# Phylogenetics of Freshwater Sculpin

**Charlene Emerson**

**Introduction.** Reconciling observed genetic variation with evolutionary history is often a daunting task. Phylogenetics is a method of study that is particularly well-suited to addressing this task by combining the measureable genetic variability of individuals, populations, or species with proposed evolutionary relationships and processes (Barraclough & Nee 2001). Historically, phylogenetic trees have been generated using a single locus approach, where variation within a single gene is used to create a phylogenetic gene tree. However, this single gene method may fail to account for the full variation across a genome.

Shortcomings of the single locus approach are most apparent in complex organisms, such as the freshwater sculpin (genus *Cottus*). Freshwater sculpin exhibit highly ambiguous morphology, as well as wide and overlapping distributions – resulting in considerable difficulty in species classification (Moyle 2002). Analyzing multiple genetic loci in such species, instead of a single locus, may provide a fuller picture of species variation.

I propose to use freshwater sculpin as a model to compare this modern multiple loci approach with the single locus approach. By comparing a species tree, generated by concatenating multiple nuclear markers, with the single gene trees, I hope to achieve greater understanding of evolutionary relationships within seven Eastern Pacific *Cottus* species, as well as discerning the advantages and disadvantages of the single locus and multiple loci approaches.

*Phylogenetics.* Before comparing the relative merits of these different approaches, it is important to understand the overall phylogenetic method. Phylogenetics compares the similarities and differences between homologous sequences of genetic material, which are sequences derived from a common ancestor (Scotland 2010). Closely related individuals, populations, or species should have fewer differences in the homologous sequence when compared to the genes of distantly related species (Lemey, Salemi & Vandamme 2009). These differences in the genetic sequence arise from the actions of several evolutionary mechanisms: mutation, natural selection, genetic drift, and gene flow (Hartl 1981). Genomic differences arising from these mechanisms are often the basis for phylogenetic analysis (Davis & Nixon 1992). In the presence of these processes, frequency of genetic alleles[1] will vary and a population will experience evolutionary change (Nei, Maruyama & Chakraborty 1975).

Mutations, or changes in genomic sequence which often occur spontaneously, are propagated or eliminated by the action of natural selection. If a mutation is deleterious to a species, natural selection causes it to be eliminated from the population. If the mutation is favorable, natural selection will allow it to become established in a population. However, a harmful mutation may propagate by the action of genetic drift. Genetic drift is the changing of allelic frequencies due to random occurrences. Effects of genetic drift are more marked in small populations, since deleterious alleles can become more easily fixed within a smaller population (Frankham 2005). Genetic drift and natural selection are generally considered to increase genetic differences between populations, eventually leading to speciation[2].

[1] Alleles are different forms of a single gene. Different alleles may create different traits in the organism in which they are carried.

[2] Speciation is the evolutionary process by which new species are formed.

In contrast, gene flow is generally considered to decrease the occurrence of speciation (Via 1999, Porter & Johnson 2007). Gene flow is the transfer of genetic alleles between populations of a species. Isolated populations experience decreased gene flow, because fewer individuals immigrate to or emigrate from the population to exchange genes. Over time, gene flow may cease and the isolated population could become reproductively isolated from the rest of their species. The absence of gene flow is considered to be a major mechanism of evolution. Together, these four major processes change the allele frequencies of a population, altering the relative relatedness of populations and species. Studying this evolutionary relatedness in the context of genetic data is one of the major applications of phylogenetics.

*Phylogenetic Trees*. The phylogenetic tree is a major tool of phylogenetic analysis. One of the most basic methods of tree-making uses a single homologous sequence – a single locus tree, and compares this sequence across members of a group, within or between populations, or between closely related species. These single locus trees can then be used to determine the overall evolutionary relatedness between members of the sample group. However, an analysis that stops with examination of a single homologous sequence would be extremely limited in scope. Limiting analysis to a single locus tree will only represent the events occurring at that gene locus, and may not necessarily represent the entire organism or species.

Analysis of multiple genetic loci is vital for a more complete understanding of evolutionary relationships, especially when considering that differing homologous sequences are subjected to differing evolutionary pressures (Brito & Edwards 2009). Because each homologous sequence changes differently, depending on evolutionary pressures exerted on that particular locus, analysis of individual sequences will create different trees for the same group of organisms (Avise 2000). For example, a gene for eye color will be selected upon very differently than a gene for limb structure, generating different phylogenetic trees. By combining, or concatenating, multiple loci into an overall species tree – rather than a single locus tree, the most accurate representation of evolutionary relationships can be determined.

Generally, phylogenetic trees are constructed using a cladistic approach, where it is assumed that members of the tree share a common evolutionary history (NCBI 2004). The cladistic approach groups members of the tree by shared common ancestry, with members of the tree gradually diverging into individual groupings, or clades. Phylogenetic trees commonly take two forms: cladograms or phylograms. Cladograms only show the order of the branching changes of the homologous sequence, whereas in phylograms, the length of each branch of the tree corresponds to the number of changes that have occurred in the sequence (Hall 2004, Lemey et al. 2009).

*Determining Accuracy of the Phylogenetic Diagram*. Phylogenetic tree diagrams can also vary according to the statistical analysis used to generate the tree. Differences in genomic sequences can vary as a result of a number of processes, making analyses of sequence changes complex. Components of the sequence cannot always be assumed to change with equal likelihood. For example, genetic changes that result in the production of a similarly functioning protein would seem more likely than changes that completely negate a protein's function.

Additionally, multiple changes may happen at a site, with no way of knowing the total number of mutation occurrences. Changes may also occur that later revert back to the original sequence, causing organisms to

seem more evolutionarily related than they actually are (Hall 2004, Lemey et al. 2009). When multiple changes occur at a nucleotide site, so that the sequence is no longer informative about true evolutionary relationships, it is called substitution saturation (Lemey et al. 2009).

