

You Can't Unscramble an Egg: Population Genetic Structure of *Oncorhynchus mykiss* in the California Central Valley Inferred from Combined Microsatellite and Single Nucleotide Polymorphism Data

Devon E. Pearse*¹ and John Carlos Garza¹

Volume 13, Issue 4 | Article 3

doi: <http://dx.doi.org/10.15447/sfew.2015v13iss4art3>

* Corresponding author: devon.pearse@noaa.gov

¹ Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, Santa Cruz, CA 95060 USA and University of California, Santa Cruz, CA 95060 USA

KEY WORDS

Evolution, genetics, adaptation, steelhead, trout, Central Valley.

ABSTRACT

Steelhead/rainbow trout (*Oncorhynchus mykiss*) are found in all of the major tributaries of the Sacramento and San Joaquin rivers, which flow through California's Central Valley and enter the ocean through San Francisco Bay and the Golden Gate. This river system is heavily affected by water development, agriculture, and invasive species, and salmon and trout hatchery propagation has been occurring for over 100 years. We collected genotype data for 18 highly variable microsatellite loci and 95 single nucleotide polymorphisms (SNPs) from more than 1,900 fish from Central Valley drainages to analyze genetic diversity, population structure, differentiation between populations above and below dams, and the relationship of Central Valley *O. mykiss* populations to coastal California steelhead. In

addition, we evaluate introgression by both hatchery rainbow trout strains, which have primarily native Central Valley ancestry, and imported coastal steelhead stocks. In contrast to patterns typical of coastal steelhead, Central Valley *O. mykiss* above and below dams within the same tributary were not found to be each others' closest relatives, and we found no relationship between genetic and geographic distance among below-barrier populations. While introgression by hatchery rainbow trout strains does not appear to be widespread among above-barrier populations, steelhead in the American River and some neighboring tributaries have been introgressed by coastal steelhead. Together, these results demonstrate that the ancestral population genetic structure that existed among Central Valley tributaries has been significantly altered in contemporary populations. Future conservation, restoration, and mitigation efforts should take this into account when working to meet recovery planning goals.

INTRODUCTION

The Central Valley of California supports both Chinook salmon (*Oncorhynchus tshawytscha*) and fish from the species *O. mykiss*, commonly known

as steelhead (anadromous life history) or rainbow trout (resident life history). Tributary rivers from the west slope of the Sierra Nevada mountain range and east slopes of the coastal mountain ranges feed into the north-flowing San Joaquin and the south-flowing Sacramento rivers, which converge in the San Francisco Bay/Delta region before finally exiting to the Pacific ocean. The watershed has been severely affected by the construction of many dams, which block movement by anadromous fish and effectively divide nearly every major tributary into separate above-barrier and below-barrier reaches. In addition, much of the flow in the system is diverted for agricultural and domestic uses through an extensive system of levees and aqueducts. Together, these effects have severely modified and reduced the habitat available to anadromous fishes (Lindley et al. 2006).

Populations of steelhead in California are divided into six Distinct Population Segments (DPSs) for management purposes; five on the coast and one in the Central Valley (Busby et al. 1996). Importantly, these DPSs specifically include only anadromous life-history fish that spawn below impassable barriers to migration; *O. mykiss* isolated above natural or artificial barriers to fish passage are excluded from the DPS and, consequently, from protection under the U.S. Endangered Species Act (ESA; Federal Register 2006). The California Central Valley Steelhead DPS also includes fish produced by two of the four artificial propagation programs in the Central Valley—the Feather River Fish and Coleman National Fish hatcheries—but not those spawned at the Nimbus or Mokelumne River hatcheries. This DPS was listed as “Threatened” under the ESA in 1998 and this status was reaffirmed in 2006 (Federal Register 2006).

Hatchery rainbow trout have been heavily stocked in the reservoirs above nearly all of the Central Valley dams for more than 100 years (Busack and Gall 1980; California HSRG 2012). These captive hatchery trout broodstock strains were domesticated from diverse geographic and phylogenetic sources, but many originated from fish collected from streams that drain into the Central Valley (Needham and Behnke 1962; Leitritz 1970). Similarly, steelhead and other anadromous salmonids have been propagated at several hatcheries in the Central Valley since the late

1800s, and four Central Valley hatcheries (Coleman, Feather, Nimbus, and Mokelumne), currently release approximately 1.5 million yearlings annually (Brown 2005; California HSRG 2012). For both steelhead and hatchery trout strains, it has been common practice to move eggs among hatcheries within the Central Valley and to import eggs from outside sources (Leitritz 1970; California HSRG 2012). Nimbus Hatchery on the American River has been a substantial producer of steelhead in the Central Valley since 1955 (Leitritz 1970) and, for many years, imported eggs from coastal steelhead sources, primarily the Eel and Mad rivers (California HSRG 2012). However, the extent to which such interbasin transfers have influenced population structure of *O. mykiss* in the Central Valley has not been carefully evaluated.

Numerous genetic analyses of salmonid population structure in California have relied on microsatellite markers, because such multi-locus data can identify population genetic structure at both larger scales (Aguilar and Garza 2006; Clemento et al. 2009; Garza et al. 2014) and at relatively fine ones (Deiner et al. 2007; Pearse et al. 2007, 2009; Kinziger et al. 2013), including within the Central Valley (Banks et al. 2000; Nielsen et al. 2005). Recently, another class of genetic markers, single nucleotide polymorphisms (SNPs), has been used increasingly in population genetics and has proven useful in assessments of population structure (Morin et al. 2004), introgressive hybridization (Stephens et al. 2009; Finger et al. 2011), and pedigree reconstruction (Abadía-Cardoso et al. 2013). Though microsatellites and SNPs each have advantages and disadvantages in terms of cost, genotyping errors, polymorphism, etc., when a large number of both types of loci is available, this combination provides the most statistical power for understanding population genetic relationships (Narum et al. 2008).

Here we attempt to “unscramble” the population genetic structure of Central Valley *O. mykiss* using a combination of more than 100 microsatellite and SNP loci on a comprehensive set of Central Valley trout and steelhead populations. We compare these data with genotypes from a representative set of hatchery trout strains and coastal California steelhead populations (Aguilar and Garza 2006; Pearse et al.

2007, 2009; Clemento et al. 2009; Garza et al. 2014). Analysis of this combined dataset provides insight into the the historical relationships of Central Valley *O. mykiss*, as well as the relationships of Central Valley populations with those from other parts of California.

METHODS

Sampling

Samples were taken from populations of *O. mykiss* at one or more locations in 15 tributary sub-basins of the Sacramento and San Joaquin rivers that drain the Central Valley (Figure 1; Table 1), including locations both above and below barriers to anadromy in most tributaries. Most fish were captured using either electrofishing or hook-and-line capture techniques. Small pieces of caudal fin tissue were then excised and preserved through desiccation on blotter paper. Fish sampled in multiple years in the same location were combined for analysis, after verifying that they were taken from the same underlying population. These groups of fish are all referred to as populations for convenience and without additional assumptions about the biological details underlying this designation.

Genetic Data Collection

Nucleic acid extraction and microsatellite and SNP genotyping followed Arciniega et al. (2016). Genotypic data from 18 microsatellite loci were collected for all samples. This set of loci has been used in numerous previous studies of *O. mykiss* in California (Aguilar and Garza 2006; Deiner et al. 2007; Pearse et al. 2007, 2009, 2011a; Garza et al. 2014). All samples were also genotyped with the panel of 96 SNP loci used by Abadía-Cardoso et al. (2013). The 96 SNPs include 95 loci from Aguilar and Garza (2008), Campbell et al. (2009), and Abadía-Cardoso et al. (2011), as well as an assay that includes a Y-chromosome marker developed by Brunelli et al. (2008) that identifies gender. All 96 loci were genotyped using 5' nuclease TaqMan assays (Applied Biosystems) on 96.96 Dynamic Genotyping Arrays in the EP1 Genotyping System (Fluidigm

Corporation). Two negative controls were included in each array and genotypes were called using Fluidigm SNP Genotyping Analysis Software v3.1.1.

