologic and ecological studies to determine if the cryptic genetic groups may warrant formal taxonomic recognition in the future. Fine-scale sampling throughout the Klamath–Trinity Basin combined with mitochondrial and nuclear markers was crucial for discerning genetic structure and cryptic diversity. The genetic structure of speckled dace in the Klamath–Trinity Basin and other western basins is much more complicated than contemporary drainage patterns would suggest, reflecting historical climatic and geological upheaval of the west. Our study and other recent studies (Hoekzema and Sidlauskas 2014) have uncovered far more diversity within speckled dace than previously thought. Possibly, the single taxonomic designation of speckled dace needs to be revised to accurately represent this diversity.

Acknowledgements The authors would like to thank the following for help with field collections: Conrad Newell, Sam Rizza, and Robbie Mueller (Humboldt State University (HSU)), Bret Harvey (US Forest Service), Rodney Nakamoto (US Forest Service), Bill Tinniswood (Oregon Department of Fish and Wildlife). Thanks to Tom Hutson for help drafting Fig. 6, Chloe Joesten (HSU) for her help in the laboratory and to Dana Herman (HSU) for providing assistance with GIS analysis. Thanks to Stewart Reid (Western Fishes Inc.) for his assistance and collection of Rogue River speckled dace and to Thomas Dowling (Wayne State University) for supplying the Rogue River mtDNA sequences.

Additionally, we would like to thank the three anonymous reviewers for their comments on previous versions of this manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Aalto KR (2006) The Klamath peneplain: a review of JS Diller's classic erosion surface. *Geol Soc Am Special Papers* 410:451–463
- Abell RM, Thieme ML, Revenga C, Bryer M, Kottelat M, Bogutskaya N, Coad B, Mandrak K, Contreras Balderas S, Bussing W, Stiassny M, Skelton P, Allen R, Unmack P, Naseka A, Ng R, Sindorf N, Robertson J, Armijo E, Higgins J, Heibel J, Wikramanake E, Olson D, Lopez H, Reis R, Lundberg J, Sabaj Perze M, Petry P (2008) Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation. *Bioscience* 58:403–414. <https://doi.org/10.1641/B580507>
- Anderson TK (2008) Inferring bedrock uplift in the Klamath Mountains Province from river profile analysis and digital topography. Master's thesis, Texas Tech University
- Ardren WR, Baumsteiger J, Allen CS (2010) Genetic analysis and uncertain taxonomic status of threatened Foskett Spring speckled dace. *Conserv Genet* 11:1299–1315. <https://doi.org/10.1007/s10592-009-9959-0>
- Avise JC (2004) Molecular markers, natural history and evolution. Springer Science & Business Media, Chicago
- Baerwald MR, May B (2004) Characterization of microsatellite loci for five members of the minnow family Cyprinidae found in the Sacramento–San Joaquin Delta and its tributaries. *Mol Ecol Notes* 4:385–390. <https://doi.org/10.1111/j.1471-8286.2004.00661.x>
- Baldwin EM (1981) Geology of Oregon. Kendall/Hunt Publishing, Iowa
- Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Mol Ecol* 13:729–744. <https://doi.org/10.1046/j.1365-294X.2003.02063.x>
- Baltz DM, Moyle PB, Knight NJ (1982) Competitive interactions between benthic stream fishes, riffle sculpin, *Cottus gulosus*, and speckled dace, *Rhinichthys osculus*. *Can J Fish Aquat Sci* 39:1502–1511. <https://doi.org/10.1139/f82-202>
- Bernardo J (2011) A critical appraisal of the meaning and diagnosability of cryptic evolutionary diversity, and its implications for conservation in the face of climate change: the systematics association, Cambridge University Press, Cambridge
