The evolutionary history of the embiotocid surfperch radiation based on genome-wide RAD sequence data

Gary Longo, Giacomo Bernardi

Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA, USA

Abstract

The radiation of surfperches (Embiotocidae) in the temperate North Pacific has been suggested to be the product of ecological competition and niche partitioning. Surfperches are a family of viviparous marine fishes, which have been used to study multiple paternity, sperm competition, and population genetics. Phylogenetic inference is essential for interpreting the evolutionary context of embiotocid life history traits and testing alternative scenarios, yet previous studies have yielded phylogenies with low support and incongruent topologies. Here we constructed reduced representation genomic libraries using restriction-site associated DNA (RAD) sequence markers to infer phylogenetic relationships among all genera and 22 out of 24 embiotocid species. Orthologous markers retained across 91% of sampled species, corresponding to 523 loci, yielded trees with the highest support values. Our results support a scenario where embiotocids first diverged into clades associated with sandy and reef habitats during the middle Miocene (13–18 Mya) with subsequent invasions of novel habitats in the reef associated clade, and northern range expansion in the Northwest Pacific. The appearance of California kelp forests (Laminariales) in the late Miocene (8–15 Mya) correlates with further proliferation in the reef associated clade. In all cases, radiations occurred within specific habitats, a pattern consistent with niche partitioning. We infer fine scale species relationships among surfperches with confidence and corroborate the utility of RAD data for phylogenetic inference in young lineages.

1. Introduction

Adaptive radiations result from divergence of an ancestral population into an array of species that inhabit a variety of environments and that differ in traits used to exploit those environments (Schluter, 2000), such as the Galapagos finches (Grant, 1999; Lack, 1947), the Hawaiian silverswords (Baldwin and Sanderson, 1998), the Caribbean anoles (Losos, 2000), and the great African rift lake cichlids (Fryer and Iles, 1972). Adaptive radiations need to fulfill three requirements: multiplication of species and common descent, adaptation, and extraordinary diversification (Glor, 2010). There are relatively few described cases in marine fishes, such as the New Zealand triplefins (Hickey et al., 2009), California Sebastes rockfish (Johns and Avise, 1998), Antarctic notothenioids (Janko et al., 2011), Caribbean hamlets (Puebla et al., 2008) and South African clinids (von der Heyden et al., 2011). The paucity of adaptive radiations in marine fishes may be due to a number of factors including life history characteristics that are conducive to panmixia (Bernardi, 2013). Indeed most marine fishes have a bipartite life history where adults exhibit a mostly sedentary life while larvae remain pelagic for days to months (Leis, 1991). Protracted pelagic larval stages often result in high dispersal potential accompanied by high levels of gene flow (Reece et al., 2010; Selkoe and Toonen, 2011) potentially hindering adaptation to particular environments and therefore limiting local adaptation. Therefore, marine fishes that lack a pelagic larval stage may provide unique insights into such studies, such as the surfperches (Embiotocidae).

The family Embiotocidae comprises 24 species that are found in the temperate coastal waters of the North Pacific from Mexico to Japan but are absent from the higher latitudes along the Aleutian Islands. The center of distribution and the only known embiotocid fossils are located in central California, it has therefore been assumed that this is also their center of origin (David, 1943; Tarp, 1952) (Fig. 1). Reproductive courtship occurs in the winter and females retain sperm from multiple matings for up to several months before fertilization, and later give live birth from spring through late summer depending on the taxa (Bennett and
Therefore, unlike most marine fishes, surfperches do not undergo a dispersive pelagic larval stage. As expected, this alternative life history strategy results in restricted gene flow and may increase the potential for local adaptation (Bernardi, 2005, 2000).

Fig. 1. Embiotocid species richness map in the eastern Pacific based off distributions from Love (2011).
Surfperches along with damselfishes (Pomacentridae), cichlids (Cichlidae), and wrasses (Labridae) formed the traditionally recognized Labroids (Lauder and Liem, 1983). However, subsequent phylogenetic work rendered this group paraphyletic, separating wrasses from the closely related damselfishes, cichlids, and surfperches (Mabuchi et al., 2007) and grouping the latter taxa with a few other families in a larger cluster called Ovalentaria (Betancur-R et al., 2013; Wainwright et al., 2012). Compared to damselfishes (390+) and cichlids (1670+), embiotocids include very few species (24), yet surfperches have invaded very diverse habitats including deep zones, coastal pelagic zones, kelp forest rocky reefs, sandy bottoms, shallow seagrass beds, estuarine zones and even inland freshwater (Allen and Pondella, 2006; Tarp, 1952). These diverse environments may have served as adaptive zones where ancestral surfperches radiated to take advantage of under exploited niches. Generally, niche partitioning and habitat use can either arise from a single lineage invading an open niche and diversifying in situ, or from multiple lineage invasions. These processes are not mutually exclusive and a combination of scenarios can give rise to observed radiations.