Substitution rates of nucleotide sites are modeled by several mathematical formulas, with each model representing a different relative rate of change. Models are chosen for a specific data set using statistical selection software (Posada 2008). Among the most popular are the MODELTEST and JMODELTEST. Model tests analyze the nucleotide sequences in the data set and select the model of nucleotide substitution that best fits the existing data.

The simplest model of nucleotide substitution is the Jukes-Cantor model, commonly called JC69 (Jukes & Cantor 1969). JC69 assumes equal nucleotide base frequencies and equal mutation rates for each adenine, thymine, cytosine, or guanine[3] in the genetic code. Models become more complex as other assumptions are made about the base sequence and substitution rate. Felsenstein (1981) created a model (F81) which assumes unequal nucleotide base frequencies. The Kimura model (K80) assigns different likelihoods to transitions between the purines – adenine and guanine, and pyrimidines[4] – cytosine and thymine (Kimura 1980). Subsequent models, including HKY85, TN93, and GTR, continued to vary and combine assumptions about base sequence and substitution rate (Tavare 1986, Tamura & Nei 1993).

After the designation of a nucleotide substitution model, sequence data is used to construct the phylogenetic tree. Several different methods have been developed to construct trees (Lemey et al. 2009, Rosenberg & Kumar 2001). Among these methods are the Maximum Likelihood (ML) and Bayesian analyses. Both are considered to be "discrete character" methods, where compared homologous sequences are aligned – each position is considered to be a "character" and the nucleotide in that position is a "state" (Lemey et al. 2009). Character-states are analyzed independently to determine relatedness between samples.

Maximum Likelihood examines different possible tree formations and searches for the most likely tree, according to a particular evolutionary model. Likelihood of possible trees is calculated, according to an algorithm, and the most likely tree is selected. A ML generated tree can be supported by the use of bootstrap resampling. Bootstrapping takes a subsample of character-states and creates a tree based upon this subsample. The bootstrapping process is replicated numerous times, providing support for the final chosen tree. In contrast, a Bayesian analysis does not search for a single best tree. Bayesian analysis targets a distribution of possible trees by using prior probabilities. After analyzing possible tree formations, a consensus tree is created based on the highest posterior probabilities of each branch or node. (Hall 2004, Lemey et al. 2009).

*Freshwater Sculpin.* Structuring evolutionary processes into a phylogenetic tree has applications across a number of biological fields, including: molecular biology, evolution and development, epidemiology, ecology and conservation biology (NCBI 2004). A particularly engaging application of phylogenetic methods is the use of trees to delineate the distribution and genetic variation of cryptic species. Cryptic species, by definition, are notoriously difficult to classify, usually because of ambiguous morphological features. In some cases, understanding the

[3] See purines and pyrimidines in 4.

[4] Purines and pyrimidines are two groups of nitrogenous bases that are part of the variable components of DNA. Adenine and guanine are purines, while cytosine and thymine are pyrimidines.

distribution and improving classification of these species can lead to insight about important historical evolutionary and geographic processes, as well as potential human environmental impacts.

One such cryptic group of species, freshwater sculpin (genus *Cottus*), are an ideal candidate for study of species variation and distribution. Sculpin exist in a variety of habitats, with many freshwater sculpin species inhabiting various inland rivers and streams, freshwater lakes, and brackish coastal waters (McGinnis 1984, Moyle 2002). Freshwater sculpin have value as a potential *indicator species,* a species that characterizes certain environmental conditions in a particular ecosystem. Indicator species can provide valuable information about a wide range of factors contributing to species distribution, including geographic, environmental and anthropogenic effects (Lindenmayer, Margules & Botkin 1990, Noss 1990).

Utilization of indicator species is an especially pertinent strategy in California, which is the site of one of the most complex water storage and transport systems in the world (McClurg 2000). California's water system necessarily impacts natural environmental conditions, with implications that are difficult to measure fully. Discerning the value of freshwater sculpin as an indicator species in California waterways would offer more headway in measuring these effects. *Understanding these environmental impacts is vital to maintaining a water system that is both efficient and sustainable.*

Distribution of freshwater sculpin species is poorly understood at this point in time, with recent genetic studies conflicting with earlier morphological definitions -- which were ambiguous enough to begin with (Baumsteiger unpublished, Kinziger, Wood & Neely 2005, Moyle 2002). To date, there are no studies using nuclear DNA markers to analyze California freshwater sculpin. By constructing phylogenetic trees using the nuclear DNA of freshwater sculpin, new and valuable information about the ecosystems of California waterways may become available. Understanding the evolutionary progress of this genus in California will illuminate the factors – man-made or natural – contributing to its distribution in California waters.

**Method**. Eighteen freshwater sculpin DNA samples were used to conduct the study. Samples were either collected directly from the field or provided from established museum collections. Species and locations included 2 prickly sculpin *(Cottus asper)* from the San Joaquin River, 2 riffle (*C. gulosus)* sculpin from the Sacramento River, 2 prickly sculpin from the Smith River, 2 riffle sculpin from the Kings River, 1 rough sculpin (*C. asperrimus*), 1 riffle sculpin from Washington, 1 reticulate sculpin (*C. perplexus*), 2 riffle sculpin from the Russian River, 2 prickly sculpin from Clear Lake, 1 coastrange sculpin (*C. aleuticus*), 1 mottled sculpin (*C. bairdi*), and 1 margined sculpin (*C. marginatus*) (see Figure 9). All individuals came from California waters, except for the Washington riffle sculpin, mottled sculpin, margined sculpin, and reticulate sculpin samples.

Ten previously identified nuclear markers were developed from a 454 sequencing run of prickly sculpin DNA (Baumsteiger unpublished data). Using the software MSATCOMMANDER (Faircloth 2008), markers were selected for: conservation across all sample species, sequence variability between species, repeatability, and an overall length of between 400 and 500 base pairs. Each marker was optimized for a polymerase chain reaction (PCR) to amplify the desired DNA. Once amplified, the DNA was analyzed through gel electrophoresis[5] to compare amplification. Amplified products were

[5] Gel electrophoresis is a method of separating DNA fragments in order to analyze them.

further purified with ExoSap[6] and submitted for sequencing on an ABI 3730 Sequencer at UCLA's Genome Core facility.