- Bickford D, Lohman DJ, Sodhi NS, Ng P, Meier R, Winker K, Ingram K, Das I (2007) Cryptic species as a window on diversity and conservation. *Trends Ecol Evol* 22:148–155. <https://doi.org/10.1016/j.tree.2006.11.004>
- Billman EJ, Lee JB, Young DO, McKell M, Evans R, Shiozawa D (2010) Phylogenetic divergence in a desert fish: differentiation of speckled dace within the bonnevillie, lahontan, and upper snake river basins. *Western North Am Nat* 70:39–47
- Bond CE (1994) Keys to Oregon freshwater fishes. OSU Book Stores, Corvallis
- Brown LR, Moyle PB (1997) Invading species in the Eel River, California: successes, failures, and relationships with resident species. *Environ Biol Fishes* 49:271–291. <https://doi.org/10.1023/A:1007381027518>
- Buggs RJA (2007) Empirical study of hybrid zone movement. *Heredity* 99:301–312

- Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach. Springer, Berlin
- Carstens BC, Knowles LL (2007) Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus grasshoppers*. *Syst Biol* 56:400–411. <https://doi.org/10.1080/10635150701405560>
- Castric V, Bonney F, Bernatchez L (2001) Landscape Structure and Hierarchical Genetic diversity in the brook charr *Salvelinus Fontinalis*. *Evolution* 55:1016–1028. <https://doi.org/10.1111/j.0014-3820.2001.tb00618.x>
- Chapin FS III, Zavaleta ES, Eviner VT et al (2000) Consequences of changing biodiversity. *Nature* 405:234–242. <https://doi.org/10.1038/35012241>
- Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9:1657–1659
- Dowling TE, Naylor GJ (1997) Evolutionary relationships of minnows in the genus *Luxilus* (Teleostei: Cyprinidae) as determined from cytochrome b sequences. *Copeia* 1997(4):758–765
- Dowling TE, Tibbets CA, Minckley WL, Smith GR, McEachran JD (2002) Evolutionary relationships of the plagioperins (Teleostei: Cyprinidae) from cytochrome b sequences. *Copeia* 2002:665–678
- Dudgeon D, Arthington AH, Gessner MO et al (2006) Freshwater biodiversity: importance, threats, status and conservation challenges. *Biol Rev* 81:163–182. <https://doi.org/10.1017/S1464793105006950>
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361. <https://doi.org/10.1007/s12686-011-9548-7>
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol* 14:2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol Bioinformatics* 1:47–50
- Felsenstein J (2005) Phylip (phylogeny inference package) version 3.6 (2004) distributed by the author. Department of Genome Sciences, University of Washington, Seattle
- Girard P, Angers B (2006) Characterization of microsatellite loci in longnose dace (*Rhinichthys cataractae*) and interspecific amplification in five other *Leuciscinae* species. *Mol Ecol Notes* 6:69–71. <https://doi.org/10.1111/j.1471-8286.2005.01140.x>
- Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. *J Hered* 86(6):485–486
- Harrison RG (1990) Hybrid zones: Windows on evolutionary process, vol 7. In: Futuyama D, Antonovics J (ed) Oxford surveys in evolutionary biology. Oxford University Press, Oxford, pp 69–128
- Harvey BC, White JL, Nakamoto RJ (2004) An emergent multiple predator effect may enhance biotic resistance in a stream fish assemblage. *Ecology* 85:127–133. <https://doi.org/10.1890/03-3018>
- Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Mol Biol Evol* 27:570–580
- Hershey OH (1900) Ancient alpine glaciers of the Sierra Costa Mountains in California. *J Geol* 8:42–57
- Hoekzema K, Sidlauskas BL (2014) Molecular phylogenetics and microsatellite analysis reveal cryptic species of speckled dace (Cyprinidae: *Rhinichthys osculus*) in Oregon's Great Basin. *Mol Phylogenet Evol* 77:238–250. <https://doi.org/10.1016/j.ympev.2014.04.027>