In fishes, the first stage of adaptation often includes ecological niche partitioning, as seen in threespine sticklebackes along the benthic–limnetic axis and later in several other systems (Schluter, 2000). Niche partitioning has also been suggested as a potential mechanism for diversification in surfperches (Ebeling and Laur, 1986). Accordingly, surfperches display an array of morphologically divergent mouth shapes, dentitions, and pharyngeal jaws, which allow for diverse diets and feeding mechanisms among species and has been characterized as an adaptive radiation previously (De Martini, 1969; Drucker and Jensen, 1991). Ecological competition between closely related species (Hixon, 1980; Holbrook and Schmitt, 1992) and sexual selection on color patterns (Cummings and Partridge, 2001; Froeschke et al., 2007; Nakazono et al., 1981; Westphal et al., 2011) have also played an important role in diversification of surfperches.

To better understand the underpinnings of this radiation, a robust species level phylogeny is necessary. Currently there are three family wide surfperch phylogenies based on morphological data, mitochondrial sequence, or a combination of both. Tarp proposed a species level morphological phylogeny in his thorough revision of surfperches, where the family was divided into two sub-families, the Amphistichinae and the Embiotocinae (Tarp, 1952). This division was later corroborated using osteological characters and sequence data were proposed (Cassano, 1998) as were relationships restricted to a single genus, Embiotoca and Ditrema, or a single subfamily, Amphistichinae (Bernardi, 2009; Katufchi et al., 2010; Westphal et al., 2011). Although some topological consensus exists among these studies, definitive relationships among embiotocid taxa remain contentious. Here, we use hundreds of genome wide RAD markers to infer a robust species tree that sheds light on surfperch evolution that is consistent with patterns of niche partitioning and adaptive radiation, and provides a time-frame for those events.

### 2. Materials and methods

#### 2.1. Sampling

The proposed phylogeny is based on full representation of species for the Amphistichiinae and 16 of 18 species from the Embiotocinae. The amphistichine surfperches consist of six species divided into two genera: *Amphistichus argenteus, A. koelzi,* and *A. rhodoterus* and *Hyperprosopon anaele, H. argenteum,* and *H. ellipticum.* Embiotocines are divided into 11 genera, with 15 species: *Brachyistius alutes* and *B. frenatus; Cymatogaster aggregata; Ditrema jordani, D. temminckii,* and *D. viride; Embiotoca jacksoni* and *E. lateralis; Hypsurus caryi; Hysterocarpus anthracus* and *M. minimus; Neoditrema ransonneti; Phanerodon atripes* and *P. furcatus; Rhacochilus toxotes* and *R. vacca; and *Zalembius rosaceus.* Initially the western Pacific genus *Ditrema* was considered monotypic (*Ditrema temminckii*) but it was later split into three species: *Ditrema temminckii,* *D. jordani,* and *D. viride,* which have been shown to be very closely related taxa (Katufchi et al., 2010). We included *D. temminckii* as a representative of this clade resulting in 22 out of 24 species of surfperches for phylogenetic inference. Two individuals of each species were sequenced (Table 1) for a total of 44 individuals in the complete data set.

#### 2.2. Molecular methods

Tissue samples were stored in 95% ethanol and DNA was extracted from fin clips or liver tissue using DNeasy Blood & Tissue kits (Qiagen) according to manufacturer’s protocol. We constructed RAD libraries using a variation of the original protocol used for the genome sequencing project (Wang et al., 2014) (Table 1). A total of 44 surfperch individuals were analyzed for RAD data, including 22 species from the Amphistichinae and 18 species from the Embiotocinae. We included *Hyperprosopon anae* as a representative of this clade.