Sequences were compared and aligned in several programs. First, sequences were aligned by forward and reverse alignment of each sample, to ensure that complete and readable sequence was present for analysis in SEQUENCHER (Gene Codes Corp., Ann Arbor, MI). After forming a contiguous sequence for each sample, sequences were then aligned according to nuclear marker and trimmed to uniform length with MUSCLE (Edgar 2004) or MEGA (Kumar, Dudley, Nei & Tamura 2008). Alignment by nuclear marker confirmed that the markers chosen were homologous sequences for each of the freshwater sculpin samples.

Once each marker was aligned for all 18 individual samples, JMODELTEST (Posada 2008) was used to determine the best nuclear substitution models for each locus. All ten loci were then concatenated at each sample to create one continuous nuclear sequence using FASconCAT (Kuck and Meusemann 2010).

The concatenated homologous sequence, created from these 10 markers was used to create the multi-locus phylogenetic species tree. Single-gene trees were generated from each of the nuclear markers as well. Both approaches used Maximum Likelihood analysis in PHYML (Guindon & Gascuel 2003) and Bayesian analysis in MRBAYES (Huelsenbeck & Ronquist 2001) to generate the single gene tree and the multi-locus species tree. For all trees, the mottled sculpin (*C.* bairdi) was used as an outgroup[7] for tree rooting based on previous phylogenetic analyses (Kinziger and Wood 2005).

**Results.** After completion of the JMODELTEST for each marker, JC69 was the model chosen for markers 502, 505, 507, 518, and 520, with HKY chosen for markers 508 and 514, F81 chosen for marker 510 and K80 for marker 516. Finally, TN93 was used for marker 517 and the overall concatenated sequence. If usable sequence could not be generated for a sample, that sample was omitted from the single gene tree. The majority of the markers did not show major differences between Bayesian and Maximum Likelihood analyses.

Only two of the ten nuclear markers generated trees with major differences when created in PHYML compared to trees created in MRBAYES (shown in Figure 1, next page). For all trees, support values for branch separation were generated as measureable evidence for proposed relationships. In Maximum Likelihood trees, branch value bootstraps closer to 100 indicate greater support for a given separation. In Bayesian trees, branch value probabilities approaching 1.0 indicate strongly supported separations.

Several trends emerged when comparing the single gene trees. All ten trees showed recent common ancestry between riffle and prickly sculpin. Nine trees showed the coast-range sculpin as the most distantly related species, other than the root defining mottled sculpin. Margined, reticulate, Washington riffle, and rough sculpin were also frequently shown to have more distant ancestry to the prickly and California riffle groups. Seven trees depicted prickly sculpin in a distinct clade, while five trees showed the California riffle in a distinct clade. None of the 10 trees grouped the Washington riffle with California riffle (shown in Figures 2 and 3).

[6] ExoSap is a common procedure used to sequence PCR products for more accurate sequencing.

[7] In this study, the mottled sculpin was used as an outgroup to provide perspective when comparing the relatedness of the different sculpin species. It is known that the mottled sculpin is distantly related to the samples in the study; however, it is still closely enough related to allow genetic comparisons.

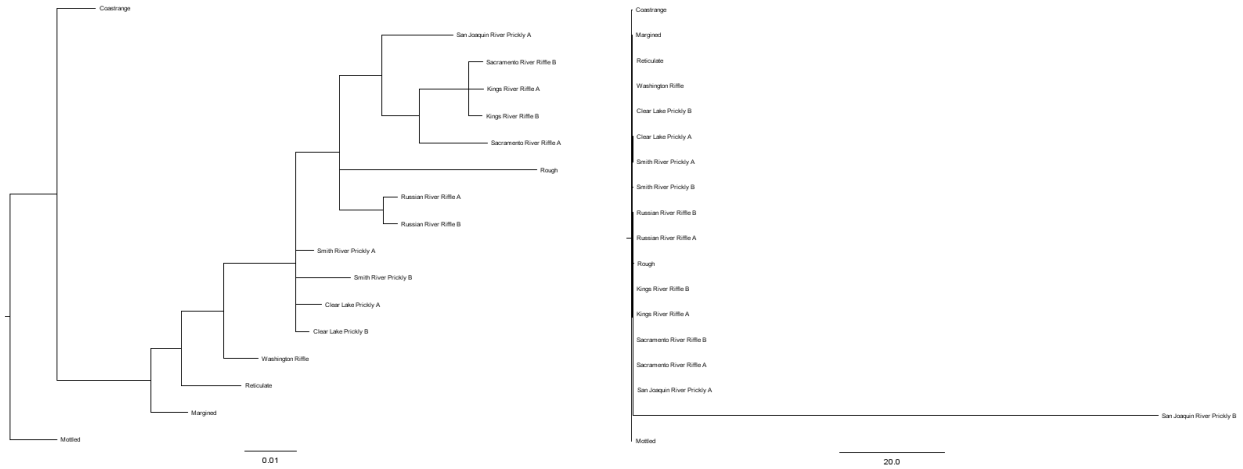


**Figure 1:** Marker 507 generated extremely different trees in Bayesian analysis (left) and in Maximum Likelihood (right)

*riffle sculpin*

*prickly sculpin*

| Marker | Length (bp) | Samples Omitted | Evolutionary Model |
|---|---|---|---|
| 502 | 417 | None | JC69 |
| 505 | 475 | None | JC69 |
| 507 | 493 | San Joaquin River Prickly B | JC69 |
| 508 | 474 | Clear Lake Prickly A | HKY |
| 510 | 489 | None | F81 |
| 514 | 486 | Kings River Riffle A and B, Russian River Riffle A | HKY |
| 516 | 447 | None | K80 |
| 517 | 433 | None | TN93 |
| 518 | 436 | None | JC69 |
| 520 | 435 | None | JC69 |
| All | 4585 | None | TN93 |

