Conservation Genetics **3:** 261–276, 2002.

Morphological and genetic analysis of the Red Hills roach (Cyprinidae: *Lavinia symmetricus*)

W.J. Jones[1*], B.D. Quelvog[2] & G. Bernardi[1]
[1]*Biology Department, University of California, Santa Cruz, California 95064+0001, USA;* [2]*California Department of Fish and Game, 21278 Phoenix Lake Road, Sonora, California 95370+8226, USA (*Author for correspondence: E-mail: william.jones@uni-konstanz.de)*

Received 18 June 2001; accepted 20 October 2001

Key words: California, *Hesperoleucus*, PCA, mtDNA, serpentine

Abstract

We performed a morphological and phylogenetic analysis of a recently discovered population of California roach (Red Hills roach; Cyprinidae: *Lavinia symmetricus*) to determine if the degree of separation of these populations warrants subspecies status. Previous morphological analysis by Brown et al. (1992) suggested that *L. symmetricus* from the Red Hills roach type locality (Horton Creek) were different for Principal Discriminant Scores (based on 15 morphological characters) from neighboring populations. Similarly, recent work by our lab on the phylogenetics of the putative subspecies of *L. symmetricus* has revealed that the Red Hills roach appears reciprocally monophyletic for assayed mitochondrial DNA markers. In addition to performing further morphological and genetic analysis of the Red Hills roach, we increased our sampling effort in the Red Hills region to determine the distribution of this undescribed putative subspecies. Our morphological results are generally in agreement with Brown et al. (1992) except in regards to Horton Creek. A significant difference between the two studies exists for all Horton Creek multivariate analyses as well as frequency of a presumably derived character (chisel lip). Our genetic results also support our previous findings that the Red Hills roach is a diagnosible, genetically distinct population. The phylogenetic analysis suggests that populations on the other side of a large reservoir were recently connected via gene flow to the Red Hills populations.

Introduction

The California roach (*Lavinia symmetricus*) is a cyprinid (minnow) often less than 10 cm standard length (SL) with a range of colors from dark brown to light silver. It represents the only freshwater dispersant (Moyle and Cech 2000) fish found in several rivers and creeks of northern California (Moyle 1976). The California roach is a benthic omnivore (Moyle 1976) which often feeds on algal tufts (Powers 1990). The California roach is of little economic value making funding and conservation of rare populations especially difficult. Despite these drawbacks, *L. symmetricus* represents one of the few species of native western freshwater fishes that has maintained a relatively widespread geographic distribution although this may be due to the presence of multiple taxa (Jones and Bernardi submitted). As with other freshwater fishes, it has been and continues to be impacted by habitat degradation and water diversion.

The Red Hills roach is a recently discovered population of *L. symmetricus* (Brown et al. 1992). Known populations of Red Hills *L. symmetricus* appear to be restricted to spring-fed pools and tributaries of Six Bit Gulch in Tuolumne County, California (Brown et al. 1992; Moyle et al. 1995). The Red Hills region is currently listed as a potential Area of Critical Environmental Concern by the Bureau of Land Management (BLM) as well as an Aquatic Diversity Management Area (Moyle 1996). Currently, a portion (ca. 15%) of the Red Hills region occurs on BLM land. However, the main portion of aquatic diversity occurs in the

lower course of Six Bit Gulch (private land) which neighbors the BLM land. In addition to the Red Hills roach, the Red Hills region contains several endemic plant species (i.e., *Verbena californica*; California verbena) as well as many native perennials which appear to survive due to the fact that most invasive plant species cannot tolerate the harsh soil conditions of the Red Hills. In addition to plants, several rare animal species such as wintering bald eagles and the foothill yellow-legged frog occur within the Red Hills region.

On the basis of frequency of a cartilaginous plate on the lower jaw (i.e., "chisel lip") which is presumably an adaptation for scraping algae from rocks and Principle Component Analysis (PCA), Brown et al. (1992) concluded that each *L. symmetricus* population surveyed (including the Red Hills) was morphologically distinct from all others. However, two main problems exist with basing management decisions on the conclusions drawn in Brown et al. (1992). First, morphological differences among populations may represent simply a phenotypic response to environmental conditions as opposed to genetic differences. Second, the data in Brown et al. (1992) represent a single sampling period. Whether or not morphological differences between populations is consistent over time is unclear. In order to create a management plan for conservation of the Red Hills roach, these key questions must first be addressed.

Since recent results from our ongoing research on the population genetics of *L. symmetricus* (Jones and Bernardi submitted), we increased our sampling effort in the Red Hills region. Previous analysis of mitochondrial (mtDNA) and nuclear DNA (nDNA) sequence data suggested that the Red Hills roach is reciprocally monophyletic for mtDNA relative to all other putative subpecies of *L. symmetricus*. Previous work by Avise (1975) showed little promise of detecting genetic differences among putative subspecies of *L. symmetricus* using allozymes. However, with the development of suitable DNA markers, several key questions regarding the evolutionary distinctiveness of the Red Hills roach can be addressed. First, based on a neutrally evolving marker (mtDNA), are there genetic differences between the Red Hills *L. symmetricus* and neighboring populations? The answer appears to be "yes" but sample sizes were small (Jones and Bernardi, submitted). Second, is there a correlation between morphological and genetic distinctiveness? If the unique morphology of the Red Hills roach is a phenotypic response to an environmental parameter (i.e., serpentine geology of Red Hills), populations in the same geographic region (Red Hills) but in separate river drainages (Tuolumne and Stanislaus Rivers) should show a similar morphology but be distinct genetically. Third, how much gene flow occurred among populations before the 1921 construction of the Don Pedro dam? The reservoir created by the formation of the dam is the main barrier to movement of individuals between the Red Hills and neighboring populations (Brown et al. 1992). Fourth, is the Red Hills population characterized by the presence of novel haplotypes? From the distribution of mtDNA haplotypes (Jones and Bernardi submitted), three haplotypes were found in the Red Hills populations. These three haplotypes were not found in any of the neighboring *L. symmetricus* populations suggesting that all haplotypes observed in the Red Hills populations were novel. Finally, is there evidence from DNA sequence data that the Red Hills *L. symmetricus* has undergone a reduction in effective population size and genetic diversity as a result of isolation? Given the restricted distribution and small census size (hundreds of individuals; Brown et al. 1992) of known Red Hills roach populations, a low number of haplotypes in the Red Hills roach relative to neighboring populations would be expected.

In this present paper, we had three main goals. The first was to increase sampling area to determine the geographic distribution of the Red Hills roach. Since Brown et al. (1992), additional populations of *L. symmetricus* have been discovered in Amber and Roach Creeks neighboring Horton Creek, the Red Hills roach type location (all tributary of Six Bit Gulch; Moyle et al. 1995; B. Quelvog, unpublished data), as well in creeks on the other side of the Don Pedro Reservoir opposite the Red Hills (ca. 3 miles to Becca B Creek and 8 miles to Hatch, First, and Second Creeks). Our second objective was to duplicate the morphological study of Brown et al. (1992) by analyzing samples collected from 1992–2000. Our goal was to determine if the Red Hills roach was morphologically distinct from neighboring populations as suggested by Brown et al. (1992). Finally, our third objective was to collect and analyze additional DNA sequences to determine if the Red Hills roach is reciprocally monophyletic as suggested by our earlier study.