<table>
<thead>
<tr>
<th>Species</th>
<th>Sample names</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Amphistichus argenteus</em></td>
<td>AAR_CLP &amp; AAR_SIV</td>
<td>Monterey Bay, CA</td>
</tr>
<tr>
<td><em>Amphistichus koelzi</em></td>
<td>AKO_MBA1 &amp; AKO_MBA2</td>
<td>Monterey Bay, CA</td>
</tr>
<tr>
<td><em>Amphistochis rhodoterus</em></td>
<td>ARL_CAP &amp; ARH_NSS</td>
<td>Monterey Bay, CA &amp; Nehalem</td>
</tr>
<tr>
<td><em>Brachyistus alutes</em></td>
<td>BAL_GUA1 &amp; BAL_GUA2</td>
<td>Isla Guadaloupe, Mexico</td>
</tr>
<tr>
<td><em>Brachyistus frenatus</em></td>
<td>BFR_CAT7 &amp; BFR_CAT8</td>
<td>Catalina Island, CA</td>
</tr>
<tr>
<td><em>Cymatogaster aggregata</em></td>
<td>CAG_CB1 &amp; CAC_SD1</td>
<td>Monterey Bay, CA &amp; San Diego, CA</td>
</tr>
<tr>
<td><em>Ditrema temminckii</em></td>
<td>DTE_J103 &amp; DTE_J104</td>
<td>Japan</td>
</tr>
<tr>
<td><em>Embiotoca jacksoni</em></td>
<td>EJA_MBA &amp; EJA_PBA</td>
<td>Monterey Bay, CA &amp; Punta</td>
</tr>
<tr>
<td><em>Embiotoca lateralis</em></td>
<td>ELA_MBA &amp; ELA_PBA</td>
<td>Monterey Bay, CA &amp; Punta</td>
</tr>
<tr>
<td><em>Hyperprosopon anae</em></td>
<td>HAR_MIB1 &amp; HAR_MIR2</td>
<td>Monterey Bay, CA</td>
</tr>
<tr>
<td><em>Hyperprosopon argenteum</em></td>
<td>HAR_MBA &amp; H MIB</td>
<td>Monterey Bay, CA</td>
</tr>
<tr>
<td><em>Hyperprosopon ellipticum</em></td>
<td>HEL_MBA &amp; HEL_MIB</td>
<td>Monterey Bay, CA &amp; Pacifica, CA</td>
</tr>
<tr>
<td><em>Hypsurus caryi</em></td>
<td>HCA_SCH1 &amp; HCA_SCH2</td>
<td>Monterey Bay, CA</td>
</tr>
<tr>
<td><em>Hysterocarpus traskii</em></td>
<td>HTR_01 &amp; HTR_02</td>
<td>Monterey Bay, CA</td>
</tr>
<tr>
<td><em>Micrometros aurora</em></td>
<td>MAU_MBA1 &amp; MAU_MBA2</td>
<td>Monterey Bay, CA</td>
</tr>
<tr>
<td><em>Micrometros minimus</em></td>
<td>MML_MBA &amp; MML_MIB</td>
<td>Monterey Bay, CA &amp; San Diego, CA</td>
</tr>
<tr>
<td><em>Neoditrema ransonneti</em></td>
<td>NRA_J1 &amp; NRA_J2</td>
<td>Japan</td>
</tr>
<tr>
<td><em>Phanerodon atripes</em></td>
<td>PAT_MON &amp; PAT_NR1</td>
<td>Monterey Bay, CA &amp; Santa</td>
</tr>
<tr>
<td><em>Phanerodon furcatus</em></td>
<td>PFU_SCH1 &amp; PFU_SCH2</td>
<td>Monterey Bay, CA</td>
</tr>
<tr>
<td><em>Rhacochilus toxotes</em></td>
<td>RTO_MBA &amp; RTO_PBA</td>
<td>Monterey Bay, CA &amp; Punta</td>
</tr>
<tr>
<td><em>Rhacochilus vacca</em></td>
<td>RVA_MBA &amp; RVA_PBA</td>
<td>Monterey Bay, CA &amp; Punta</td>
</tr>
<tr>
<td><em>Zalembius rosaceus</em></td>
<td>ZRO_M12 &amp; ZRO_M13</td>
<td>Monterey Bay, CA</td>
</tr>
</tbody>
</table>
(Baird et al., 2008; Miller et al., 2007) with restriction enzyme SbfI as described in Miller et al., 2012 with minor revisions reported below. Initial genomic DNA concentrations for each individual were 400 ng. We physically sheared libraries on a Covaris S2 sonicator with an intensity of 5, duty cycle of 10%, cycles/burst of 200, and a cycle time of 30 s. The final PCR amplification step was carried out in 50 μl reaction volumes with 18 amplification cycles. For all size selection and purification steps we used Ampure XP beads (Agencourt). Samples used in this study were sequenced in one of two libraries, each containing 96 individually barcoded samples. Each library was sequenced in a single lane on an Illumina HiSeq 2000 at the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley.