Summary of results for genetic markers

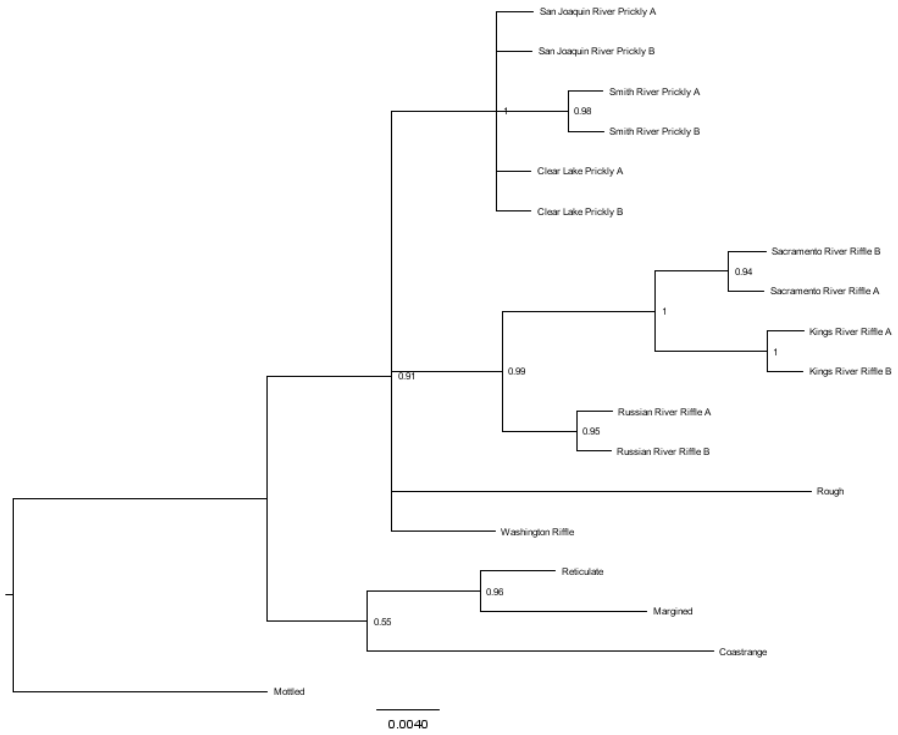


**Figure 2:** Single gene tree for marker 510, created with Bayesian analysis.

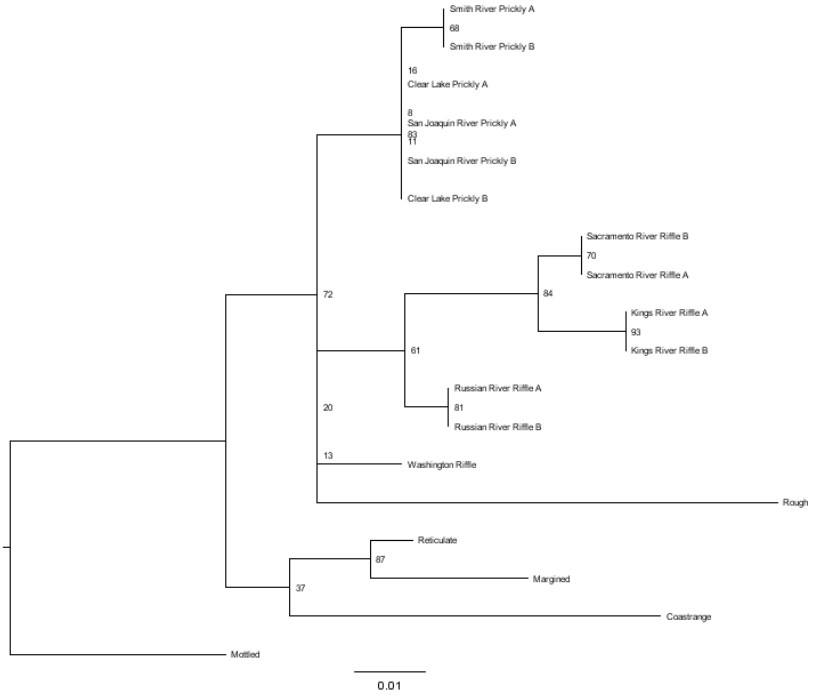


**Figure 3:** Single gene tree for marker 510, created with Maximum Likelihood.

Within proposed riffle sculpin collected in California, the Russian River samples were observed to form their own clade in eight of the single gene trees (Figures 2 and 3). Riffle from the Sacramento River and the Kings River were also observed to diverge into separate clades frequently (Figure 4) – in five of the ten trees. However, support values for this separation varied greatly. Similarly, within the prickly sculpin, Clear Lake prickly formed a separate clade in four trees with strong support values for the separation (Figure 5). However, in three trees the Clear Lake samples did not group together, making conclusions difficult (Figure 6).