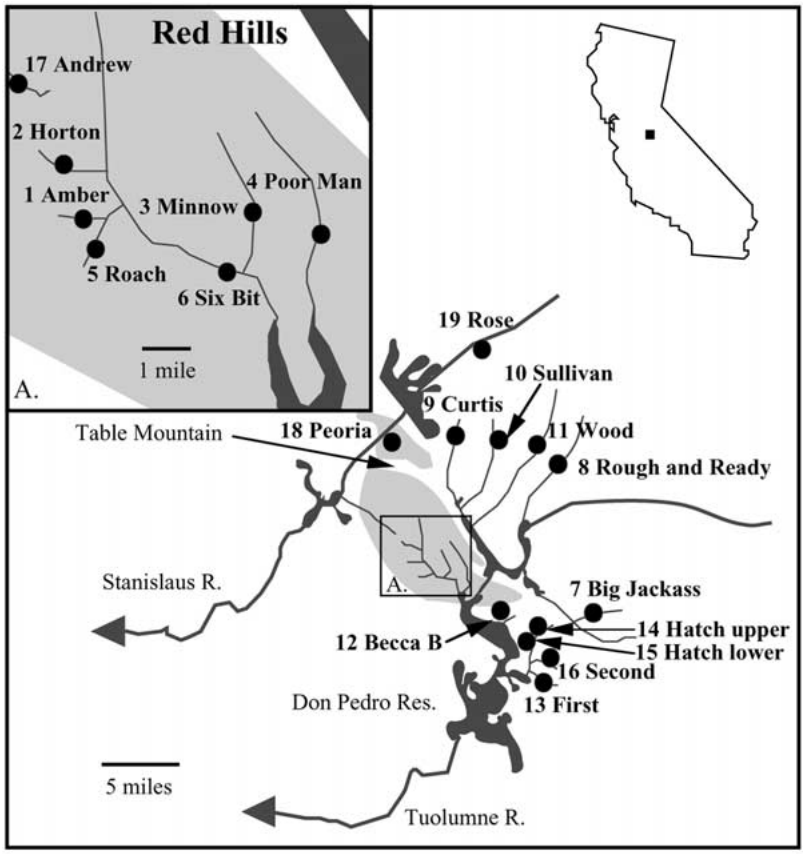

Figure 1. Sampling localities for *L. symmetricus* in the Tuolumne River, Red Hills, and Stanislaus Rivers. Inset (A) shows sampling locations within the Six Bit Gulch and Andrew's Creek portion of the Red Hills. Specific names for localities are listed according to numbers in Table 1. Shaded area indicates approximate distribution of the Red Hills serpentine outcropping.

Methods

Sampling

Fish were collected from a total of 16 populations for morphological analysis and 19 populations for genetic analysis (Figure 1; Table 1). Sampling localities were partitioned into 5 geographic areas for analysis: Red Hills, East Tuolumne, North Tuolumne, South Tuolumne, and Stanislaus River (Table 1). Individuals were collected using backpack electroshocker or baited minnow traps. Samples for morphological analysis were preserved in 10% buffered formalin and then transferred to 70% ethanol for storage except for Rough and Ready Creek (1999), Becca B. Creek (2000), and Rose Creek (2000) which were placed directly into 70% ethanol. Tissue samples for genetic analysis were collected by fin clipping a lower portion of the caudal fin then releasing the individual except for Rough and Ready Creek (1999), Becca B. Creek (2000), and Rose Creek (2000) where fin clips were taken from samples collected for morphological analysis. All fin clips were stored in 95% ethyl alcohol.

Morphological analysis

Morphological characters were measured to the nearest 0.001 inch using vernier calipers by one of us (BDQ). For comparison with the Brown et al. (1992) dataset, all measurements were converted to the nearest 0.1 mm. Characters measured included standard length (SL), head length (HDL – the distance from the snout to the most posterior point of the operculum), head width (HDW), body depth (BD), body width (BW), trunk shape (TS – the distance from the point of maximum body depth to the snout), caudal peduncle length (PL – the distance from a line perpen-

Table 1. Sampling sites, codes, corresponding locality numbers to Figure 1, population designation, and year and sample size (N) for morphology and genetic survey of *L. symmetricus*

Site	Code	Figure 1	Population	Year morphology	N morphology	Year genetics	N genetics
Amber Creek	AMB	1	Red Hills	1992, 1999	30	1998, 2000	5, 3
Horton Creek	HOR92	2	Red Hills	Brown et al. 1992	50	n/a	n/a
" " "	HOR00	2	Red Hills	1999	36	1998, 1999	4, 9
Minnow Gulch	MIG	3	Red Hills	1992	30	1999	3
Poor Man's Gulch	PMG	4	Red Hills	1993	10	1999	6
Roach Creek	ROA	5	Red Hills	1992	30	1998, 2000	3, 2
Six Bit Gulch	SBG	6	Red Hills	n/a	n/a	1999	5
Big Jackass Creek	BJC	7	East Tuolumne	1998, 2000	23	1999	5
Rough and Ready Creek	RRC92	8	East Tuolumne	Brown et al. 1992	8	n/a	n/a
" " "	RRC00	8	East Tuolumne	1999	25	1999	5
Curtis Creek	CUR92	9	North Tuolumne	Brown et al. 1992	4	n/a	n/a
" " "	CUR00	9	North Tuolumne	1999	23	1999	5
Sullivan Creek	SUL	10	North Tuolumne	n/a	n/a	1998	3
Wood Creek	WOC92	11	North Tuolumne	Brown et al. 1992	31	n/a	n/a
" " "	WOC00	11	North Tuolumne	1993	21	1998	5
Becca B Creek	BBC	12	South Tuolumne	2000	23	2000	15
First Creek	FIR	13	South Tuolumne	1995	23	1999, 2000	5, 5
Hatch Creek upper	HAT	14	South Tuolumne	1995	30	1999	5
Hatch Creek lower	HAT	15	South Tuolumne	n/a	n/a	2000	5
Second Creek	SEC	16	South Tuolumne	1995	24	1999, 2000	5, 5
Andrew's Creek	AND	17	Stanislaus River	1992	30	1999	5
Peoria Creek	PEO	18	Stanislaus River	1997	30	2000	5
Rose Creek	ROS92	19	Stanislaus River	Brown et al. 1992	28	n/a	n/a
" " "	ROS00	19	Stanislaus River	2000	10	2000	5
Total				All samples	519		123
				2000 samples only	398		

dicular to the body axis located at the most posterior part of the insertion of the anal fin to the end of the vertebral column), caudal peduncle depth (PD – depth at the midpoint of the caudal peduncle), pectoral fin length (PCFL), pectoral fin width (PCFW), pelvic fin length (PVFL), eye diameter (ED), and interorbital distance (ID) (descriptions taken from Brown et al. 1992). Occurrence of the chisel lip condition was recorded as not present (0), poorly developed (1), well developed (2), or extremely well developed (3). Differences in untransformed morphological characters were compared using pairwise t-tests (α = 0.05; Bonferroni corrected; Rice 1989). Differences between populations for the chisel lip character was tested using chisel lip frequency in a chi-square test.

In order to correct for differences in overall size (i.e., standard length) between sampling localities (see Results), we chose to use two multivariate methods, Burnaby size correction and Canonical discriminant analysis. As discussed in Brown et al. (1992) and Rohlf and Bookstein (1987), the Burnaby algorithm (Burnaby 1966) is the most appropriate method for size adjustment. Burnaby Principal Component Analysis (PCA) of the morphological and meristic data was performed with the computer program, Burnaby (N. MacCleod, personal communication), for all 15 measured morphological characters. In addition, we excluded four characters (standard length (SL), pectoral fin length (PCFL), pectoral fin width (PCFW), and pelvic fin length (PVFL)) from the dataset because of possible intersex

differences (Murphy 1943; Brown et al. 1992). This reduced dataset of 11 characters was analyzed in the same manner as the full dataset. As in Brown et al. (1992), the mean ±2 standard errors for each population was estimated for each of the first two discriminant axes (PC1 and PC2) calculated from Burnaby. These values were used to plot population ellipses in discriminant space to determine if there were similarities in morphologies among sampling localities.