Data Analysis

We combined the microsatellite and SNP data collected from the Central Valley *O. mykiss* populations with previously collected data from coastal California steelhead populations and hatchery trout strains commonly stocked in California. In analyzing these data, we first removed from most analyses three SNP loci that have been shown to be influenced by selection on life-history patterns in *O. mykiss* (Pearse et al. 2014). Two of these loci in particular, *SH121006-131* and *SH114448-87*, are in strong linkage disequilibrium (LD) with a genomic region on chromosome Omy5 that was recently found to be associated with resident and anadromous life-history in coastal California steelhead populations (Pearse et al. 2014). These two loci were analyzed separately to evaluate patterns of LD between them in Central Valley populations using the R package genetics (Warnes and Leisch 2005). Finally, we removed three microsatellite loci (*OtsG401*, *Omy27*, and *Ots1b*) and two SNP loci (*SH127645-308* and *SH128996-481*) for which at least one of the population samples was not genotyped. Together, these removals left a total of 105 loci (15 microsatellites and 90 SNP loci), and we conducted all further population genetic analyses on this combined dataset. The gender identification locus was also excluded from the population genetic analyses.

We calculated expected heterozygosity (Nei 1987), observed heterozygosity, and number of alleles for each sample population, and estimated allelic richness (Ar) with the rarefaction method in the program HP-Rare (Kalinowski 2005) based on a sample of 25 gene copies. We quantified pairwise differentiation between all populations with *F_{ST}*, using Weir and Cockerham's (1984) estimator, and assessed significance by the permutation algorithm in the genetix software package (Belkhir et al. 2004) with 100 replicates. We used a Mantel test implemented in the program *ISOLDE* of the GenePop software package (Raymond and Rousset 1995) to evaluate the correlation between genetic and geographic distance

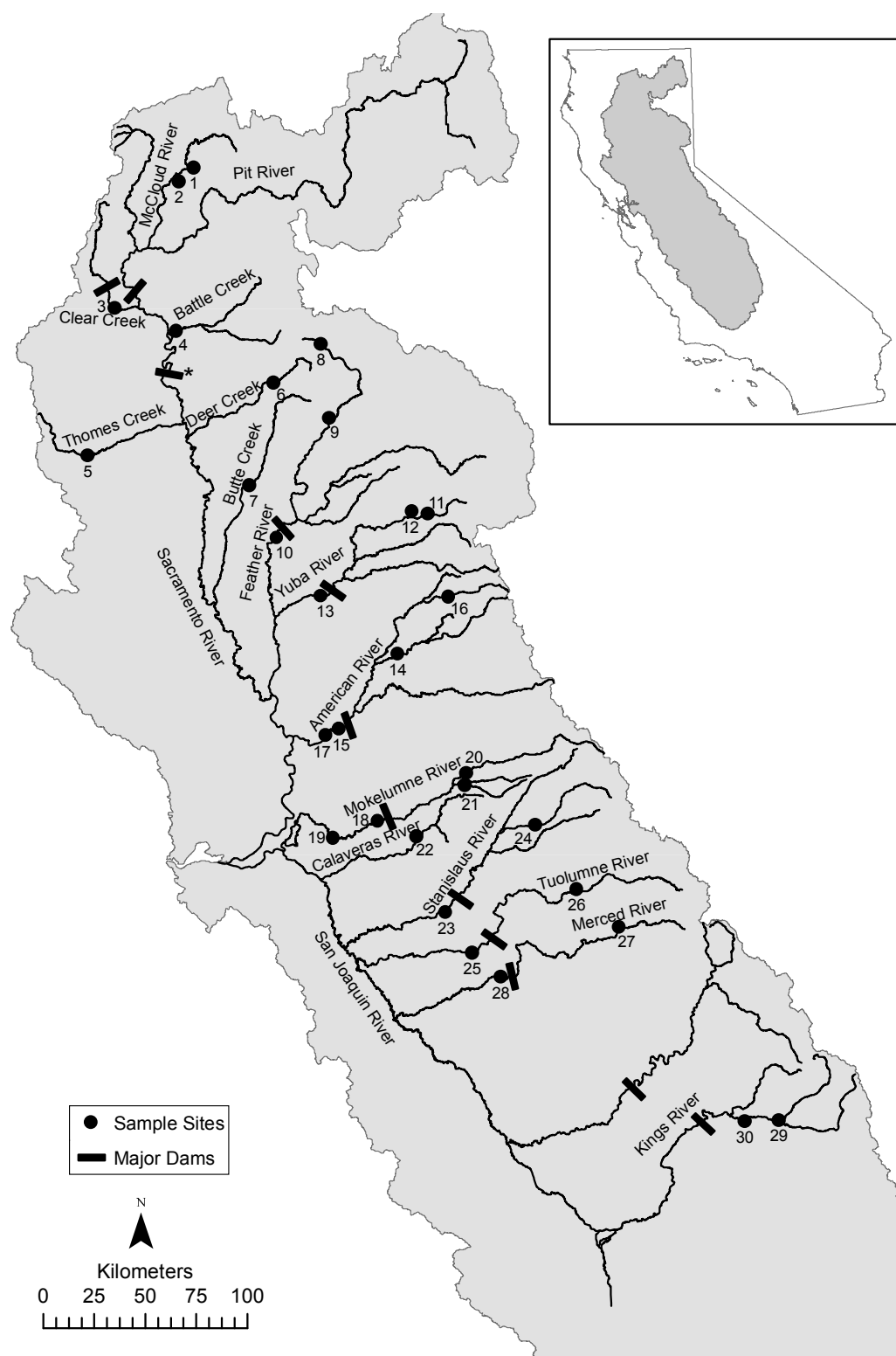


Figure 1 Map showing Central Valley of California, including major barrier dams and rivers tributary to the Sacramento-San Joaquin system. Sample sites are numbered as in Table 1.

Table 1 Population samples used in the present study, organized by region (coastal versus Central Valley), and listed north to south. Within each basin, population samples are grouped with respect to location above and below major dams, or at hatcheries. For each population, number of samples (N), expected and observed heterozygosities (He and Ho), microsatellite and SNP allelic richness (Ar), proportion genetically self-assigned, and linkage disequilibrium (r^2) between two SNP loci located on chromosome Omy5.