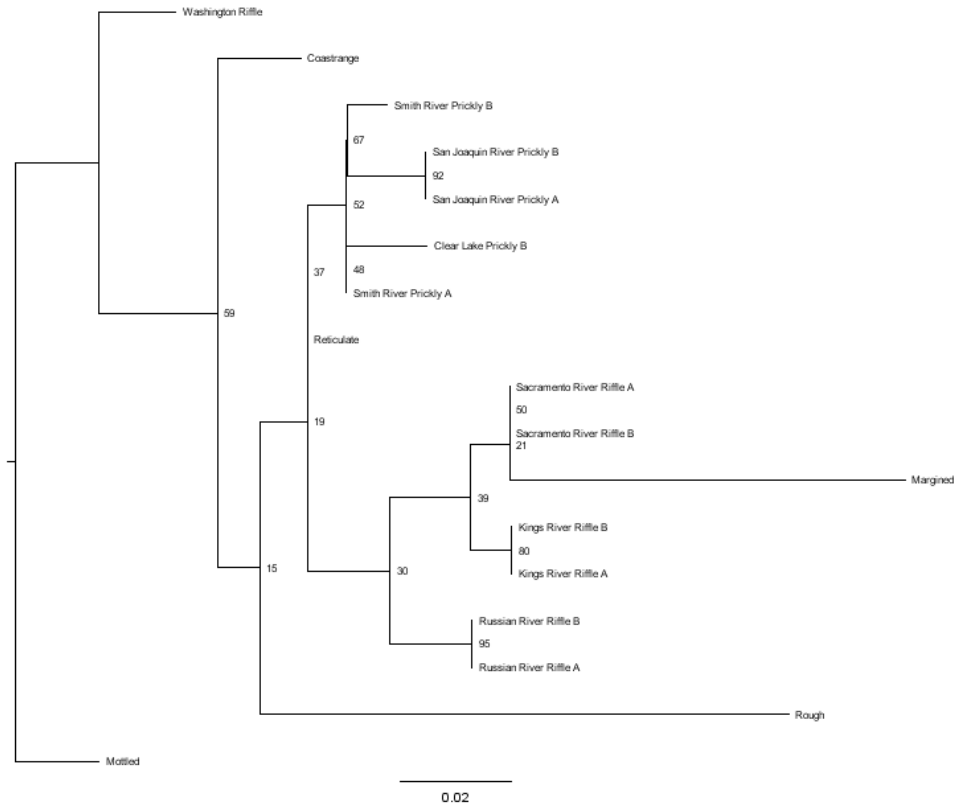


**Figure 4:** Single gene tree for marker 508, created with Maximum Likelihood.

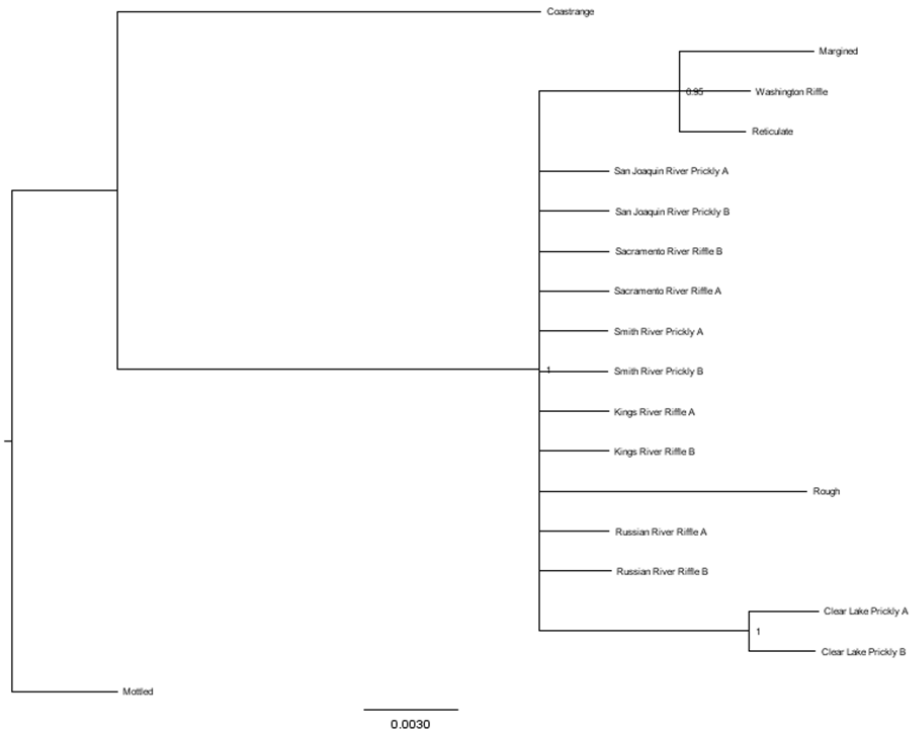


**Figure 5:** Single gene tree for marker 516, created with Bayesian analysis.

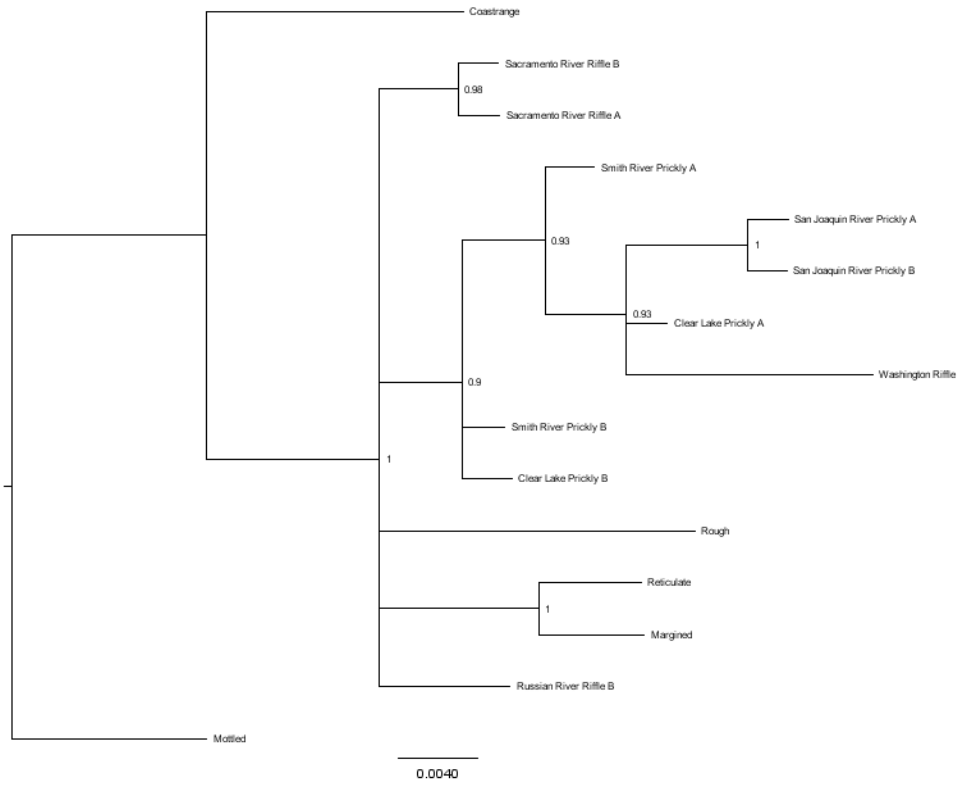


**Figure 6:** Single gene tree for marker 514, created with Bayesian analysis.

Identical multi-locus trees were generated by the Bayesian and Maximum Likelihood methods (shown in Figure 7 and 8). Support values for both trees were strong.