Canonical discriminant analysis (JMP Statistics and Graphics Guide 1994) was also used to group individuals among populations. Percent correct classification of each population was assessed using Canonical discriminant analysis on both the full and the reduced dataset as described above. The first two discriminant scores (Canonical 1 and Canonical 2) were used to plot 95% ellipses for each of the sampling localities. Character loading was also analyzed to determine which of the measurements contributed to the Canonical discriminant scores.

Genetic analysis

Samples of Oregon pikeminnow (*Pytocheilus oregonensis*, n = 1), Sacramento pikeminnow (*P. grandis*, n = 1), and hardhead (*Mylopharodon concephalus*, n = 2) were used as outgroup taxa (Avise, 1975, T. Dowling, personal communication) to root the gene trees. Approximately 5 mm^2 of fin from each individual *L. symmetricus* was placed at 37 °C for 5 minutes or until all alcohol had evaporated. DNA was extracted by digesting the dried fin clipping in 750 μl of extraction buffer (10 mm Tris HCl, pH 8.2; 10 mm EDTA; 200 mm NaCl; 0.5% SDS; 200 mg/ml Proteinase K) at 55 °C for approximately 12 hours. The solution was extracted once with an equal volume of chloroform and once with an equal volume of isoamyl alcohol. Genomic DNA was precipitated with 200 mm NaCl and two volumes of 100% ethanol and centrifuged at 13,000 RPM for 30 minutes at room temperature. The DNA pellet was washed with 70% ethanol (–20 °C), dried in a vacuum centrifuge, and resuspended in 100 μl of autoclaved MilliQ water and stored at –20 °C.

Polymerase chain reactions (PCR) were performed as 50 μl reactions in a Perkin Elmer Thermal Cycler (480 or 9700) with negative controls. A portion of the mitochondrial NADH subunit 2 (NADH-2) was amplified using the NDR1 and NDF1 primers for 35 cycles (94 °C 1 minute, 55 °C 1 minute, 72 °C 2 minutes) (Jones and Bernardi submitted). The quality of the PCR product was checked by electrophoresing 5 μl of the PCR product on a 2% 1 X TBE agarose gel and the bands visualized with ethidium bromide. Purified PCR products (Qiagen) were labeled using dye-labeled nucleotides (Applied Biosystems). PCR products were sequenced in one direction consistently following manufacturer's recommendations and protocols (Applied Biosystems). Each unique haplotype was sequenced in the opposite direction to confirm mutations. Sequencing products were electophoresed on a 4% bis-acrylamide gel (38:2 acrylamide:bis-acrylamide) for 16 hours (ABI 373). Electropherograms were imported into Sequence Navigator and all mutations checked by hand. Corrected sequences were copied into a NEXUS file. Individuals possessing the same haplotypes were identified in MacClade 3.05 (Maddison and Maddison 1992). Estimates of population subdivision (Fst) were performed in Arlequin 2.001 (Schneider et al. 1999). Gene flow (*Nfm*) was estimated from Fst by:

$$Nf M = (1 - \mathrm{FST})/2\ \mathrm{FST}$$

(Page & Holmes 1998; Table 4.3)

Phylogenetic relationships among mtDNA haplotypes were estimated using two methods. First, maximum parsimony trees were estimated in PAUP* with transversions (TV) weighted equal to transitions (TS) as well as weighted (TV:TS $\approx$ 7:1; according to maximum likelihood estimates) using heuristic searches with 10 random additions to increase accuracy (Swofford 1993, 1998). Second, maximum likelihood trees were estimated using the Kimura (1980) model with equal base frequencies, two substitution types, gamma-distributed rates with estimated shape parameter (0.27), and estimated transition:transversion ratio (7.0585) (model suggested under the Hierarchical Likelihood Ratio Tests (hLRTs); MODELTEST version 3.04; Posada and Crandall 1998).

Bootstrapping (Felsenstein 1985) was used to estimate the reliability of phylogenetic reconstructions (500 replicates) of the maximum parsimony trees. Partitioning of genetic variation (AMOVA; Excoffier et al. 1992) was assessed using the five groupings (Table 1) as implemented in Arlequin 2.001 (Schneider et al. 1999).

All morphological datafiles (Tab delimited format) and genetic datasets (NEXUS format and AMOVA format) with associated settings and parameters are available at <http://www.biology.ucsc.

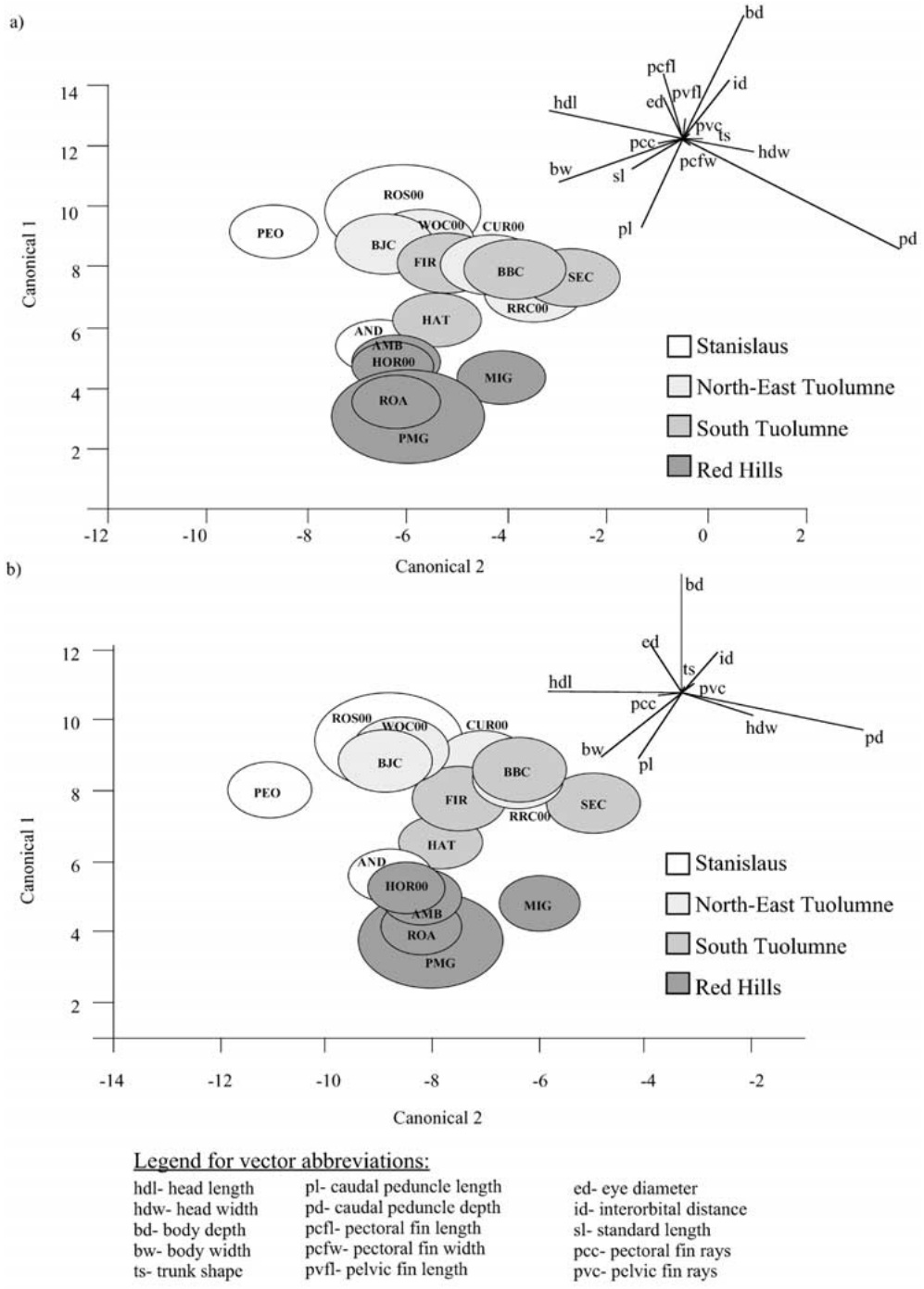

Figure 2. (a) Canonical principle component graphs for the first two canonical scores (Canonical 1 and Canonical 2) using the full morphological dataset (15 characters). Ellipses represent 95% density ellipses for each population. Population codes are the same as in Table 1. Vectors indicate relative contribution of characters (see legend for full names) to canonical scores. (b) Canonical principle component graphs for the first two canonical scores (Canonical 1 and Canonical 2) using the reduced morphological dataset (11 characters). Labeling and shading as in Figure 2a.

edu/people/bernardi/jones/red_hills/> or via e-mail request from WJJ.