Population	Area	N	He	Ho	Ar (micros)	Ar (SNPs)	% Self	r^2 (Omy5)
North Coast								
Smith R.	Coastal	32	0.38	0.37	7.99	1.89	93.8	0.99
Klamath R. (Blue Ck.)	Coastal	32	0.39	0.37	8.64	1.93	96.9	0.93
Klamath R. (Hunter Ck.)	Coastal	28	0.39	0.39	6.68	1.92	85.7	0.79
Redwood Ck. (Lost Man Ck.)	Coastal	31	0.42	0.39	8.27	1.94	87.1	1.00
Mad R.	Coastal	31	0.39	0.41	6.77	1.94	74.2	1.00
Eel R. (Hollow Tree Ck.)	Coastal	28	0.39	0.39	7.04	1.91	71.4	1.00
Eel R. (Lawrence Ck.)	Coastal	30	0.41	0.40	7.50	1.94	90.0	0.74
Mattole R.	Coastal	31	0.39	0.40	6.36	1.92	93.5	1.00
Noyo R.	Coastal	31	0.41	0.42	7.62	1.96	93.5	0.82
Gualala R.	Coastal	29	0.43	0.45	6.65	1.98	86.2	1.00
Redwood Ck. (Marin Co.)	Coastal	30	0.44	0.44	7.28	1.98	83.3	0.86
Miller Ck. (Marin Co.)	Coastal	31	0.39	0.41	5.66	1.94	100.0	1.00
Central Valley								
1. McCloud R. (Butcherknife Ck.)	Above	21	0.21	0.21	3.75	1.61	100.0	—
2. McCloud R. (Claiborne Ck.)	Above	33	0.37	0.37	6.00	1.88	97.0	—
3. Clear Ck.	Below	94	0.39	0.37	8.00	1.96	93.6	0.12
4. Battle Ck.	Below	94	0.41	0.40	8.43	1.97	71.3	0.25
5. Thomas Ck.	Below	51	0.37	0.36	6.25	1.92	100.0	1.00
6. Deer Ck.	Below	45	0.42	0.43	8.64	1.98	71.1	0.40
7. Butte Ck.	Below	47	0.40	0.39	7.99	1.95	87.2	0.91
8. Feather R. (Above Lake Almanor)	Above	16	0.29	0.27	6.23	1.86	100.0	1.00
9. Feather R. (Chips Creek)	Above	31	0.37	0.36	7.49	1.91	90.3	0.49
10. Feather River Hatchery	Hatchery	30	0.41	0.40	7.01	1.96	33.3	0.17
11. Yuba R. (Upper)	Above	26	0.43	0.43	6.85	1.98	84.6	1.00
12. Yuba R. (Pauley Ck.)	Above	25	0.35	0.33	5.71	1.85	92.0	—
13. Yuba R.	Below	90	0.45	0.44	7.80	1.99	87.8	0.68
14. American R. (MF)	Above	58	0.42	0.40	7.58	1.97	91.4	0.93
15. American R.	Below	19	0.43	0.44	7.18	1.98	42.1	0.25
16. American R. (NF)	Above	49	0.38	0.38	6.29	1.92	93.9	0.64
17. Nimbus Hatchery	Hatchery	98	0.43	0.42	7.53	1.97	86.2	0.39
18. Mokelumne Hatchery	Hatchery	162	0.41	0.40	7.31	1.97	77.8	0.09
19. Mokelumne R.	Below	63	0.43	0.42	7.61	1.98	74.6	0.36
20. Mokelumne R. (North Fork)	Above	51	0.38	0.38	7.63	1.94	80.4	0.84
21. Mokelumne R. (South Fork)	Above	49	0.36	0.36	6.84	1.93	91.8	0.65
22. Calaveras R.	Below	47	0.41	0.41	6.91	1.96	95.7	0.13
23. Stanislaus R.	Below	80	0.44	0.44	7.66	1.99	91.3	0.26
24. Stanislaus R. (Upper)	Above	52	0.37	0.37	7.02	1.94	95.7	0.75
25. Tuolumne R.	Below	112	0.43	0.43	6.62	1.97	95.5	0.64
26. Tuolumne R. (Upper)	Above	47	0.39	0.36	7.28	1.93	91.5	0.87
27. Merced R. (Upper)	Above	35	0.36	0.32	6.04	1.92	90.5	0.41
28. Merced R.	Below	83	0.34	0.32	5.71	1.90	51.8	0.22
29. Kings R. (Deer Cove Ck.)	Above	33	0.38	0.37	4.70	1.95	90.9	0.32
30. Kings R. (Mill Flat Ck.)	Above	26	0.41	0.42	6.94	1.98	92.3	0.82
South Coast								
San Franciscoquito R.	Coastal	24	0.39	0.41	5.34	1.91	100.0	1.00
San Lorenzo R.	Coastal	32	0.42	0.43	5.80	1.97	71.9	1.00
Carmel R.	Coastal	32	0.42	0.42	6.96	1.97	96.9	0.68
Big Sur R.	Coastal	31	0.43	0.43	6.97	1.97	77.4	1.00
Hatchery trout strains								
Kamloops	Trout	47	0.29	0.28	5.51	1.75	100.0	—
Mt. Shasta	Trout	92	0.35	0.35	4.26	1.84	98.9	0.01
Eagle	Trout	47	0.30	0.29	4.35	1.83	87.2	1.00
Coleman	Trout	47	0.37	0.37	5.04	1.91	100.0	0.19
Moccasin	Trout	47	0.30	0.30	4.72	1.78	100.0	0.11
All CV		1,667					84.7	
All samples		2,430					86.9	

for the naturally spawning populations below barriers, using river distances separating the confluences of each major tributary along the mainstem of the Sacramento–San Joaquin River system.

We used two individual-based assignment methods to evaluate both recent gene flow among populations and to identify hatchery rainbow trout individuals among the naturally spawning populations. The first analysis, implemented in the model-based clustering program *structure* (version 2.2; Pritchard et al. 2000), was used to fractionally assign the genome of individual fish to a hypothesized number of genetic clusters, K , in the dataset and to identify population associations. This analysis did not use information about *a priori* population designations, so it truly assigns the ancestry of each individual fish without regard to its origin. We evaluated the data using a range of values of $K=2$ –14 to qualitatively document consistent patterns of population association. The second assignment analysis, implemented in the program *gsi_sim* (Anderson et al. 2008), uses the population genotype data as references to assign each individual fish to its most likely population of origin based on the method of Rannala and Mountain (1997). This approach evaluates the likelihood of assignment of each individual to every population, providing an evaluation of the composition of each population sample.

We constructed phylogeographic trees based on matrices of Cavalli-Sforza and Edwards' (1967) chord distance using the software package PHYLIP (v. 3.69c; Felsenstein 2005). This genetic distance was chosen because of its accuracy and ability to reliably recover the correct topology for phylogeographic trees (Takezaki and Nei 1996; Felsenstein 2003). We used the neighbor-joining algorithm (Saitou and Nei 1987) to determine tree topology, and derived a consensus tree from 1,000 bootstrap samples of the distance matrix with the CONSENSE program of PHYLIP. Finally, we conducted a correspondence analysis (CA) on the full dataset to qualitatively evaluate population relationships in the absence of a constrained tree structure. This analysis was conducted using the R-based software package *adeget* 1.3-4 (Jombart 2008; Jombart and Ahmed 2011).

RESULTS

Individual-Based Analysis

The final dataset contained genotypes of 2,430 individuals from 51 sample groups, including 1,667 fish from Central Valley populations. Model-based assignments from the program *structure* over the range of K -values employed clearly identified hatchery rainbow trout sampled among the naturally spawned fish (Figure 2). This analysis was used to identify 14 hatchery-origin rainbow trout in the Upper Merced population sample, six in the Upper Stanislaus, and 11 sampled at Nimbus Hatchery. The large number of hatchery trout identified in the Upper Merced River (14 of 35, 40%) were all sampled on the same day, separately from the rest of the fish in that population sample, and likely represent a distinct group of planted hatchery trout. Hatchery rainbow trout identified with *structure* were removed from the dataset in all subsequent analyses, with the exception of fish in the Lower Merced River sample, which had a strong and uniform hatchery influence, so no individuals could be singled out for removal.

Individual assignment tests provided high accuracy of self-assignment to Central Valley *O. mykiss* populations. The overall accuracy of assignment to population of origin was 84.7% (Table 1). Assignment accuracy for individual populations ranged from 100% for the McCloud R.–Butcherknife Ck., Thomes Creek, and the Feather River-above-Lake-Almanor samples to 33% for the Feather River Hatchery stock, in which many fish assigned to the Mokelumne Hatchery, and vice versa. Similarly, a substantial number of individuals cross-assigned between the American River and Nimbus Fish Hatchery samples, reflecting the strong similarities between these groups of fish.