2.3. Quality filtering and marker discovery

Raw reads were trimmed to 92 bp on the 3’ end, quality filtered, and then split according to the 6 bp unique barcode using Miller et al. (2012) custom Perl scripts. Sequences were dropped if the product of quality scores for their 92 bases was below 80%. The barcode (6 bp) and restriction site residue (6 bp) were then removed from the 5’ end, resulting in a final sequence length of 80 bp.

We used the software program Stacks version 1.13 (Catchen et al., 2013, 2011) to identify orthologous sequences among surfperch taxa. We first ran the program denovo_map.pl, which runs all three Stacks components in a pipeline (i.e., ustacks, cstacks, and sstacks). We set a minimum stack depth (−m) of three, a maximum of three mismatches per loci for each individual (−M), and allowed up to seven mismatches when building catalog loci (−n). We excluded highly repetitive stacks, set the max number of stacks per locus to two, and disabled calling haplotypes from secondary reads. We then ran multiple iterations of the Stacks program populations to generate output files for input into downstream phylogenetic programs. This internal program allows for fine control of which markers will be exported by specifying the number of species (−p) and the percentage of individuals in each species (−r) that must possess that marker. Due to generally high coverage across individuals, we increased the minimum stack depth (−m) to ten for all populations runs. In order to generate datasets that reflected various levels of orthologous sequence retention, we ran the program with −p set at 14, 16, 18, 20 & 21, which corresponds to about 64%, 73%, 82%, 91%, and 96% of surfperch species possessing the marker, respectively. Additionally, we ran −r set to 50% and 100% (i.e., one or both of the individuals in each species possesses the marker) for each of the five −p settings, resulting in 10 total datasets. Full sequence RAD markers of each individual were exported for downstream analyses. The quality filtered sequences are deposited at the National Center for Biotechnology Information short-read archive (accession no. SRP056799).

2.4. Phylogenetic inference

For each Stacks populations parameter set, we built supermatrices with complete RAD sequences (80 bp) and identified phylogenetically informative sites using FASconCAT-G (Kück and Longo, 2014). Supermatrices were generated both with individual sequence data and species consensus sequence data with IUPAC ambiguity codes for polymorphic data.

We used both maximum likelihood methods as implemented in PhyML (Guindon et al., 2010) and Bayesian phylogenetic inference as implemented in MrBayes (Ronquist et al., 2012) to assess relatedness within Embiotocidae. For MrBayes analyses we partitioned the dataset by each 80 bp locus, in FASconCAT-G, which allowed each locus to be assigned its own GTR + Γ + I model and parameters. We selected a Markov Chain Monte Carlo (MCMC) search algorithm with a chain length of 1,000,000 using four chains with a sampling frequency of 1000. In PhyML, we selected the GTR model of sequence evolution, six substitution rate categories, set the initial tree to random, and performed 100 bootstrap replicates. Phylogenetic trees and corresponding support values were visualized using FigTree v1.4.0 (Rambaut, 2014). Trees were midpoint-rooted due to the constraints of using an outgroup with RAD data, which relies on retaining orthologous restriction sites.

2.5. Estimating divergence times

Divergence times were estimated using standard models of evolution implemented in BEAST (Drummond et al., 2012) assuming mutual independence among sites. Exploratory runs showed a random local clock (RLC) model, in combination with a birth–death (BD) prior for rates of cladogenesis (Drummond and Rambaut, 2007) as appropriate for our data set. Three runs were conducted with 20 million generations each, with sampling every 1000 generations. The software Tracer v1.5 (Rambaut and Drummond, 2007) was used to quantify effective sample sizes (ESS) for model parameters, and the ‘compare’ command in AWTY (Nylander et al., 2008) was used to assess convergence, with 10% of each run discarded as burn-in. Runs were combined using LogCombiner v1.7.5 (Drummond et al., 2012), and a time tree was obtained using TreeAnnotator v1.7.5 (Drummond et al., 2012).