Results

Morphology

Analysis of overall size differences prior to size correction between the 16 sampling localities for morphological analysis revealed that there was a significant difference between localities for standard length (SL) with 15 of 153 pairwise comparisons significant after Bonferroni correction (Rice 1989; data available upon request from WJJ). Analysis of character loading for both the Burnaby and Canonical analysis revealed relatively equal weighting except for trunk shape (TS) and the two meristic counts (pectoral fin rays (PCC) and pelvic fin rays (PVC)) (data not shown). The vectors showing the relative contribution of characters to the canonical analysis are shown in Figure 2.

Using the Burnaby and Canonical multivariate analysis on the combined dataset from the Tuolumne and Stanislaus River drainages from Brown et al. (1992) and this study, a total of 519 individuals were available for morphological analysis. Excluding the samples from Brown et al. (1992) left 398 individuals. Comparison of our results with the results of Brown et al. (1992) were not directly comparable for the Rough and Ready Creek and the Curtis Creek samples due to small sample sizes in Brown et al. (1992) (n = 8 and 4, respectively; Table 1). However, the remaining three populations (Rose, Woods, and Horton Creeks) had comparable sample sizes between the two studies and were considered for between year multivariate analysis.

Morphological differences between sampling periods

For Rose Creek (Stanislaus River) and Woods Creek (Tuolumne River), the principal component scores for the first two Burnaby discriminant axes were similar between years (Figure 3). Similarly, the chisel lip condition was variable between years, but generally in low frequency for Rose and Woods Creeks (ROS92 – 17.86% vs. ROS00 – 0.00% and WOC92 – 3.23% vs. WOC00 – 14.29%; Table 2). Although collected from the same locality, there was a significant difference between years for the canonical discriminant analysis. Analysis of the population ellipses comparing sites between years revealed a distinct grouping of Brown et al.'s sites versus those in this study (data and figures available from WJJ). Similarly, percent classification for populations were generally high for a given year rather than by site suggesting that individuals from the same locality were different in canonical scores between years (data not shown). Multivariate classification of samples to the correct creek locality were moderately high for both Rose and Woods Creeks with no large difference in classification between years (ROS92 – 58.0–66.3% vs. ROS00 – 56.0–64.2% and WOC92 – 66.2–68.6% vs. WOC00 – 63.1–63.6%; Table 2).

In contrast to Rose and Woods Creeks, Horton Creek showed a highly significant difference between years for both the Burnaby and the Canonical analysis. For the first two Burnaby discriminant axes for both the full and the reduced datasets, there was little to no overlap between the HOR92 and HOR00 density ellipses (Figure 3). Further, a significant difference in the frequency of the chisel lip condition between years for the Horton Creek locality was evident. In the Brown et al. (1992) dataset, nearly all (94%) Horton Creek individuals exhibited the chisel lip condition with many (32%) showing an extremely well-developed chisel lip (Table 2). Samples from the Horton Creek locality in 1999 (HOR00) showed a much lower incidence of the chisel lip condition (66.7%) and no occurrence of the extremely well-developed chisel lip (Table 2).

Finally, the morphological results of the Horton Creek samples differed dramatically between years. First, analysis of untransformed morphological values revealed a highly significant difference in interorbital distance (ID; $p = 0.0009$), caudal peduncle length (PL; $p = 0.0027$), pectoral fin length (PCFL; $p = 0.0087$), and pelvic fin ray count (PVC; $p = 0.0036$) between Horton Creek samples from the two sampling periods ($\alpha = 0.05$; Bonferroni corrected). The difference in the interorbital distance between years for Horton Creek does not appear to be due to size (standard length; $p = 0.5964$) or other measurements associated with the head such as eye diameter (ED) or head width (HDW) ($p = 0.9764$ and $p = 0.1366$; respectively). Similarly, sampling size seems to not be an important factor in between-year differences for Horton Creek (Table 1), nor does measurement bias since the Horton Creek samples from Brown et al. (1992) and this study were measured by the same individual (B. Quelvog). Other between-year comparisons revealed only one significant comparison for Woods Creek (head width (HDW); $p = 0.0032$).