Population Genetic Diversity

Allelic richness within populations was strongly correlated for microsatellite and SNP loci ($r^2=0.453$, $p<0.001$; Table 1). For the microsatellite loci, allelic richness ranged from a low of 3.75 (McCloud R.–Butcherknife Ck.) to a high of 8.64 (Deer Ck.). Since these SNPs have a maximum of two alleles, their allelic richness ranged from 1.6 (McCloud R.–

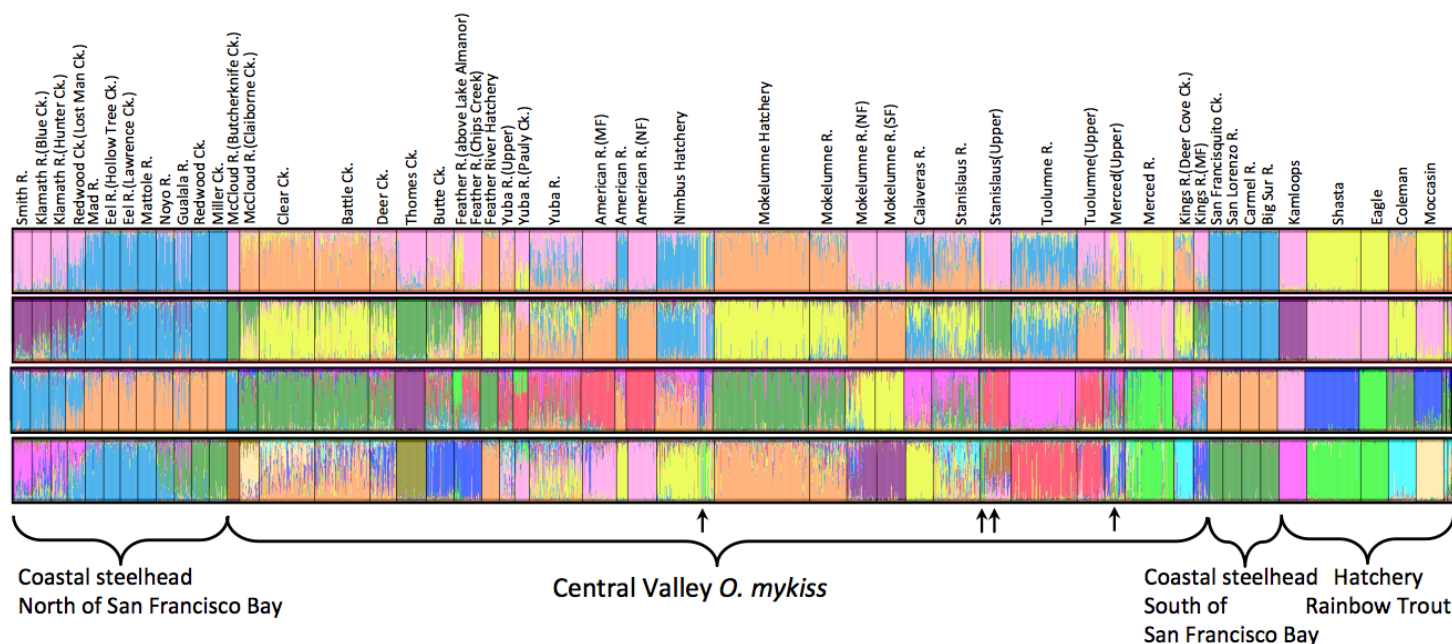


Figure 2 Graphical results of model-based clustering method implemented in structure (Pritchard et al. 2000) for a hypothesized number of genetic clusters, K, of 4, 6, 10 and 14. This analysis fractionally assigns each individual to the set of genetic clusters without prior consideration of geographic or population information, providing an unbiased assessment of individual ancestry. Arrows show hatchery trout identified and removed from other analyses.

Butcherknife Ck.) to >1.9 (many populations; Table 1). For both marker types, the range of allelic richness values observed in Central Valley populations encompassed the full range seen in coastal *O. mykiss* populations (Table 1). Observed heterozygosity across all loci ranged from a low of 0.212 in the McCloud R.–Butcherknife Creek population to 0.443 in the lower American, Stanislaus, and Yuba rivers (Table 1). Both microsatellite allelic richness and observed heterozygosity were significantly lower in the populations above dams than those in sites below dams (AR: 6.42 vs. 7.40, t -test, $p < 0.05$. Ho: 0.355 vs. 0.401; t -test, $p < 0.05$), while the five hatchery trout strains had the lowest average values for any of the groups for both genetic diversity measures (Table 1). Finally, unlike the strong LD between the two SNP loci on chromosome Omy5 that was observed in coastal *O. mykiss* populations (mean $r^2 = 0.92$, range = 0.68 to 1.00), the strength of LD between these two SNP loci varied widely among Central Valley populations (mean $r^2 = 0.54$, range = 0.09 to 1.00; Table 1). Nonetheless, as in

coastal populations, the mean frequency of alleles associated with anadromy at these adaptive loci was significantly higher in Central Valley below-barrier populations than above-barrier ones (0.48 vs. 0.17; t -test, $p < 0.01$), consistent with the influence of dams on life-history variation (Pearse et al. 2014).

Population Structure

We examined pairwise values of F_{ST} , the standardized variance in allele frequencies between populations, for patterns of population structure. All pairwise F_{ST} values were significantly greater than zero based on permutation tests, with the highest values found between above-barrier populations (0.34, McCloud R., Butcherknife Ck. and Yuba River–Upper) and the lowest values involving below-barrier hatchery populations (0.005, Feather River Hatchery and Mokelumne Hatchery; 0.01, Nimbus Hatchery and American River) and below-barrier natural populations (0.015, Battle and Deer creeks). Notably, the lower

statistically significant, relationships were consistent with the trees shown in [Figures 3](#) and [4](#). In general, the phylogeographic trees did not cluster populations by basin of origin, with little or no statistical support for most internal branching relationships. We found strong bootstrap support primarily for nodes joining pairs of population samples above the same barrier dam. For example, the relationships between the two upper American River populations—American-NF and American-MF, and the two upper Mokelumne River populations, Mokelumne-NF and Mokelumne-SF—were both strongly supported in all trees ([Figures 3](#) and [4](#)). There was also a well-supported association between the Upper Yuba (Pauley Creek), Upper Feather River (both samples), Eagle Lake hatchery strain, and Lower Merced River samples, which consistently clustered, even when Eagle Lake was excluded from the analysis. Among the below-barrier populations, the American River—Lower and Nimbus

Phylogeographic trees were created for Central Valley populations only (Figure 3) and also with coastal California steelhead included (Figure 4). We also constructed trees using the microsatellite and SNP data separately, and with the hatchery rainbow trout strains included and excluded (data not shown). Regardless of which populations were included, there were only minor differences in the relationships inferred in the different trees, and all the major,