- Hohler DB (1981) A dwarfed population of *Catostomus rimiculus* (Catostomidae: Pisces) in Jenny Creek, Jackson, County, Oregon. MS Thesis, Oregon State University
- Houston DD, Evans RP, Shiozawa DK (2012) Evaluating the genetic status of a Great Basin endemic minnow: the relict dace (*Relictus solitarius*). *Conserv Genet* 13:727–742. <https://doi.org/10.1007/s10592-012-0321-6>
- Jenkins M (2003) Prospects for biodiversity. *Science* 302:1175–1177. <https://doi.org/10.1126/science.1088666>
- Jensen JL, Bohonak AJ, Kelley ST (2005) Isolation by distance, web service. *BioMed Central Genet* 6:13. <https://doi.org/10.1186/1471-2156-6-13>
- Jombart T (2008) Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BioMed Central Genetics* 11:94. <https://doi.org/10.1186/1471-2156-11-94>
- Kalinowski ST (2005) hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes* 5:187–189. <https://doi.org/10.1111/j.1471-8286.2004.00845.x>
- Kinziger AP, Nakamoto RJ, Anderson EC, Harvey BC (2011) Small founding number and low genetic diversity in an introduced species exhibiting limited invasion success (speckled dace, *Rhinichthys osculus*). *Ecol Evol* 1:73–84. <https://doi.org/10.1002/ece3.8>
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452. <https://doi.org/10.1093/bioinformatics/btp187>
- Maddison WP, Knowles LL (2006) Inferring Phylogeny despite incomplete lineage sorting. *Syst Biol* 55:21–30. <https://doi.org/10.1080/10635150500354928>
- McNeely JA, Miller KR, Reid WV, Mittermeier RA, Werner TB (1990) Conserving the world's biodiversity. World Conservation Union (IUCN) & World Wide Fund for Nature (WWF), Gland, Suiza & World Resources Institute (WRI), Conservation International (CI) & World Bank. EUA, Washington, DC
- McPhail JD, Taylor EB (2009) Phylogeography of the longnose dace (*Rhinichthys cataractae*) species group in northwestern North America: the origin and evolution of the Umpqua and Millicoma dace. *Can J Zool* 87:491–497. <https://doi.org/10.1139/Z09-036>
- Meirmans PG, Van Tienderen PH (2004) genotype and genotype: two programs for the analysis of genetic diversity of asexual organisms. *Mol Ecol Notes* 4:792–794. <https://doi.org/10.1111/j.1471-8286.2004.00770.x>
- Minckley WL, Hendrickson DA, Bond CE (1986) Geography of western North American freshwater fishes: description and relationships to intracontinental tectonism. In: Hocutt CH, Wiley EO (eds) The zoogeography of North American freshwater fishes. Wiley, New York, pp 519–613
- Moyle PB (2002) Inland fishes of California. Univ of California Press
- Moyle PB, Katz JVE, Quiñones RM (2011) Rapid decline of California's native inland fishes: a status assessment. *Biol Conserv* 144:2414–2423. <https://doi.org/10.1016/j.biocon.2011.06.002>
- Moyle PB, Quiñones RM, Katz JV, Weaver J (2015) Fish species of special concern in California. California Department of Fish and Wildlife, Sacramento
- Murphy NP, King RA, Delean S (2015) Species, ESUs or populations? Delimiting and describing morphologically cryptic diversity in Australian desert spring amphipods. *Invertebrate Syst* 29:457–467. <https://doi.org/10.1071/IS14036>
- Myers N, Mittermeier RA, Mittermeier CG, Da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature* 403:853–858. <https://doi.org/10.1038/35002501>

- Narum SR (2006) Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conserv Genet* 7:783–787. <https://doi.org/10.1007/s10592-005-9056-y>