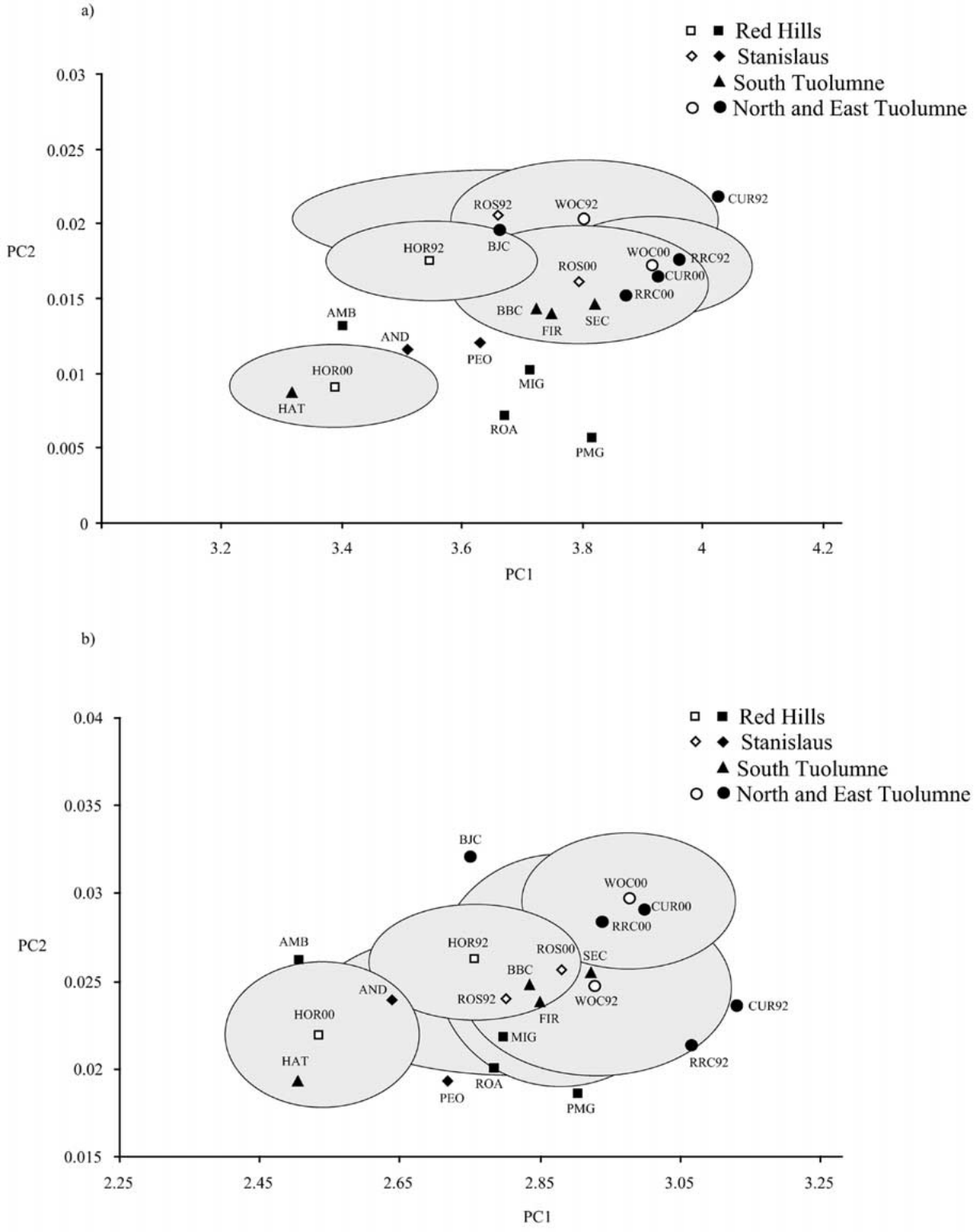

Figure 3. (a) Burnaby principle component graphs for the first two principle component scores (PC 1 and PC 2) using the full morphological dataset (15 characters). Means for each population are plotted. Standard error (±2) ellipses are given only for between year comparisons for figure clarity. Means for between year comparisons are indicated by white symbols while all other means are indicated by black symbols. Population codes are the same as in Table 1. (b) Burnaby principle component graphs for the first two principle component scores (PC 1 and PC 2) using the reduced morphological dataset (11 characters). Symbols as in Figure 3a.

Table 2. Chisel lip frequency in surveyed *L. symmetricus* populations

	Population	Character state					% Lip	Group mean
		0	1	2	3	Total		
AMB	Red Hills	27	3	0	0	30	10.00	
HOR92		3	21	10	16	50	94.00	
HOR00		12	18	6	0	36	66.67	
MIG		4	24	1	1	30	86.67	
PMG		9	1	0	0	10	10.00	
ROA		21	8	1	0	30	30.00	49.56
BJC	East Tuolumne	21	2	0	0	23	8.70	
RRC92		8	0	0	0	8	0.00	
RRC00		25	0	0	0	25	0.00	2.90
CUR92	North Tuolumne	4	0	0	0	4	0.00	
CUR00		20	3	0	0	23	13.04	
WOC92		30	0	1	0	31	3.23	
WOC00		18	3	0	0	21	14.29	7.64
BBC	South Tuolumne	23	0	0	0	23	0.00	
FIR		23	0	0	0	23	0.00	
HAT		21	8	1	0	30	30.00	
SEC		24	0	0	0	24	0.00	7.50
AND	Stanislaus	23	5	1	1	30	23.33	
PEO		30	0	0	0	30	0.00	
ROS92		23	2	3	0	28	17.86	
ROS00		10	0	0	0	10	0.00	10.30
	Total	379	98	24	18	519		

For the HOR92 samples, classification to Horton Creek was high for both the full (99.9%) and reduced (97.6%) datasets. For the HOR00 samples, classification to Horton Creek were some of the lowest of all observed values in the entire dataset (51.9% full and 38.4% reduced; Table 3). This low classification score for HOR00 samples was due to the fact that samples from Amber and Roach Creeks (both in the Six Bit Drainage of the Red Hills) were often misclassified into the Horton Creek category (data not shown). Perhaps due to the fact that the only Red Hills population sampled in Brown et al. (1992) was Horton Creek, the values for correct multivariate classification are lower in this study due to increased sampling of the Red Hills region.

Morphological differences between Red Hills and neighboring populations

As in Brown et al. (1992), principle component analysis revealed that the Red Hills populations (Amber, Horton, Minnow Gulch, Poor Man Gulch, Roach, and Six Bit Gulch) were morphologically distinct from all other populations. For the most part, all Tuolumne River samples cluster in the same Burnaby grouping (high PC1 and high PC2 scores; Figure 3) while the Red Hills populations generally had low PC1 and low PC2 scores. For the samples collected from the South Tuolumne populations (Becca B, First, Hatch, Second), Burnaby PCA discriminant scores were generally intermediate the Red Hills and the North-East Tuolumne populations with all of the South Tuolumne populations grouping

Table 3. Percent correct classification according to sampling locality based on canonical discriminant analysis

	Population	n	*a priori*	Full dataset	Reduced dataset
AMB	Red Hills	30	0.06	57.1	48.4
HOR92		50	0.10	99.9	97.6
HOR00		36	0.07	51.9	38.4
MIG		30	0.06	79.8	76.1
PMG		10	0.02	71.3	65.5
ROA		30	0.06	59.9	53.7
BJC	East Tuolumne	23	0.04	54.5	32.3
RRC92		8	0.02	67.3	60.7
RRC00		25	0.05	63.6	57.0
CUR92	North Tuolumne	4	0.01	95.8	94.9
CUR00		23	0.04	45.0	42.6
WOC92		31	0.06	68.6	66.2
WOC00		21	0.04	63.6	63.1
BBC	South Tuolumne	23	0.04	67.9	56.0
FIR		23	0.04	61.4	50.4
HAT		30	0.06	67.1	50.3
SEC		24	0.05	72.0	60.0
AND	Stanislaus	30	0.06	61.2	50.4
PEO		30	0.06	93.2	86.8
ROS92		28	0.05	66.3	58.0
ROS00		10	0.02	64.2	56.0
Total		519		68.2	60.2

closely to the remaining Tuolumne populations except for Hatch Creek. In addition, samples from the Stanislaus River drainage (Andrew, Peoria, and Rose), were intermediate in Burnaby PCA scores between the Red Hills and all of the Tuolumne populations (South, East, and North). Of particular interest is the Andrew's and Peoria Creek populations which occur in the Red Hills regions (Figure 1) but presently flow into the Stanislaus River. For the Burnaby PCA graphs, Andrew's and Peoria Creek fall well within the multivariate space of the Red Hills populations (Figure 3). The other Stanislaus River population, Rose Creek (ROS), falls within the Burnaby PCA distribution of the Tuolumne River populations.

Canonical discriminant analysis results were generally concordant with the Burnaby analysis (Figure 2). Separation of populations was greater in the canonical analysis but groupings were generally the same as in the Burnaby analysis. The Red Hills populations all grouped very closely together except for Minnow Gulch (MIG) which was slightly different for the reduced dataset (Figure 2b). North and East Tuolumne populations grouped together with three of the South Tuolumne populations (Becca B., First, and Second). The remaining South Tuolumne population (Hatch Creek) was located intermediate the Tuolumne grouping and the Red Hill's grouping. As in the Burnaby analysis, the Andrew's Creek population was more similar based on canonical scores to the Red Hills populations than to the other Stanislaus populations (Peoria and Rose Creeks).

Genetics

A total of 123 individuals were sequenced for 339 base pairs (bp) of the mtDNA NADH-2 subunit corresponding to positions 5170 through 5508 of *Cyprinus carpio* (GenBank X61010). Within the ingroup, there were a total of 17 variant sites with 13 being at the 3rd position, 3 at the 1st position, and 1 at the 2nd position defining 16 haplotypes. Of the 16 haplotypes present in the assayed individuals (Table 5), 3 have

been previously observed in *L. symmetricus* (RHR1 = N16, RHR2 = N17, and RHR3 = N20, Jones and Bernardi submitted). These mutations resulted in five amino acid substitutions corresponding to amino acid positions 86, 87, 124, 148, and 179 in *C. carpio* mtDNA NADH subunit 2 (SWISS PROT P24972). Two of these amino acid changes involved a change in polarity of the amino acid residue (position 86: Met to Ile and position 148: Ala to Thr).