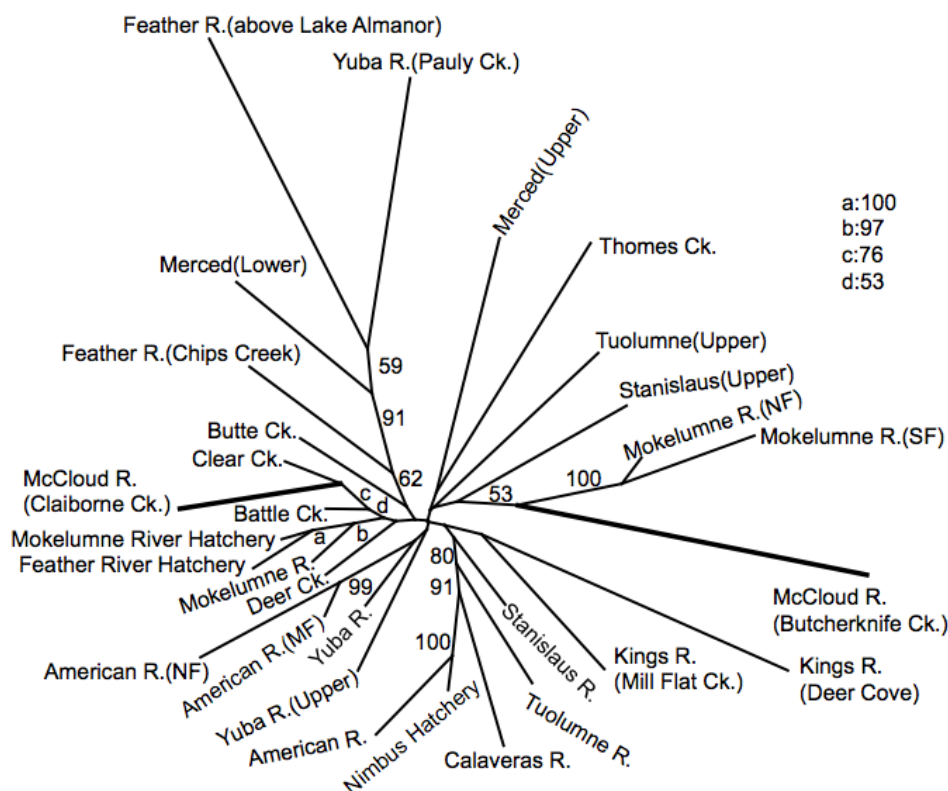


Figure 3 Neighbor-joining network showing only Central Valley *O. mykiss* populations, constructed with chord distances and original sub-basin groupings. Thick lines for McCloud populations shown at half length for display purposes. Bootstrap consensus values from 1000 bootstrap replicates shown. Only bootstrap values above 50% are reported.

Hatchery samples were closely associated with strong bootstrap support in all trees, as were the Mokelumne River, Mokelumne River Hatchery, and Feather River Hatchery samples (Figures 3 and 4).

The phylogeographic analysis that included coastal California steelhead populations revealed that, in general, Central Valley *O. mykiss* populations, both above and below dams, are more closely related to each other than to coastal populations outside of the Central Valley. Similarly, all of the hatchery strains cluster with the Central Valley populations in those analyses, as expected, given that most strains of hatchery rainbow trout used in California were domesticated from Sacramento River tributary

populations (Busack and Gall 1980). The reduced LD between the two Omy5 loci in the hatchery trout strains is also consistent with their Sacramento River basin origins (Table 1). In addition, the American River–Lower and Nimbus Hatchery samples fall in a position intermediate between the coastal steelhead populations and the rest of the Central Valley, consistent with the founding of the current Nimbus Hatchery stock with eggs imported from coastal populations, primarily the Eel and Mad rivers. Finally, correspondence analysis displayed the same general relationships among Central Valley and coastal *O. mykiss* populations as the phylogenetic analysis, without the constraints of a tree (Figure 5).

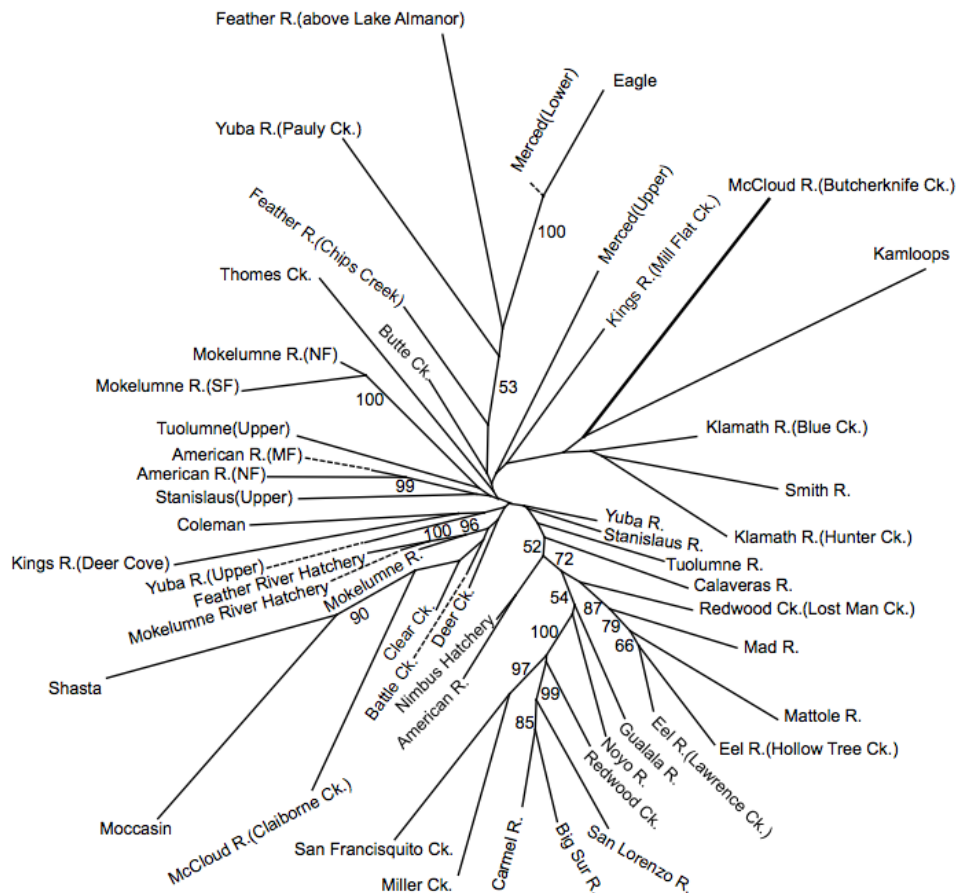


Figure 4 Neighbor-joining network of coastal and Central Valley *O. mykiss* populations constructed with chord distances and original sub-basin groupings. The thick line for McCloud R. (Butcherknife Ck.) is shown at half scale, and dotted lines connect names to branch tips for display purposes. Bootstrap consensus values from 1000 bootstrap replicates shown. Only bootstrap values above 50% are reported.



Figure 5 Correspondence analysis, showing population relationships on the first two axes of differentiation. Coastal steelhead populations that are not contained in the oval shown are labeled individually.

DISCUSSION

In contrast with the patterns typically found in natural populations, genetic analysis of Central Valley *O. mykiss* populations with more than 100 markers found a general lack of geographically associated population structure. This likely reflects more than a century of habitat modification and stocking/hatchery practices that together have altered the historical genetic relationships among *O. mykiss* populations in at least three ways. First, unlike the close relationships typically found between coastal *O. mykiss* populations above and below barriers within the same watershed (Clemento et al. 2009; Pearse et al. 2009), Central Valley populations separated by dams are usually not each other's closest relatives. Second, the relationships among below-barrier Central Valley populations do not fit a pattern of isolation-by-distance, as has been found among *O. mykiss* and other salmonid populations both within and among watersheds (Primmer et al. 2006; Palstra et al. 2007; Pearse et al. 2007; Pearse et al. 2011b; Garza et al. 2014), as well as in a recent study of Central Valley giant gartersnakes (*Thamnophis gigas*) inhabiting the same geographic area (Wood et al. 2015). Finally, some below-barrier Central Valley *O. mykiss* populations, particularly in the lower American River, are clearly derived primarily from populations from the northern California steelhead DPS, presumably through past importation of eggs from the Eel and Mad rivers. Like scrambling an egg, these genetic effects are largely irreversible, and future management must take them into account while recognizing that the historical relationships cannot be completely restored. However, such genetic effects are also not static, making efforts to use science-based recovery planning essential for the restoration of the adaptive potential of *O. mykiss* populations in the Central Valley (Meek et al. 2014).