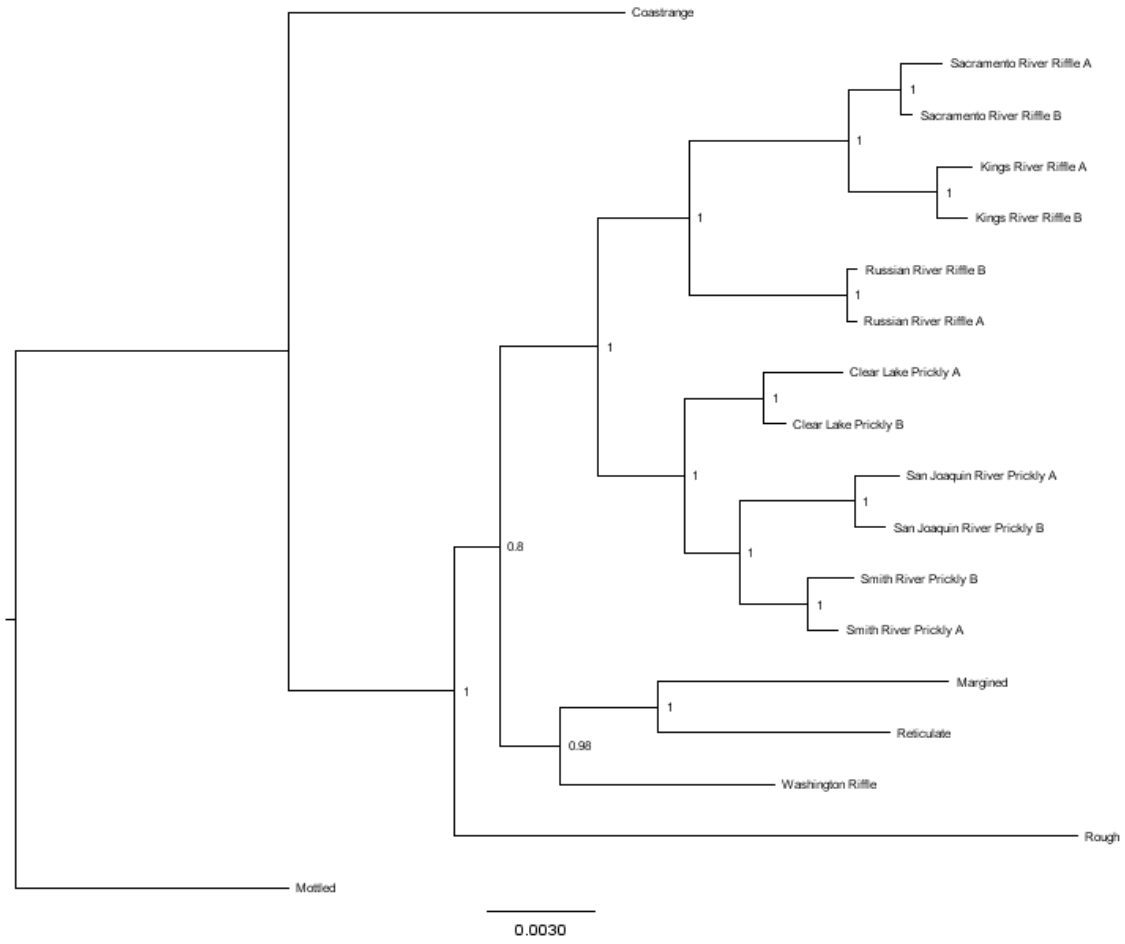


**Figure 7:** Multi-gene tree, generated with partitioning and Bayesian analysis.

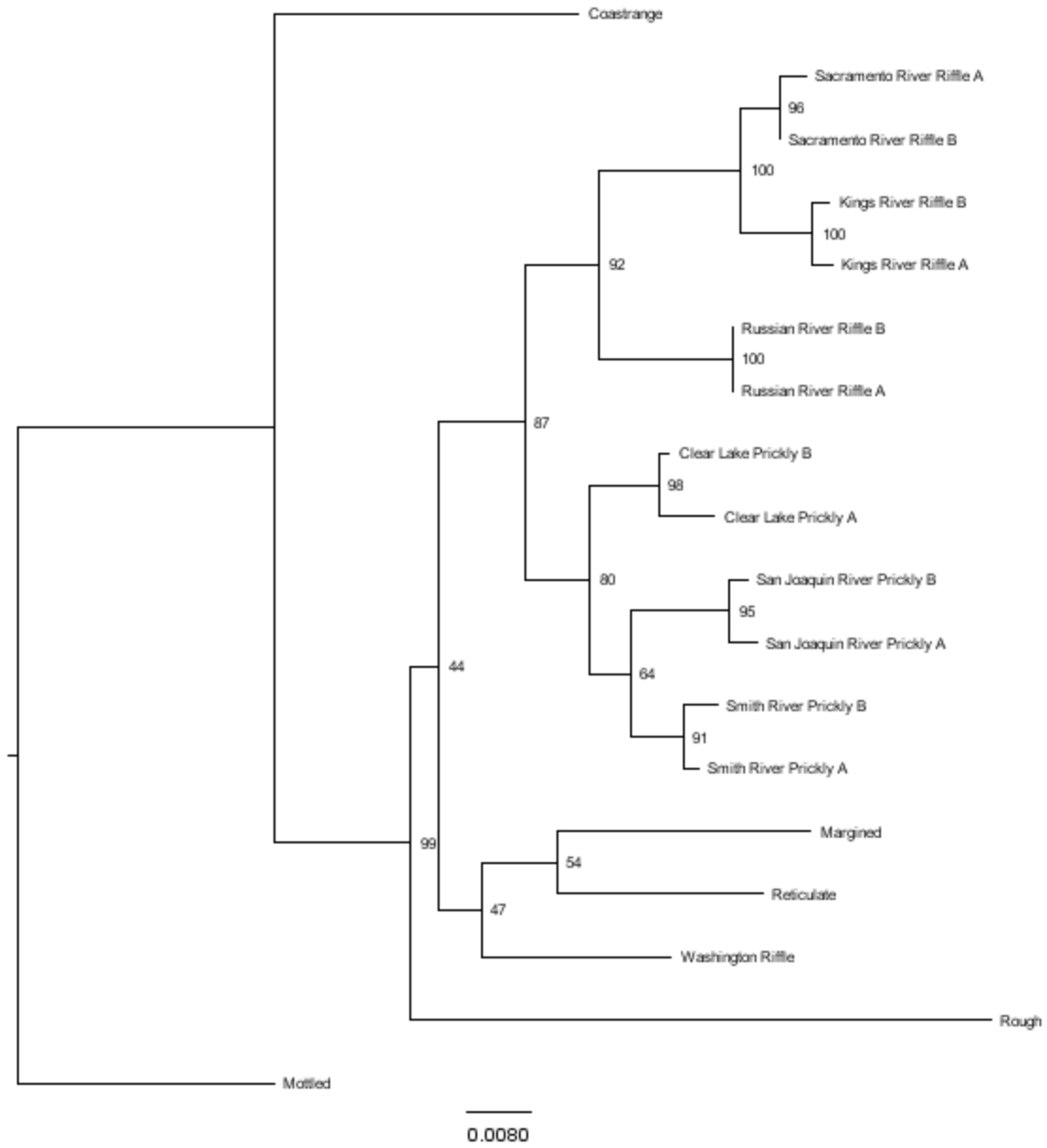


**Figure 8:** Multi-gene tree, generated with Maximum Likelihood.

The multi-loci tree showed many of the same trends depicted in the single gene trees. Coastrange, rough, Washington riffle, reticulate, and margined sculpin have more distant ancestry to the riffle and prickly sculpin. Riffle – (except for the Washington riffle), and prickly each formed their own clade. Within the riffle clade, Russian River riffle were more distant from the Kings and Sacramento River riffle. In the prickly clade, Clear Lake prickly were more distant from the San Joaquin and Smith River prickly.

**Discussion.** The multi-locus tree appears to provide a better representation of the evolutionary relationships of 7 species of Eastern Pacific *Cottus* than the single-locus

trees. Trends found scattered through groupings of the single gene trees are clearly represented in the concatenated species tree with extremely strong support values for the branching. Many of the genetic relationships shown are strongly supported by their presence in both the single and the multi-locus tree. However, the multi-locus trees provide clearer depictions of the relationships. Furthermore, our findings do not align with current morphological species classifications.

Freshwater sculpin are highly suspected of cryptic speciation (Moyle 2002). This species ambiguity is illustrated by our genetic analyses of Californian sculpin, which show complex and unexpected speciation patterns. Existing classifications of sculpin, such as riffle or prickly, do not seem to convey the full variation present in each species in the state. Many of these discrepancies can be tied to Califonia's phylogeography. Principally, phylogeography seeks to understand how historical processes of changing allele frequencies have left behind evolutionary implications on geographic species distribution (Avise 2000). Within California, this phylogeography is highly complex and subject to strong anthropogenic influences, causing additional ambiguity.