Distribution of haplotypes was highly structured geographically. Within the six Red Hills sampling locations, a total of seven haplotypes were observed (RHR1-7). RHR1 and RHR2 were the most common haplotypes observed in the Six Bit Gulch drainage of the Red Hills (40% for each haplotype). Within the East Tuolumne River samples, there were no unique haplotypes. The two haplotypes present (BBC1 and TUOL3) were also found throughout the North Tuolumne, South Tuolumne, and Stanislaus sampling localities. However, the North Tuolumne individuals possessed haplotypes that were not found in any other sampling region (TUOL2, TUOL4, TUOL5, and TUOL6). The Southern Tuolumne region possessed a unique haplotype (BBC2) that was present in all assayed Hatch and Second Creek individuals. The BBC2 haplotype was also present in 40% of the Becca B. Creek individuals. The other haplotype present in Becca B. Creek was the BBC1 haplotype that was present in First Creek, Peoria Creek, Rose Creek, and the two East Tuolumne sites (BJC and RRC). The TUOL1 haplotype comprised the remaining samples of First Creek as well as several individuals from Peoria and Rose Creeks. The AND haplotype was only observed in the Andrew's Creek populations.

Analysis of population differentiation

Population comparison of FST revealed that there has been very little gene flow occurring between the Red Hills populations and neighboring populations (Table 4). Gene flow between Red Hills populations, including Poor Man's Gulch, has remained relatively high being infinite in nearly all of the cases. Only in two instances within the Red Hills is migration not infinite (HOR-PMG: 27.51 and AMB-PMG: 93.84; Table 4). Within the Tuolumne River drainage, significant gene flow between Curtis-Woods and Curtis-Rough and Ready Creeks has occurred (Table 4). Within the South Tuolumne, there are high levels of gene flow between two of the creeks (Hatch and Second Creeks) but not with Hatch-Second Creeks and First Creek (FST = 0.68104 and 0.69711, respectively; Table 4). However, significant gene flow has occurred between First Creek and two of the Stanislaus River sites (Rose and Peoria Creeks).

Analysis of molecular variance (AMOVA) using the five geographic groupings (Table 1) revealed a highly significant Φct (0.34327; $p < 0.0001$). Analysis of only the Red Hills populations suggested no significant partitioning of genetic variation among the upper and the lower sampling sites in the Six Bit drainage (Horton, Amber, Roach versus Six Bit, Minnow Gulch, and Poor Man's Gulch) (ΦCT = 0.05355; $p = 0.2747$). However, when the six Red Hills populations are grouped against the creeks directly opposite of the reservoir (Becca B., First, Hatch, and Second Creeks), a significant amount of genetic variance is explained (ΦCT = 0.2442; $p = 0.0068$). Exclusion of Becca B and First Creek from this AMOVA analysis continued to suggest that a significant level of genetic variance exists between the Red Hills populations and Hatch-Second Creeks (ΦCT = 0.41477; $p = 0.0323$).

Phylogenetic analysis

Maximum parsimony analysis of the 339 bp of NADH-2 subunit revealed a total of 3 most parsimonious trees (113 steps, 0.841 CI, 0.860 RI) (Figure 4). Weighting of transversions (TV:TS = 7:1) revealed 9 most parsimonious trees (226 steps, 0.845 CI, 0.802 RI) which were in agreement with the unweighted parsimony topologies. Bootstrap values were generally low but provided support (61–63% bootstrap) for a reciprocally monophyletic Red Hills grouping. Further, all phylogenetic analyses produced a topology consisting of two Red Hills clades (α and β). The Red Hills β clade is defined by a change in amino acid polarity (Met to Ile) (Figure 4). Both Red Hills clades are relatively widespread in the Red Hills populations (Table 5). Maximum-likelihood analysis (see Materials and Methods for criteria) produced a phylogeny (–ln likelihood = 1011.68616) consistent with the parsimony findings (Figure 4).

Discussion

Morphology as an ecophenotypic response

One of the primary goals for this study was to determine if there was evidence that the distinct morphology of the Red Hills *L. symmetricus* (as

Table 4. Pairwise estimates of population subdivision (Fst; below diagonal) and gene flow (*Nfm*; above diagonal)

	ROA	AMB	MIG	SBG	HOR	PMG	WOC	SUL	CUR	RRC	BJC	BBC	HAT	SEC	FIR	PEO	ROS	AND
ROA	—	inf	inf	inf	inf	inf	0.2933	0.1121	0.5435	0.5738	0.0750	0.6751	0.4685	0.3931	0.2580	0.1400	0.2442	0.0758
AMB	0.0000	—	inf	inf	inf	93.8396	0.4066	0.2745	0.6250	0.7020	0.2299	0.7801	0.7998	0.7540	0.3420	0.2895	0.3752	0.1664
MIG	0.0000	0.0000	—	inf	inf	inf	0.3488	0.1071	0.7201	0.7580	0.0542	0.7729	0.4914	0.4014	0.2900	0.1517	0.2931	0.0729
SBG	0.0000	0.0000	0.0000	—	inf	Inf	0.2933	0.1121	0.5435	0.5738	0.0750	0.6751	0.4685	0.3931	0.2580	0.1400	0.2442	0.0758
HOR	0.0000	0.0000	0.0000	0.0000	—	27.5112	0.5446	0.4214	0.7540	0.8613	0.3770	1.0051	1.6108	1.4149	0.4378	0.4052	0.4804	0.2463
PMG	0.0000	0.0053	0.0000	0.0000	0.0179	—	0.3111	0.1520	0.5296	0.5508	0.1174	0.6213	0.4533	0.3977	0.2626	0.1642	0.2576	0.0972
WOC	0.6303	0.5515	0.5890	0.6303	0.4787	0.6164	—	9.3756	inf	2.8029	0.1923	0.5193	0.2551	0.2300	0.3285	0.2489	0.3571	0.1508
SUL	0.8168	0.6456	0.8235	0.8168	0.5427	0.7668	0.0506	—	8.0763	1.0452	0.0184	0.3526	0.1656	0.1380	0.1924	0.0958	0.1816	0.0473
CUR	0.4792	0.4444	0.4098	0.4792	0.3987	0.4856	0.0000	0.0583	—	inf	0.6000	0.9457	0.3684	0.3372	0.6559	0.5714	0.8478	0.2917
RRC	0.4657	0.4160	0.3975	0.4657	0.3673	0.4758	0.1514	0.3236	0.0000	—	1.8333	2.2059	0.3962	0.3567	1.6641	1.2931	2.8029	0.3421
BJC	0.8696	0.6850	0.9023	0.8696	0.5701	0.8099	0.7222	0.9646	0.4546	0.2143	—	1.2209	0.1287	0.1062	2.2360	0.4000	2.4999	0.0500
BBC	0.4255	0.3906	0.3928	0.4255	0.3322	0.4459	0.4905	0.5864	0.3459	0.1848	0.2905	—	0.7293	0.6821	1.4514	1.3380	1.8483	0.3382
HAT	0.5163	0.3847	0.5043	0.5163	0.2369	0.5245	0.6622	0.7512	0.5758	0.5579	0.7953	0.4067	—	inf	0.2342	0.1798	0.2281	0.1104
SEC	0.5598	0.3987	0.5547	0.5598	0.2611	0.5570	0.6850	0.7837	0.5972	0.5837	0.8248	0.4230	0.0000	—	0.2172	0.1587	0.2074	0.0954
FIR	0.6596	0.5938	0.6329	0.6596	0.5332	0.6557	0.6035	0.7221	0.4326	0.2310	0.1828	0.2562	0.6810	0.6971	—	14.7207	inf	0.2478
PEO	0.7813	0.6333	0.7672	0.7813	0.5524	0.7528	0.6676	0.8392	0.4667	0.2789	0.5556	0.2720	0.7355	0.7590	0.0329	—	inf	0.1260
ROS	0.6719	0.5713	0.6305	0.6719	0.5100	0.6600	0.5833	0.7336	0.3710	0.1514	0.1667	0.2129	0.6867	0.7068	0.0000	0.0000	—	0.2230
AND	0.8684	0.7503	0.8728	0.8684	0.6700	0.8372	0.7683	0.9136	0.6316	0.5938	0.9091	0.5966	0.8191	0.8398	0.6687	0.7987	0.6916	—