Our results are largely concordant with previous genetic studies of Central Valley *O. mykiss* (e.g., Nielsen et al. 2005). However, the increased power of the combined microsatellite and SNP data used in the present study, as well as the inclusion of multiple stocks of hatchery rainbow trout and population samples above barriers to anadromy, offer increased resolution, especially given the complementary characteristics of these two types of marker (Narum et

al. 2008). Nonetheless, unlike the well-supported relationships and strong isolation by distance found among coastal populations, there was only weak statistical support for most phylogenetic relationships among Central Valley *O. mykiss* populations. Thus, the lack of strong population structure found in this study likely represents an accurate depiction of the current population genetic relationships among Central Valley *O. mykiss* populations, while also showing that the overall genetic distinction between coastal and Central Valley DPS *O. mykiss* remains. Moreover, the majority of the genetic diversity found among the Central Valley steelhead/rainbow trout populations studied here was found at the level of the individual sample sites, all of which were significantly differentiated, contributing to high rates of self-assignment for most populations (Table 1). Accurate population self-assignments are useful because they indicate that the underlying genetic data can be used as a reference baseline for genetic stock identification techniques to determine basin and tributary of origin for individual fish in management or forensic applications (e.g., Seeb et al. 2007).

As noted above, one salient result of the present study is that populations above and below barrier dams in the same basins are not closely related in most of the major tributaries. Instead many of the above-barrier populations appear to be more genetically similar to each other than to any of the below-barrier populations, a pattern also observed by Nielsen et al. (2005). However, that study did not evaluate relationships between Central Valley trout and hatchery rainbow trout, leaving uncertainty about the phylogenetic origin of the above-barrier populations (Lindley et al. 2006). In the present study, most above-barrier populations are clearly genetically distinct from the hatchery trout strains, supporting the hypothesis that hatchery rainbow trout stocked in the reservoirs and elsewhere above dams in the region have not replaced the native *O. mykiss* populations that residualized following dam construction. Thus, our results suggest that native *O. mykiss* dominate the existing populations represented above the dams, as has been documented in coastal California basins (Clemento et al. 2009). However, it should be noted that detecting the influence of hatchery strains

is complicated by the close relationship of most hatchery trout strains to their Central Valley origins, and substantial past introgression by hatchery trout into some or all above-barrier populations can not be completely ruled out.

In several sub-basins, we sampled and analyzed multiple above-barrier populations and the results were not all consistent. For example, in the Kings River, samples from Deer Cove and Mill Flat creeks both showed some similarity to hatchery trout strains, and the two populations were closely associated in some, but not all, analyses. On the other hand, pairs of samples from different tributaries above Folsom and Pardee dams, in the American and Mokelumne rivers, respectively, were closely related in all analyses. Pairwise F_{ST} values were very low, 0.03 and 0.04 between the middle and north forks of the American River and north and south forks of the Mokelumne River, respectively, and both pairs also cluster with high confidence in all phylogeographic analyses, indicating a common genetic ancestry and/or recent gene flow between them.

Artificial propagation of *O. mykiss* began in the Central Valley with the establishment of the Baird Station on the McCloud River in 1872. Since then, millions of juvenile fish have been released annually in Central Valley rivers, streams, lakes and reservoirs (Leitritz 1970). This massive propagation and stocking effort, much of it sparsely documented, significantly complicates efforts to disentangle historical population structure. Based on individual genotypic assignments, few hatchery trout were found amongst the population samples, with almost all identified hatchery trout sampled in three locations (Upper Stanislaus River, Upper Merced River, and Nimbus Hatchery). However, two populations showed significant associations with one or more hatchery trout strains. The population from Deer Cove Creek on the Kings River clustered with hatchery strains in some analyses, suggesting likely hatchery trout ancestry, even though no hatchery trout were identified individually. More strikingly, the sample from the Lower Merced River associated strongly with the Eagle Lake hatchery trout strain in both phylogenetic and correspondence analyses, as well as containing a significant number of individuals that assigned to the Eagle Lake strain.

Thus it appears that the fish sampled in the Lower Merced River are almost exclusively descended from this hatchery trout strain.

Introgression of hatchery rainbow trout into natural steelhead/rainbow trout populations and hatchery production is potentially detrimental, because of their reduced genetic variation, history of hatchery selection, and potential for a genetic predisposition against anadromy. Here, among the 31 sampled adults that entered Nimbus Fish Hatchery in 2005–2006, nine were identified as hatchery rainbow trout (Garza and Pearse 2008). These individuals were generally smaller than the steelhead, but there was significant overlap in the size distributions, suggesting that such fish might be mistaken for small steelhead and incorporated into the broodstock. However, a separate genetic analysis of steelhead broodstock from all four Central Valley steelhead hatcheries from 2011–2013 identified the presence of fewer than 10 hatchery rainbow trout among the more than 7,000 adult spawners assayed (four at Nimbus among > 1,100 spawners; Vendrami et al., unpublished data, see “Notes”). In addition, recent changes in California state hatchery protocols call for stocking only sterile, triploid, trout in waters where they may come into contact with naturally spawning steelhead (Fish and Game Code Chpt. 7.2, §1725–1730; see http://www.leginfo.ca.gov/.html/fgc_table_of_contents.html). Thus, it is unlikely that the incorporation of hatchery trout into broodstock of Central Valley steelhead hatchery programs is a widespread or ongoing problem.

Steelhead propagation and hatchery practices have directly affected the population structure of *O. mykiss* below barriers in the Central Valley in several ways. First, after the collapse of the steelhead program at Mokelumne Hatchery, the broodstock was replaced with fish from Feather River Hatchery, and these two populations are now extremely similar genetically, despite being geographically separated by >100 km. Similarly, Nimbus Hatchery on the American River has been a substantial producer of steelhead in the Central Valley since 1955 (Leitritz 1970) and, during the first several decades of operation, broodstock was imported periodically from coastal steelhead populations, including the Eel, Mad and Russian rivers (Lee and Chilton 2007). The effects of this out-of-basin

stocking are apparent in both individual and population analyses, in which the Nimbus and American-Lower populations are intermediate between the coastal steelhead populations and all other Central Valley populations. Notably, the closest relationship of the American River populations outside of the Central Valley is to fish from Northern California, in the group that includes the Eel and Mad rivers, rather than to more geographically proximate populations in San Francisco Bay (e.g., Los Trancos, Miller Creek, Redwood Creek; [Figures 3 and 4](#)). The clustering of other Central Valley below-barrier populations with Nimbus and American River samples, particularly those from the Calaveras and Tuolumne Rivers, indicates that introgression of natural populations by fish with coastal steelhead ancestry has occurred through straying/migration of Nimbus Hatchery steelhead. Conversely, individual assignment tests and the reduced LD between the two *Omy5* loci show that the *O. mykiss* now propagated at Nimbus Hatchery have diverged significantly from their coastal steelhead origins, likely through interbreeding with native Central Valley fish.

CONCLUSION

Our genetic results indicate small population sizes and reduced genetic diversity in above-barrier populations relative to below-barrier populations, consistent with the decreased connectivity and lost influx of new genes through migration after dam construction, factors that can contribute to population extirpation (Srikwan and Woodruff 2000). Facilitating fish migration across barriers is one way to mitigate such effects, and might also counteract adaptation of above-barrier populations in response to the strong selection against anadromy in these populations (Pearse et al. 2014). However, re-establishing connectivity of above-barrier populations trout with steelhead populations below dams should be carefully monitored because the consequences of such integration are not known, and could range from beneficial increases in genetic diversity and effective size, to negative changes in life history of the below-barrier populations, decreased fitness of hybrids, and adverse ecological interactions such as competition or direct predation.