Internally calibrating divergence in surfperches is complicated because very few fossils are available. Three late Miocene (5.3 Mya) fossils were found by Eric Knight Jordan in Lompoc, California, and described by his father David Starr Jordan, as Eriquius plectrodes (one fossil) (Jordan, 1924), and Erisceles pristinus (two fossils) (Jordan, 1925). Upon reexamination in 1941, all three fossils were assigned to the single species Eriquius plectrodes (David, 1943) and represent the only reliable surfperch fossil remains (David, 1943; Tarp, 1952). David (David, 1943) proposed the fossils most closely resembled the genus Embiotoca, however, upon our own examination of all three specimens, we could not substantiate that claim. Although dorsal ray counts are more similar to embiotocin surfperches, other characters (e.g., body depth, caudal peduncle length, and overall shape) approximate to amphistichin surfperches. We therefore declined to use these embiotocid fossil remains for internal calibration due to their uncertain taxonomic assignment. Instead we used external calibration based on published molecular phylogenies that include diverse embiotocid taxa. Specifically, two molecular studies on bony fishes and one on pomacentrids, which all used several molecular markers, allowed us to estimate the crown age of Embiotocidae to approximately 13–18 Mya (Betancur-R et al., 2013; Frédérich et al., 2013; Near et al., 2013). These dates are consistent with a published molecular phylogeny of the family based on mitochondrial (Bernardi and Bucciarelli, 1999). Therefore this prior with a normal distribution was used as a calibration point, with the minimum and maximum bounds implemented with the 95th percentile of the distribution.

3. Results

3.1. RAD sequences

The final filtered library contained 57,124,651 reads among 44 individuals. Coverage ranged from 293,878 to 4,945,582 with an average of 1,289,287 filtered reads per individual (median = 1,042,690) (Fig. S1). A potential source of variance in coverage could be due to variable quality of starting genomic DNA of each individual as well as variability in the amount of data generated from different Illumina runs. The denovo_map.pl program detected
between 34,439 and 104,494 unique, genome-wide RAD markers (i.e., loci) in each individual, which corresponded to between 568 and 5758 polymorphic loci in each individual (Table S1). Individuals with low coverage typically yielded lower numbers of loci. As we increased the stringency of populations filter parameters in Stacks (i.e., higher –p and –r values), the number of markers and overall size of the concatenated supermatrix decreased. Among the different datasets, the number of loci retained ranged from 116 to 30,629 while the number of parsimony informative nucleotide sites pulled from those loci ranged from 96,368 (Table S2).

3.2. Phylogenetic relationships

Phylogenetic reconstruction using more stringent datasets resulted in nearly identical topologies while low stringency datasets resulted in poorly resolved trees with low support (Fig. S2). The best supported tree (Fig. 2 & Fig. S3) was generated with the populations parameters –p 20 & –r 1 dataset (Table S2) (nexus file available at <http://dx.doi.org/10.6084/m9.figshare.1365521>). The inferred topology recovered monophyletic groups for the subfamilies Amphistichinae and Embiotocinae with high support. Within Amphistichinae, the genus Hyperprosopon was found to be paraphyletic with H. anale branching first from the rest of the amphistichines with high bootstrap support. However, the remaining Amphistichinae included two clades, which comprised on one hand the sister species Hyperprosopon argenteum and H. ellipticum, and on the other hand the Amphistichus clade: A. rhodoterus, A. argenteus, and A. koelzi. The sub-family Embiotocinae consists of two major clades. The first is comprised of four species with small body sizes including, Micrometrus aurora and M. minimus joined as sister taxa, and another sub-clade with Cymatogaster aggregata joined with the freshwater Tule perch, Hysterocarpus traskii. The second major clade indicates a divergence between the western Pacific species, Ditrema temminckii and Neoditrema ransonnetii, and a more diverse clade of eastern Pacific species. The eastern Pacific species separate into two sub-clades. One sub-clade includes the silvery open-water species, loosely associated with kelp (Laminariales) and rocky reefs, in the genera Phanerodon, Rhacochilus, and the deep water species Zalembius rosaceus, and the other sub-clade includes the species more strongly associated with kelp reefs in the genera Embiotoca, Hypsurus, and Brachyistius. Within this clade, the two species of Brachyistius are very closely related with a sequence divergence of only 0.025%. This is slightly larger than the divergence between individuals within a species sampled at large distances from each other (Monterey, California, and Punta Banda, Mexico) such as Embiotoca jacksoni (0.015%), E. lateralis (0.012%), Rhacochilus toxotes (0.007%), and R. vacca (0.010%). It is, however, a value that is about one order of magnitude smaller than the average sequence divergence observed in other surfperch sister species (0.317%, Table S3).