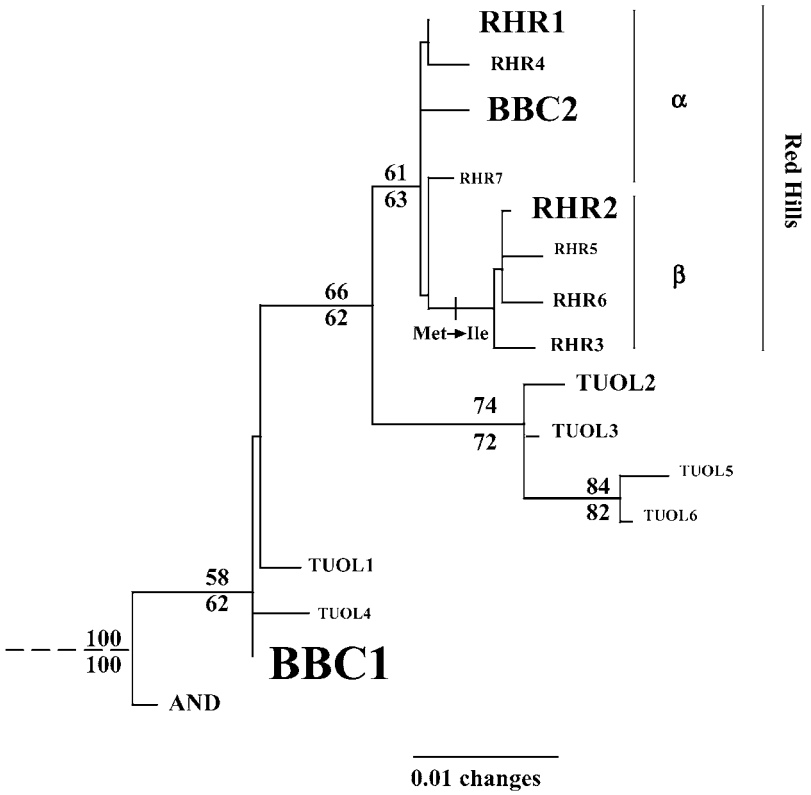

Figure 4. Maximum likelihood tree of the 17 ingroup haplotypes. Rooting produced by outgroups is indicated by the dashed line. Bootstrap values (500 replicates) for equal weighting and weighted (TV:TS = 7:1) parsimony are indicated above and below the branch, respectively. Font size of haplotype name indicates relative abundance of a given haplotype in the overall total. See Table 5 for specific values and geographic distribution of haplotypes. The two Red Hills groups (α and β) are indicated along with the phylogenetic location of the amino acid change that defines the Red Hills β group.

represented by Horton Creek in Brown et al. (1992)) was a product of the unusual environment of the Red Hills region (serpentine rock). In this study, eight populations were sampled in areas considered part of the Red Hills formation (shaded areas in Figure 1). One of these populations (Peoria Creek) occurs in a disjunct outcropping of the Red Hills separated from the larger Red Hills region (Figure 1) by Table Mountain. Both the Burnaby and the Canonical discriminant analysis suggested that the Andrew's Creek population, although physically closest to other Stanislaus River populations, appears morphologically very similar to the type locality of the Red Hills (Horton Creek) and other sites sampled in the Six Bit Gulch drainage as opposed to the two other sampled Stanislaus River sites (Peoria and Rose Creeks).

The phylogenetic analysis of the mtDNA sequence data suggested that the Andrew's Creek population has low haplotype diversity and is genetically distinct from all other populations sampled including Peoria and Rose Creeks which shares haplotypes with Big Jackass, Rough and Ready, Woods, Becca B., and First Creeks. In all cases, phylogenetic analysis suggested that Andrew's Creek was basal to all other sampled populations in the Stanislaus and Tuolumne Rivers. Additional sampling in the Stanislaus River drainage (especially near Andrew's Creek) is required to determine the distribution of the observed haplotypes (AND, BBC1, and TUOL1; Table 5). Taken together, the morphological, genetic, and geographic sampling suggest that the Red Hills morphology may be in large part a factor of the Red Hills environment. From the distribution of haplotypes and phylogenetic analysis, there does not appear to have been any recent gene flow or migration between *L. symmetricus* in the Six Bit Gulch portion of the Red Hills and populations in the Stanislaus River drainage (including Andrew's Creek).

Table 5. Haplotype distribution among populations of *L. symmetricus*

Haplotype	AMB	HOR	MIG	PMG	ROA	SBG	BJC	RRC	CUR	SUL	WOC	BBC	FIR	HAT	SEC	AND	PEO	ROS	Total
RHR1	3	6	1	2	2	2													**16**
RHR2	3	3	2	3	2	3													**16**
RHR3				1	1														**2**
RHR4	1	1																	**2**
RHR5	1																		**1**
RHR6		2																	**2**
RHR7		1																	**1**
AND																5	**5**		**10**
BBC1							5	3			1	9	6				3	4	**31**
BBC2												6		10	10				**26**
TUOL1													4				2	1	**7**
TUOL2									2	1									**3**
TUOL3								2	1	2	1								**6**
TUOL4									2										**2**
TUOL5											1								**1**
TUOL6											2								**2**
	8	13	3	6	5	5	5	5	5	3	5	15	10	10	10	5	5	5	**123**

Morphological differences consistent between years

A second goal of this study was to determine if the morphological distinctiveness of the Red Hill *L. symmetricus* was consistent between the two sampling periods (Brown et al. 1992 vs. this study). One of the first problems encountered in comparing the two datasets was the difference in canonical discriminant scores. Although Burnaby multivariate analysis did not indicate any substantial difference between years (except for Horton Creek), canonical analysis showed a clear grouping of populations according to sampling period. Two possible reasons can explain the discrepancy in the canonical analysis.

First, measurement error is the primary concern that would cause a discrepancy. However, the difference between years for a given site is consistent for all three of the comparable sites in this paper (positive shift in Canonical 1 but none in Canonical 2). In addition, two of the sites (ROS92 and WOC92) were measured by one person (W. Bennett from Brown et al. 1992) while all morphological measurements for this study (i.e., ROS00, WOC00, and HOR00) as well as HOR92 were measured by B. Quelvog. The alternative explanation for the difference in between-year canonical scores is that populations in the sampling region are undergoing slight, but significant changes in morphology over short periods of time. Of particular interest is that the primary characters explaining the difference between years was a change in caudal peduncle length (PL) and body depth (BD) (data not shown). Although the three comparable populations represented geographically diverse sampling localities (Rose Creek – Stanislaus River, Horton Creek – Red Hills, and Woods Creek – North Tuolumne), the change in the Canonical 1 score was consistent for all populations suggesting that similar changes happened to all three populations between the two sampling periods.

Variation in the chisel lip character and frequency between years was low for two of the three compared populations (Woods and Rose Creeks). However, Horton Creek *L. symmetricus* showed a dramatic shift to lower occurrence of the chisel lip character as well as a shift in the distribution of character states. Similarly, the Horton Creek population showed a significant difference in 4 of the 15 morphological characters. One of these characters, interorbital distance (ID), showed a highly significant difference between sampling periods suggesting that the morphology of the Horton Creek population may be changing rapidly in response to environmental changes. The environment of Horton Creek as well as many other of the Red Hills populations is extremely harsh. Oxygen levels and temperature may reach levels intolerable to most other aquatic organisms (1–2 ppm and 30–

35 °C; Moyle et al. 1995). Although the amount of time between these two studies is short (4–7 years), sufficient time for several generations of *L. symmetricus* has passed (time to maturity ca. 2 years; Moyle 1976).