Lastly, because of confusing morphology, it is possible that some identification of samples may be erroneous – as may be the case in several of the riffle sculpin. Confusion of species definitions is not limited to morphological discrepancies, however, as previous species definitions of sculpin (Kinziger et al. 2005) conflict with initial phylogenetic trees generated in this study.

A number of Riffle sculpin classifications are called into question by our latest findings, though this is not entirely surprising given the life history requirements of these fish. Riffle sculpin need cold, clean, highly oxygenated water to exist and given the current distribution of riffle sculpin along the west coast, gene flow between populations would be highly unlikely (Moyle 2002). For example, the sample from Washington, called "Washington riffle" may be wrongly classified. The Washington "riffle" did not group with the other riffle species, indicating that it is not the same species as these other fish. Some genetic variation would be expected between the proposed Washington riffle and the California riffle because of geographic distance (Hutchison & Templeton 1999); however, the observed distance between Washington and California riffle is too great to be completely accounted for by the geographic distance effects and the loss of gene flow.

Additional discrepancies exist in riffle sculpin in the Russian River, where preliminary results indicate that they too may not be a "true" riffle sculpin. The separation of the Russian River riffle may have a phylogeographic explanation. A coastal mountain range separates the Russian River riffle from the Kings and Sacramento River riffle. Geographic variation in the American Pacific Northwest has been linked to pronounced genetic variation between species in the area (Brunsfeld et al. 2001).

Lastly, the margined sculpin sample also contradicts existing classifications of this species of sculpin. According to Kinziger, Wood, and Neely (2005), the margined sculpin shares a clade with riffle sculpin. However, in the trees generated by this study, the margined sculpin is in a more distantly related clade with the Washington riffle and reticulate sculpin.