3.3. Habitat partitioning

We simplified surfperch habitat use into broad categories and color-coded accordingly on the phylogenetic tree described above (Fig. 2). Phylogenetic clusters strongly partitioned by habitat: sandy, seagrass, rocky, and kelp. Interestingly the shiner perch, which has the widest range of salinity tolerance and is known to inhabit estuaries, and sporadically freshwater (Love, 2011), partitioned with the only freshwater species, the Tule perch Hysterocarpus traskii.

![Fig. 2. Phylogeny of Embiotocidae inferred using genome-wide RAD markers (p20_r1 supermatrix) with maximum likelihood and Bayesian methods. Species consensus sequences were used for phylogenetic inference here. Node values represent posterior probability and bootstrap support (top and bottom, respectively). Taxa are colored coded based on habitat preference.](https://example.com/fig2.png)
3.4. Estimating divergence times

In order to estimate divergence times among embiotocid taxa we used a Bayesian multispecies coalescent approach, as implemented in BEAST (Drummond et al., 2012), based on a 13–18 Mya external calibration for absolute age of the family (Fig. 3). Our data suggest the crown Amphistichinae clade diverged approximately 5 Mya. The earliest Embiotocinae divergence occurred approximately 10 Mya and split the group into the smaller surfperches (i.e., Micrometrus, Cymatogaster, Hysterocarpus) on one hand and the relatively larger embiotocines on the other hand. The larger surfperches subsequently diverged approximately 7 Mya to give rise to the western Pacific and the eastern Pacific species. Roughly 5 Mya the eastern Pacific species diverged into two groups, the rocky reef species and the kelp reef species, each of which diversified approximately 2.5–3 Mya.

4. Discussion

4.1. Taxonomic notes

We generated a very robust phylogenetic hypothesis for Embiotocidae based on thousands of genome wide, phylogenetically informative DNA bases drawn from the most complete representation of the family to date. Our results are mostly in agreement with previously published phylogenies (Bernardi and Bucciarelli, 1999; Tarp, 1952; Westphal et al., 2011) but allow for reassessment of taxonomic issues due to better resolution and support.

In the subfamily Amphistichinae, the genus Hyperprosopon is likely paraphyletic, with H. anale branching first, and sister to the two genera Amphistichus and Hyperprosopon. One year after its description as Hyperprosopon anale (Agassiz, 1861), the species was re-described as Hypocritichthys anale (Gill, 1862), an available name that could be used in light of our new findings. In the subfamily Embiotocinae, the sister species Brachyistius frenatus and B. aletes may either be considered two different valid species or two diverging populations experiencing incipient speciation. The clade that includes Phanerodon and Rhacochilus reveals that both genera are paraphyletic. A simple way to resolve this issue would be to include all members of this clade in a single genus. Using a single genus for such a diverse group is consistent with other embiotocid examples such as the genus Micrometrus, which contains two species that are more divergent than any of the species contained in this clade. The genera Phanerodon and Rhacochilus were both described in 1854, however Rhacochilus was described earlier (May vs. October) and would therefore have precedence (Agassiz, 1854; Girard, 1854). Finally, the rainbow seaperch, Hypsurus cari, originally described as Embiotoca cari, was found to be the sister species of Embiotoca jacksoni. Therefore its original name, Embiotoca cari, should be used, as previously suggested (Bernardi, 2009).