ACKNOWLEDGEMENTS

We thank Vanessa Apkenas, Anthony Clemento, Cheryl Dean, Celeste Gallardo, Libby Gilbert-Horvath, Andres Martinez, and Edith Martinez for assistance in the laboratory, and Sara John and Alicia Abadía-Cardoso for help with the analysis and figures. The samples from most of the Central Valley steelhead populations were provided by George Edwards, Michael Lacy and Katie Perry (California Department of Fish and Wildlife). The samples from Battle Creek were provided by Andrew Matala and Bill Ardren (U.S. Fish and Wildlife Service), from Thomes Creek by Lee M. Morgan. Most of the hatchery fish were sampled by Chris Donohoe. George Edwards, Tim Heyne, Michael Lacy, Dennis Lee, and Armando Quiñones provided helpful discussion and background. This work was supported in part by California Department of Fish and Wildlife contract #P0485303.

REFERENCES

- Abadía-Cardoso A, Clemento AJ, Garza JC. 2011. Discovery and characterization of single nucleotide polymorphisms in steelhead/rainbow trout, *Oncorhynchus mykiss*. *Mol Ecol Resour* 11:31–49. doi: <http://dx.doi.org/10.1111/j.1755-0998.2010.02971.x>.
- Abadía-Cardoso A, Anderson EC, Pearse DE, Garza JC. 2013. Large-scale parentage analysis reveals reproductive patterns and heritability of spawn timing in a hatchery population of steelhead (*Oncorhynchus mykiss*). *Mol Ecol* 22:4733–4746. doi: <http://dx.doi.org/10.1111/mec.12426>.
- Aguilar A, Garza JC. 2006. A comparison of variability and population structure for major histocompatibility complex and microsatellite loci in California coastal steelhead (*Oncorhynchus mykiss* Walbaum). *Mol Ecol* 15:923–937. doi: <http://dx.doi.org/10.1111/j.1365-294X.2006.02843.x>.

- Aguilar A, Garza JC. 2008. Isolation of 15 single nucleotide polymorphisms from coastal steelhead, *Oncorhynchus mykiss* (Salmonidae). *Mol Ecol Res* 8:659–662. doi: <http://dx.doi.org/10.1111/j.1471-8286.2007.02038.x>.
- Anderson EC, Waples RS, Kalinowski ST. 2008. An improved method for predicting the accuracy of genetic stock identification. *Can J Fish Aquat Sci* 65:1475–1486. doi: <http://dx.doi.org/10.1139/F08-049>.
- Arciniega M, Clemento AJ, Miller MR, Peterson M, Garza JC, Pearse DE. 2016. Parallel evolution of the summer steelhead ecotype in multiple populations from Oregon and Northern California. *Conserv Gen* doi: <http://dx.doi.org/10.1007/s10592-015-0769-2>.
- Banks MA, Rashbrook VK, Calavetta MJ, Dean CA, Hedgecock D. 2000. Analysis of microsatellite DNA resolves genetic structure and diversity of chinook salmon (*Oncorhynchus tshawytscha*) in California's Central Valley. *Can J Fish Aquat Sci* 57:915–927. doi: <http://dx.doi.org/10.1139/f00-034>.
- Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. 2004. GENETIX 4.05: logiciel sous Windows™ pour la génétique des populations. *Laboratoire Génome: populations, interactions*, CNRS UMR 5000. Montpellier (France): *Université de Montpellier II*.
- Brown R. 2005. Understanding Central Valley Chinook salmon and steelhead: it is time to get off the dime! *San Franc Estuary Watershed Sci* [Internet] [cited 2012 Sep 07];3(1). Available from: <http://escholarship.org/uc/item/60c30881> doi: <http://dx.doi.org/10.15447/sfews.2005v3iss1art2>
- Brunelli JP, Wertzler KJ, Sundin K, Thorgaard GH. 2008. Y-specific sequences and polymorphisms in rainbow trout and Chinook salmon. *Genome* 51:739–748. doi: <http://dx.doi.org/10.1139/G08-060>.
- Busack CA, Gall GAE. 1980. Ancestry of artificially propagated California rainbow trout strains. *Cal Fish Game* 66:17–24.
- Busby PJ, Wainwright TC, Bryant GJ, Lierheimer LJ, Waples RS, Waknitz FW, Lagomarsino IV. 1996. Status review of West Coast steelhead from Washington, Idaho, Oregon, and California. NOAA Tech Memo. NMFS–NWFSC–27.
- [California HSRG] California Hatchery Scientific Review Group. 2012. California Hatchery Review Report. Prepared for the US Fish and Wildlife Service and Pacific States Marine Fisheries Commission. June 2012. 100 p.
- Campbell NR, Overturf K, Narum SR. 2009. Characterization of 22 novel single nucleotide polymorphism markers in steelhead and rainbow trout. *Mol Ecol Resour* 9:318–322. doi: <http://dx.doi.org/10.1111/j.1755-0998.2008.02376.x>.
- Cavalli-Sforza LL, Edwards AWF. 1967. Phylogenetic analysis. Models and estimation procedures. *Am J Human Gen* 19:233–257.
- Clemento AJ, Anderson EC, Boughton D, Girman D, Garza JC. 2009. Population genetic structure and ancestry of *Oncorhynchus mykiss* populations above and below dams in south-central California. *Conserv Gen* 10:1321–1336. doi: <http://dx.doi.org/10.1007/s10592-008-9712-0>.
- Deiner K, Garza JC, Coey R, Girman DJ. 2007. Population structure and genetic diversity of trout (*Oncorhynchus mykiss*) above and below natural and manmade-barriers in the Russian River, California. *Conserv Gen* 8:437–454. doi: <http://dx.doi.org/10.1007/s10592-006-9183-0>.
- Federal Register. 2006. Endangered and threatened species: final listing determinations for 10 distinct population segments of West Coast steelhead. *Federal Register* 71:834–862.
- Felsenstein J. 2003. Inferring phylogenies. Sunderland (MA): Sinauer Associates.
- Felsenstein J. 2005. PHYLIP v. 3.69 c Phylogeny inference package. Seattle (WA): Department of Genetics, University of Washington. Available from: <http://evolution.genetics.washington.edu/phylip.html>

Finger AJ, Anderson EC, Stephens MR, May BP. 2011. Application of a method for estimating effective population size and admixture using diagnostic single nucleotide polymorphisms (SNPs): implications for conservation of threatened Paiute cutthroat trout (*Oncorhynchus clarkii seleniris*) in Silver King Creek, California. *Can J Fish Aquatic Sci* 68:1369–1386. doi: <http://dx.doi.org/10.1139/f2011-075>.

[Fish and Game Code] California Department of Fish and Wildlife. 2012. Fish and Game Code, Chapter 7.2, §1725–1730. Trout and Steelhead Conservation and Management Planning Act of 1979. Available from: <http://www.leginfo.ca.gov/cgi-bin/displaycode?section=fgc&group=01001-02000&file=1725-1730>

Garza JC, Gilbert-Horvath EA, Spence BC, Williams TH, Fish H, Gough SA, Anderson JH, Hamm D, Anderson EC. 2014. Population structure of steelhead in coastal California. *Trans Am Fish Soc* 143:134–152. doi: <http://dx.doi.org/10.1080/00028487.2013.822420>.

Jombart T. 2008. *adeigenet*: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405. doi: <http://dx.doi.org/10.1093/bioinformatics/btn129>.

Jombart T, Ahmed I. 2011. *adeigenet* 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27:3070–3071. doi: <http://dx.doi.org/10.1093/bioinformatics/btr521>.

Kalinowski ST. 2005. HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic diversity. *Mol Ecol Notes* 5:187–189. doi: <http://dx.doi.org/10.1111/j.1471-8286.2004.00845.x>.