Genetic distinctiveness of the Red Hills L. symmetricus

The results from the phylogenetic analysis of the Red Hills *L. symmetricus* and neighboring populations clearly indicates that the Red Hills *L. symmetricus* is a genetically distinct and identifiable entity. The theoretical time to reciprocal monophyly for mitochondrial DNA is *Nef* generations where *Nef* is the effective number of females (Avise 1994). Assuming an equal sex ratio and a historical census size similar to the current census size (ca. 200–500 individuals), the expected time to reciprocal monophyly would be 100–250 generations. Given the age of the Don Pedro Reservoir (80 years) and the generation time of approximately 2 years for *L. symmetricus* (Moyle 1976), the Red Hills *L. symmetricus* appears to have been separated from neighboring populations well before dam construction. However, the construction of the Don Pedro dam may have created a barrier (the large body of water in the reservoir) which separated the Six Bit Gulch populations of the Red Hills from closely related populations in Hatch and Second Creeks.

Further sampling in unnamed creeks bordering the Hatch-First-Second Creeks needs to be performed to determine if any RHR haplotypes occur outside of the Six Bit Gulch region. Of particular relevance to the Red Hills *L. symmetricus* is the phylogenetic position of the Hatch and Second Creek individuals whose haplotype (BBC2) fell within the Red Hills clade. In addition, Becca B Creek possessed this same haplotype as well as the BBC1 haplotype which was found throughout much of the Tuolumne and Stanislaus sampling localities. The occurrence of these two divergent haplotypes in the Becca B sampling locality suggests a possible secondary connection. In fact, the Hetch Hetchy Aqueduct (underground) flows from near the Big Jackass Creek locality (site 7, Figure 1) past the Becca B. Creek site northward. Although transfer of individuals via the aqueduct is unlikely, the possibility exists. An alternative explanation is that the BBC1 and BBC2 haplotypes were once widespread in the South Tuolumne region, but populations due to random genetic drift went to fixation for either BBC1 (i.e., First Creek) or BBC2 (Hatch and Second Creeks). Unlike other South Tuolumne populations, the Becca B. Creek population maintained a large effective population size which prevented either BBC1 or BBC2 from going to fixation. Additional sampling in Becca B. Creek and other South Tuolumne sites needs to be performed to determine the distribution of the BBC1 and BBC2 haplotypes.

In addition to being reciprocally monophyletic for the mtDNA sequences, the Red Hills *L. symmetricus* contains as many haplotypes as neighboring populations in the combined North and East Tuolumne populations (7 versus 6). From this data, there does not appear to have been a pronounced bottleneck in effective population size in the past which led to a marked decrease in haplotypic diversity. If any Red Hills populations have undergone a bottleneck, Hatch and Second Creeks would be the most likely candidates. However, this relatively high haplotypic diversity does not imply that the Red Hills *L. symmetricus* populations are safe from future threats. Loss of habitat due to development and introduction of nonnatives species such as green sunfish (*Lepomis cyanellus*) and western mosquitofish (*Gambusia affinis*) are an impending threat for *L. symmetricus* through competition and predation. By preserving the upper portions of Six Bit Gulch, much of the community of rare and endemic plants as well as a large population of Red Hills *L. symmetricus* can be maintained.

Conservation status of the Red Hills L. symmetricus

One of the problems of studying newly discovered, rare populations is that there is limited information regarding the taxon under investigation. An additional problem is that often the populations are discovered in a state at which protection of habitat or population sizes with high levels of genetic variation is impossible due to loss of habitat and corresponding groups of individuals. In this study, we have used combined morphological information from two sampling periods along with phylogenetic analysis to provide concordant evidence (sensu Grady and Quattro 1999) that the Red Hills *L. symmetricus* should be recognized as a distinct entity. We propose that the Red Hills *L. symmetricus* be considered a new subspecies given the endemic nature and its distinct morphological and genetic characteristics. Description of this subspecies will be done in a subsequent paper (Jones in preparation). Likewise, we propose that the Red Hills roach be afforded appropriate state

and federal protection before an unrecoverable loss of habitat or genetic diversity occurs.

Acknowledgements

Funding for this research was provided by the Genetic Resource Conservation Program, American Museum of Natural History (Theodore Roosevelt Fund), ACCESS, UCSC Biology Department, American Society of Ichthyologists and Herpetologists (Raney Award), Myers Oceanographic Trust, Friends of the Long Marine Lab, the David Gaines Memorial Award, and a NSF Dissertation Improvement Grant (WJJ). Samples were collected under California Department of Fish and Game Scientific Collecting Permits 803053-02, 803036-03, 803026-05, and Memorandum of Understanding following the criteria of the Chancellor's Animal Research Committee (CARC). Thanks to N. MacCleod for allowing use of his computer program, Burnaby. Special thanks to R. Vrijenhoek for information regarding the JMP software. M. Myers is acknowledged for help in data collection. T. Dowling provided primers and outgroup tissue crucial to this study. Thanks to L. Brown, G. Corrigan, T. Dowling, P. Moyle, and M. Ramon for comments on the manuscript. Sequences described in this manuscript have been submitted to GenBank (accession numbers AF392997-AF393011).

References

Avise JC (1975) *Protein Divergence and Speciation in California Minnows*. PhD thesis, University of California, Davis.

Avise JC (1994) *Molecular Markers, Natural History and Evolution*. Chapman and Hall, New York.

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

Burnaby TP (1966) Growth-invariant discriminant functions and generalized distances. *Biometrics*, **22**, 96–110.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, **39**, 783–791.

Grady JM, Quattro JM (1999) Using character concordance to define taxonomic and conservation units. *Conserv. Biol.*, **13**, 1004–1007.

JMP Statistics and Graphics Guide (1994) Version 3; SAS Institute Inc. Cary N.C., USA.

Jones WJ, Bernardi G. Phylogenetic relationships among *L. symmetricus* and *L. exilicauda* subspecies. Submitted to Molecular Phylogenetics and Evolution.

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.*, **26**, 24–33.

Maddison WP, Maddison DR (1992) MacClade 3.05. Sinauer Associates, Sunderland, MA.

Moyle PB (1976) *Inland Fishes of California*. 1st edn., University of California Press, Berkeley.

Moyle PB (1996) Potential aquatic diversity management areas. Ch. 57 In: *Sierra Nevada Ecosystem Project: Final Report to Congress, vol. II, Assessments and Scientific Basis for Management Options*. Davis: University of California, Centers for Water and Wildland Resources.

Moyle PB, Cech JJ (2001) *Fishes: An Introduction to Ichthyology*, 4th edn. Prentice Hall, Upper Saddle River, NJ.

Moyle PB, William JE, Wikramanayake ED (1995) Fish species of special concern of California (2nd edn.). Final Report, Contract No. 21281F.

Murphy GI (1943) Sexual dimorphism in the minnows *Hesperoleucus* and *Rhinichthys*. *Copeia*, **1943**, 187–188.

Page RDM, Holmes EC (1998) *Molecular Evolution – A Phylogenetic Approach*. Blackwell Science, Massachusetts.

Posada D, Crandall KR (1998) MODELTEST: testing the model of DNA substitution. *Bioinform.*, **14**, 817–818.

Power ME (1990) Effects of fish in river food webs. *Science*, **250**, 811–814.

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

Rohlf FJ, Bookstein FL (1987) A comment on shearing as a method for 'size correction'. *Syst. Zool.*, **36**, 356–367.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN ver 2.000: a software for population genetics data analysis. Department of Anthropology and Ecology, University of Geneva, Switzerland.

Swofford DL (1993) PAUP, phylogenetic analysis using parsimony, version 3.1. Illinois Natural History Survey, Champaign.

Swofford DL (1998) PAUP∗, phylogenetic analysis using parsimony, version 4.0. Illinois Natural History Survey, Champaign.