4.2. Ecological speciation and niche partitioning

One of the most salient features of the phylogenetic hypothesis presented here is the strong correlation between habitat use and phylogenetic relationships (Fig. 2). As mentioned earlier, diversification in a given habitat may happen via a single invasion followed by a radiation in situ, or by repeated invasions of different evolutionary lineages. Our data suggest that the former process repeatedly occurred during surfperch evolution, while there is no evidence of the latter process ever happening. Indeed, ancestral surfperches invaded different habitats and speciation followed

Fig. 3. Time calibrated phylogeny of Embiotocidae using an external time calibration based on two recently published, large-scale molecular phylogenies. Horizontal blue bars at the nodes represent the 95% confidence intervals for each date estimate.
within those habitats, resulting in the patterns observed here where habitat and phylogenetic clusters are perfectly correlated (Fig. 2).

The phylogeny proposed here suggests early divergence based on sandy versus shallow reef habitat followed by further specialization within each lineage. In the Amphistichinae, the spotfin surfperch, *Hyperprosopon anaele*, is sister to the remaining lineages, which was previously proposed by (Westphal et al., 2011). Little work has been done on *H. anaele*, but their diet consists largely of zooplankton (Love, 2011), while other amphistichines, such as the barreled surfperch, *Amphistichus argenteus*, feed primarily on benthic invertebrates (Carlisle et al., 1960). It is therefore plausible that the extant benthic, sand dwelling amphistichin surfperches radiated from a limnetic ancestor that fed on zooplankton. The remaining taxa sort into Amphistichus and Hyperprosopon clades. Both of these Hyperprosopon species are widely distributed from Washington or Oregon, to Baja California, Mexico (one species, *H. argenteum*, is also found on Guadalupe Island, off Baja California). In contrast, the distribution of Amphistichus is variable. While *A. koelzi* is also widely distributed from Washington to Baja California, *A. argenteus* and *A. rhodoterus*, the barreled and redfin surfperch respectively, are nearly allopatric. The redfin surfperch is distributed from British Columbia to Central California, while the barred surfperch prevails from Central California to northern Baja California, Mexico. Therefore in this group, allopatric speciation may have played a prominent role.

While the ancestors of Amphistichinae radiated over sandy habitats, the Embiotocinae lineage invaded and diversified in the near shore reef environment of the temperate North Pacific. Within this group, the clade containing the genera Micrometrus, Hysterocephalus, and Cymatogaster is sister to the rest of the lineage. Cymatogaster is a habitat generalist. However, these small surfperch tend to preferentially be found in shallow waters and often associate with seagrass (Byerly, 2001; De Martini, 1969: Love, 2011). Notably, *C. aggregata* has a high salinity tolerance permitting it to frequent estuarine and even enter freshwater for short periods of time (Love, 2011). *C. aggregata* is sister to the only obligate freshwater embiotocid, the Tule perch, *H. traskii*, indicating that ecological speciation likely drove their most recent common ancestor to split and invade freshwater as noted by Bernardi and Bucciarelli (1999). Micrometrus spp. are mostly restricted to shallow waters and are commonly found over seagrass or algae beds (Love, 2011). The next branching event led to the western Pacific surfperches, Ditremia and Neoditremia, which often associate with nearshore environments with low lying biological structure such as seagrass beds. Based on the common association with seagrass in extant embiotocines just discussed, it is plausible the most recent common ancestor of the Embiotocinae also associated with seagrass.

The next group comprises the most diverse surfperch group, the reef associated species, which is divided into two major ecological clusters. One includes the primarily rocky reef associated species (i.e., *Rhacochilus + Zalembius clade*), and the other includes the kelp associated species (i.e., *Embiotoca + Brachyistius clade*). Zalembius rosaceus differs from the rest of the clade as it is most commonly found under soft bottoms in depths greater than 50 m (Love, 2011). *Rhacochilus* and *Phanerodon* spp. are structure-oriented generalists that can be found in and around rocky reefs and kelp forests, as well as manmade structures such as pilings. The species of the last clade, *Embiotoca, Hypsurus*, and *Brachyistius*, are most typically associated with kelp. These species are very specialized and show high levels of ecological competition (Hixon, 1980). A number of feeding specializations evolved within this clade, such as winnowing (sorting food within the mouth). Winnowing is found in *R. toxotes, E. jacksoni, H. caryi*, and conspicuously absent in *E. lateralis*, which is consistent with the sister relationship of *E. jacksoni* and *E. caryi*.