- Niemiller ML, Graening GO, Fenolio DB et al (2013) Doomed before they are described? The need for conservation assessments of cryptic species complexes using an amblyopsid cavefish (Amblyopsidae: Typhlichthys) as a case study. *Biodivers Conserv* 22:1799–1820. <https://doi.org/10.1007/s10531-013-0514-4>
- Noss RF (1990) Indicators for monitoring biodiversity: a hierarchical approach. *Conserv Biol* 4:355–364
- Oakey DD, Douglas ME, Douglas MR (2004) Small fish in a large landscape: diversification of *Rhinichthys osculus* (Cyprinidae) in Western North America. *Copeia*. <https://doi.org/10.1643/CG-02-264R1>
- Pearsons TN, Li HW, Lamberti GA (1992) Influence of habitat complexity on resistance to flooding and resilience of stream fish assemblages. *Trans Am Fish Soc* 121:427–436
- Peden AE, Hughes GW (1981) Life history notes relevant to the Canadian status of the speckled dace (*Rhinichthys osculus*). *Syesis* 14:21–31
- Pfenninger M, Schwenk K (2007) Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BioMed Central Evol Biol* 7:121. <https://doi.org/10.1186/1471-2148-7-121>
- Pfrender ME, Hicks J, Lynch M (2004) Biogeographic patterns and current distribution of molecular-genetic variation among populations of speckled dace, *Rhinichthys osculus* (Girard). *Mol Phylogenet Evol* 30:490–502. [https://doi.org/10.1016/S1055-7903\(03\)00242-2](https://doi.org/10.1016/S1055-7903(03)00242-2)
- Piggott MP, Chao NL, Beheregaray LB (2011) Three fishes in one: cryptic species in an Amazonian floodplain forest specialist. *Biol J Linn Soc* 102:391–403
- Posada D (2008) jModelTest: phylogenetic model averaging. *Mol Biol Evol* 25:1253–1256. <https://doi.org/10.1093/molbev/msn083>
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Raymond M, Rousset F (1995) GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *J Hered* 86:248–249
- Ricciardi A, Rasmussen JB (1999) Extinction rates of north american freshwater fauna. *Conserv Biol* 13:1220–1222. <https://doi.org/10.1046/j.1523-1739.1999.98380.x>
- Rice WR (1989) Analyzing tables of statistical tests. *Evol Int J Org Evol* 43:223–225
- Rosenberg NA (2004) DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes* 4:137–138. <https://doi.org/10.1046/j.1471-8286.2003.00566.x>
- Sharp RP (1960) Pleistocene glaciation in the Trinity Alps of northern California. *Am J Sci* 258:305–340. <https://doi.org/10.2475/ajs.258.5.305>
- Smith GR (1973) Analysis of several hybrid Cyprinid fishes from Western North America. *Copeia* 1973:395–410. <https://doi.org/10.2307/1443102>
- Smith GR, Dowling TE (2008) Correlating hydrographic events and divergence times of speckled dace (*Rhinichthys*, Teleostei, Cyprinidae) in the Colorado River drainage. In: Reheis M, Hershler R, Miller D (eds) Late cenozoic drainage history of the southwestern Great Basin and lower Colorado River region: Geologic and biotic perspectives. Special paper 439. Geological Society of America, Boulder, pp 301–317
- Tamura K, Stecher G, Peterson D, Flipski A, Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729. <https://doi.org/10.1093/molbev/mst197>
- Turner TF, Dowling TE, Broughton RE, Gold JR (2004) Variable microsatellite markers amplify across divergent lineages of cyprinid fishes (subfamily Leuciscinae). *Conserv Genet* 5:279–281. <https://doi.org/10.1023/B:COGE.0000029998.11426.ab>
- Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes* 4:535–538
- Walsh PS, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 10:506–513
- Waters JM, Rowe DL, Burrige CP, Wallis GP (2010) Gene trees versus species trees: reassessing life-history evolution in a freshwater fish radiation. *Syst Biol* 59:504–517. <https://doi.org/10.1093/sysbio/syq031>