Relationships among prickly sculpin also show additional complexity than would be initially assumed. Clear Lake prickly sculpin form a separate clade than the San Joaquin River and Smith River prickly. This separation, based in higher numbers of genetic differences, could indicate reproductive isolation between the Clear Lake prickly and the other prickly. Reproductive isolation has been shown to arise in fish over

a relatively short period of time (Hendry, Wenburg, Bentzen, Volk & Quinn 2000). Clear Lake prickly, collected from a lake environment after thousands of years in isolation from other species of prickly, appear to have become reproductively isolated from the other prickly, certainly more so than the various species of prickly typically inhabiting river environments.


**Figure 9:** Geographic distribution of freshwater sculpin samples used in this study.

Future research, utilizing more samples from varied locations as well as different methods of analysis, may contribute to clarifying the convoluted species definitions of freshwater sculpin. Sites and Marshall (2004) suggest using multiple strategies of analysis to empirically delineate species. Combining data from morphological classification with

genetic information from several sources – including mitochondrial, nuclear, and microsatellite analyses, may provide the clearest picture of freshwater sculpin speciation and distribution. The results of this study, generated through a multi-gene species tree approach, may contribute to understanding the role of sculpin in California's waterways and their potential to demonstrate principles of phylogenetic and ecosystem complexity.

## References


Avise, J.C., Arnold, J., Ball, R.M., Bermingham, E., Lamb, T., Neigel, J.E., Reeb, C.A., and N.C. Saunders. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematic. *Annual Review of Ecological Systems* 18: 489-522.

Avise, J.C. 2000. *Phylogeography: The History and Formation of Species.* Cambridge, MA: Harvard University Press.

Barraclough, T.G., Nee, S. 2001. Phylogenetics and speciation. *TRENDS in Ecology and Evolution*. 16(7):391-399.

Baumsteiger, J. Unpublished. Speciation and phylogeography of the generalist prickly sculpin (*Cottus asper*) and two specialists – riffle (*C. gulosus*) and Pit (*C. pitensis*) sculpin.

Brito, P.H., Edwards, S.V. 2009. Multilocus phylogeography and phylogenetics using sequence-based markers. *Genetica* 2009(135):439-455.

Burnsfeld, S.J., J. Sullivan, D.E. Soltis, and P.S. Soltis. 2001. Comparative phylogeography of northwestern North America: a synthesis. In: *Integrating Ecological and Evolutionary Processes in a Spatial Context* (eds. Silvertown J., Antonovics J.), pp.319–339. Blackwell Science, Oxford.

Davis, J.I. & Nixon, K.C. 1992. Populations, genetic variation, and the delimitation of phylogenetic species. *Systematic Biology* 41(4): 421-435

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and throughput. *Nucleic Acids Research* 32(5): 1792-1797.

Faircloth, B.C. 2008. MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* (2008)8: 92-94.

Frankham, R. 2005. Genetics and extinction. *Biological Conservation* 126(2005): 131-140.

Guindon, S. & Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, 52(5): 696-704.

Hall, B.G. 2004. *Phylogenetic trees made easy: a how-to manual.* Sunderland, MA: Sinauer Associates.

Hartl, D.L. 1981. *A Primer of Population Genetics*. Sunderland, MA: Sinauer Associates.

Hendry, A.P., J.K. Wenburg, P. Bentzen, E.C. Volk, T.P. Quinn. 2000. Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science* 290(5491): 516-518.

Huelsenbeck, J.P. & Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17(8): 754-755.

Hutchison, D.W. & A.R. Templeton. 1999. Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. *Evolution* 53(6): 1898-1914.

Jukes, T.H. & Cantor, C.R. 1969. *Evolution of protein molecules.* New York: Academic Press.

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

Kinziger, A.P., Wood, R.M., Neely, D.A. 2005. Molecular Systematics of the Genus *Cottus* (Scorpaeniformes: Cottidae). *Copeia* 2005(2): 303-311.

Kumar, S., Dudley, J, Nei, M., Tamura, K. 2008. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. Briefings in *Bioinformatics* 9: 299-306.

Lemey, P., Salemi, M., Vandamme, A.M. 2009. *The Phylogenetic Handbook.* Cambridge, UK: Cambridge University Press.

Lindenmayer, D.B., Margules, C.R., Botkin, D.B. 1990. Indicators of biodiversity for ecologically sustainable forest management. *Conservation Biology* 4(4): 355-364.

McClurg, S. 2000. *Water and the Shaping of California.* Sacramento, CA: Water Education Foundation and Heyday Books.

McGinnis, S.M. 1984. *Freshwater Fishes of California.* Berkeley, CA: University of California Press.

Moyle, P.B. 2002. *Inland Fishes of California.* Berkeley, CA: University of California Press.

Nei, M., Maruyama, T., Chakraborty, R. 1975. The bottleneck effect and genetic variability in populations. *Evolution* 29(1): 1-10.

Noss, R. 1990. Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology* 4: 355-364.

Porter, A.H. & Johnson, N.A. 2007. Speciation despite gene flow when developmental pathways evolve. *Evolution* 56(11): 2103-2111.

Posada, D. 2008. jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution* 25(7): 1253-1256.

Rosenberg, M.S. & Kumar, S. 2001. Traditional phylogenetic reconstruction methods reconstruct shallow and deep evolutionary relationships equally well. *Molecular Biology and Evolution.* 18(9): 1823-1827.

Scotland, R.W. 2010. Deep homology: A view from systematics. *Bioessays* 32:438-449.

Sites, J.W. Jr, and J.C. Marshall. 2004. Operational criteria for delimiting species. *Annual Review of Ecology, Evolution, and Systematics* 35: 199-227.

Systematics and Molecular Phylogenetics. 2004. From *NCBI Resources*. Retrieved from http://www.ncbi.nlm.nih.gov/About/primer/phylo.html.

Tavare, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. In: *Some Mathematical Questions in Biology - DNA Sequence Analysis*, 57-86. Providence, RI: American Mathematical Society.

Via, S. 1999. Reproductive isolation between sympatric races of pea aphids: gene flow restriction and habitat choice. *Evolution* 53(5): 1446-1457.

***Salt Pond – Eddie Campbell***