Kinziger, AP, Hellmair M, Hankin DG, and Garza JC. 2013. Contemporary population structure in Klamath River basin Chinook salmon revealed by analysis of microsatellite data. *Trans Am Fish Soc* 142:1347–1357. doi: <http://dx.doi.org/10.1080/00028487.2013.806351>.

Leitritz E. 1970. A history of California's fish hatcheries 1870–1960. California Department of Fish and Game, Fish Bull. 150:1–125.

Lindley ST, Schick RS, Agrawal A, Goslin M, Pearson TE, Mora E, Anderson JJ, May B, Greene S, Hanson C, Low A, McEwan D, MacFarlane RB, Swanson C, Williams JG. 2006. Historical population structure of Central Valley steelhead and its alteration by dams. *San Franc Estuary Watershed Sci* [Internet]. [cited 2012 Sep 07];4(1). Available from: <http://escholarship.org/uc/item/1ss794fc> doi: <http://dx.doi.org/10.15447/sfewes.2006v4iss1art3>

Meek MH, Stephens MR, Tomalty KM, May B, Baerwald MR. 2014. Genetic considerations for sourcing steelhead reintroductions: investigating possibilities for the San Joaquin River. *San Franc Estuary Watershed Sci* [Internet]. [cited 2012 Sep 07];12(1). Available from: <http://escholarship.org/uc/item/6wn5q90h> doi: <http://dx.doi.org/10.15447/sfewes.2014v12iss1art2>

Morin PA, Luikart G, Wayne RK, SNP Workshop Group. 2004. SNPs in ecology, evolution and conservation. *Trends Ecol Evol* 19:208–216. doi: <http://dx.doi.org/10.1016/j.tree.2004.01.009>.

Narum SR, Banks M, Beacham TD, Bellinger MR, Campbell MR, DeKoning J, Elz A, Guthrie CM, Kozfkay C, Miller KM, Moran P, Phillips R, Seeb LW, Smith CT, Warheit K, Young SF, Garza JC. 2008. Differentiating salmon populations at broad and fine geographic scales with microsatellites and single nucleotide polymorphisms. *Mol Ecol* 17:3464–3477. doi: <http://dx.doi.org/10.1111/j.1365-294X.2008.03851.x/full>.

Needham PR, Behnke RJ. 1962. The origin of hatchery rainbow trout. *Prog Fish-Cult* 24:156–158. doi: [http://dx.doi.org/10.1577/1548-8659\(1962\)24\[156:TOOHR T\]2.0.CO;2](http://dx.doi.org/10.1577/1548-8659(1962)24[156:TOOHR T]2.0.CO;2).

Nei M. 1987. Molecular evolutionary genetics. New York (NY): Columbia University Press.

- Nielsen JL, Pavey SA, Wiacek T, Williams I. 2005. Genetics of Central Valley *O. mykiss* populations: drainage and watershed scale analysis. San Franc Estuary Watershed Sci [Internet]. [cited 2012 Sep 07];3(2). Available from: <http://escholarship.org/uc/item/6sc3905g> doi: <http://dx.doi.org/10.15447/sfew.2005v3iss2art3>
- Palstra FP, O'Connell MF, Ruzzante DE. 2007. Population structure and gene flow reversals in Atlantic salmon (*Salmo salar*) over contemporary and long-term temporal scales: effects of population size and life history. Mol Ecol 16:4504–4522. doi: <http://dx.doi.org/10.1111/j.1365-294X.2007.03541.x>.
- Pearse DE, Donohoe C, Garza JC. 2007. Population genetics of steelhead (*Oncorhynchus mykiss*) in the Klamath River. Environ Biol Fishes 80:377–387. doi: <http://dx.doi.org/10.1007/s10641-006-9135-z>.
- Pearse DE, Hayes SA, Anderson EC, Bond MH, Hansen CV, MacFarlane RB, Garza JC. 2009. Over the falls? Rapid evolution of ecotypic differentiation in steelhead/rainbow trout (*Oncorhynchus mykiss*). J Heredity 100:515–525. doi: <http://dx.doi.org/10.1093/jhered/esp040>.
- Pearse DE, Gunckel SL, Jacobs SE. 2011a. Population structure and genetic divergence of coastal Rainbow and Redband trout in the Upper Klamath Basin. Trans Am Fish Soc 140:587–597. doi: <http://dx.doi.org/10.1080/00028487.2011.583538>.
- Pearse DE, Martinez E, Garza JC. 2011b. Disruption of historical patterns of isolation by distance in coastal steelhead. Conserv Gen. 12:691–700. doi: <http://dx.doi.org/10.1007/s10592-010-0175-8>.
- Pearse DE, Miller MR, Abadía-Cardoso A, Garza JC. 2014. Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. Proc Roy Soc B: Biol Sci 281:20140012. doi: <http://dx.doi.org/10.1098/rspb.2014.0012>.
- Primmer CR, Veselov AJ, Zubchenko A, Poututkin A, Bakhmet I, Koskinen MT. 2006. Isolation by distance within a river system: genetic population structuring of Atlantic salmon, *Salmo salar*, in tributaries of the Varzuga River in northwest Russia. Mol Ecol 15:653–666. doi: <http://dx.doi.org/10.1111/j.1365-294X.2005.02844.x>.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.
- Rannala B, Mountain JL. 1997. Detecting immigration by using multilocus genotypes. Proc Nat Acad Sci USA 94:9197–9221.
- Raymond M, Rousset F. 1995. Genepop version 1.2: Population genetics software for exact tests and ecumenicism. J Heredity 86:248–249.
- Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425.
- Seeb LW, Antonovich A, Banks MA, Beacham TD, Bellinger MR, Campbell MR, Decovich NA, Garza JC, Guthrie CM, Moran P, Narum SR, Stephenson JJ, Supernault KJ, Teel DJ, Templin WD, Wenburg JK, Young SF, Smith CT. 2007. Development of a standardized DNA database for Chinook salmon. Fisheries 32:540–552. doi: [http://dx.doi.org/10.1577/1548-8446\(2007\)32\[540:DOASDD\]2.0.CO;2](http://dx.doi.org/10.1577/1548-8446(2007)32[540:DOASDD]2.0.CO;2).
- Srikwan S, Woodruff DS. 2000. Genetic erosion in isolated small mammal populations following rain forest fragmentation. In: Young A, Clarke G, editors. Genetics, demography and viability of fragmented populations. Cambridge (UK): Cambridge University Press. p. 149–172.
- Stephens MR, Clipperton NW, May B. 2009. Subspecies-informative SNP assays for evaluating introgression between native golden trout and introduced rainbow trout. Mol Ecol Resour 9:339–343. doi: <http://dx.doi.org/10.1111/j.1755-0998.2008.02407.x>.

Takezaki N, Nei M. 1996. Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* 144:389–399.

Warnes G, Leisch F. 2005. Genetics: population genetics. R package version 1.2.0. [cited 2012 Sep 07]. Available from: <http://cran.r-project.org/web/packages/genetics/index.html>

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population-structure. *Evolution* 38:1358–1370. doi: <http://dx.doi.org/10.2307/2408641>.

Wood, DA, Halstead, BJ, Casazza ML, Hansen EC, Wylie, GD, Vandergast AG. 2015. Defining population structure and genetic signatures of decline in the giant gartersnake (*Thamnophis gigas*): implications for conserving threatened species within highly altered landscapes. *Conserv Gen* 16:1025–1039. doi: <http://dx.doi.org/10.1007/s10592-015-0720-6>.

NOTES

Vendrami D, Pearse DE, Anderson EC, Garza JC. 2014. Analysis of Central Valley steelhead hatchery broodstock, unpublished data. NOAA/NMFS Southwest Fisheries Science Center, Santa Cruz, CA. Available from: devon.pearse@noaa.gov