4.3. Tempo and mode of evolution

Embiotocidae diverged relatively recently, approximately 13–18 Mya. The family likely originated off the coast of California as the only known fossils are from the Lompoc deposits (southern CA) (David, 1943; Jordan, 1924) and the center of distribution is central California (Tarp, 1952) (Fig. 1). Within the subfamily Embiotocinae, the group of smaller species (i.e., *Cymatogaster, Hysterocephalus, and Micrometrus*) diverged approximately 10 Mya. Within this group, divergence times between taxa are much greater compared to other surfperch groups. For example the two *Micrometrus* species, reef and dwarf surfperches, diverged between 3.5 to 4 Mya (Fig. 3). The next branching event, which occurred approximately 7 Mya, led to the western Pacific surfperches, *Ditrema* and *Neoditrema*, which may have migrated across the northern Pacific during a warmer climatic period (David, 1943: Tarp, 1952). During that time, the northern Pacific was dominated by seagrass, as evidenced by seagrass feeding marine mammals and invertebrates (Estes and Steinberg, 1988). The ocean cooling in the late Miocene, together with an influx of nutrients, led to a shift from a seagrass dominated system to a kelp dominated system (Bolton, 2010; Brasier, 1975; Estes and Steinberg, 1988). The expansion of kelp resulted in an increased breadth of niches and resources, which likely contributed to a radiation in the later surfperches. Indeed, the greatest diversity of extant embiotocin surfperches can be found in or around kelp forests. Within this new habitat, two major embiotocin groups evolved within the past 5 Mya. On one hand the rocky reef associated species (*Rhacochilus + Zalembius clade*), and on the other hand the kelp associated species (*Embiotoca + Brachyistius clade*).

At the same time, approximately 3.5–4 Mya, the Amphistichinae diverged into the two current genera *Hyperprosopon* and *Amphistichus*. Therefore most of the embiotocid diversification occurred between 5 and 3.5 Mya ago. It is interesting to note that the sole fossil representative of the surfperches, *Eriquius*, dated at 5.3 My, is also from that general era, which might suggest that this was a time of great diversification in surfperches.

4.4. Conclusion

Empirical and theoretical support for the applicability of RAD data for phylogenetic inference is growing (Emerson et al., 2010; Hipp et al., 2014; Rubin et al., 2012; Viricel et al., 2014; Wagner et al., 2013). Here we corroborate the applicability of RAD data and infer the most complete and fine scale Embiotocidae phylogeny with high support values. Embiotocids likely radiated ~13–18 Mya (Frédérich et al., 2013; Wainwright et al., 2012), making this one of the older lineages RAD data has been applied to empirically.

Surfperches diverged relatively recently and comprise comparatively few species (24). Although older, the other two major Ovamentaria families sensu Betancur-R et al. (2013), Pomacentridae and Cichlidae, comprise one to two orders of magnitude more species. Compared to oviparous groups, it is likely that the life history strategy of the livebearing surfperches does slow down speciation rates, yet, even within Embiotocidae, we can see some remarkable ecological diversity that in many respects encompasses what is observed in more diverse families.

The adaptive radiations of cichlids in the great African lakes show remarkable examples of convergent evolution, where mouth, teeth, and pharyngeal jaw shapes are strikingly similar among similar ecotypes (Kocher et al., 1993). Such convergence is not
completely surprising and recent theoretical models have shown that relatively high levels of convergence should be expected in adaptive radiations (Muschick et al., 2012; Scheffer and van Nes, 2006; terHorst et al., 2010). Morphological convergence of feeding structures between fish families has been shown before as well (Norton and Brainerd, 1993). As in cichlids, surfperches mouth, teeth, and pharyngeal jaw shape have been correlated to their diverse feeding mechanisms and ecological niches (De Martini, 1969). Recently, genetic and genomic approaches have been used to pinpoint the regions responsible for the evolution of structures involved in feeding specializations in cichlids (Albertson et al., 2003a,b; Brawand et al., 2014; Muschick et al., 2012) It will be interesting to determine if the same genomic regions are also involved in the surfperch radiation. It appears that major divergences in surfperches, which strongly correspond to habitat, occurred relatively recently and were followed by subsequent specialization. Dispersal during favorable climatic conditions likely led to the divergence of eastern and western Pacific clades with the remaining clades likely arising through niche specialization. As for many other lineages, the diversity of surfperches has arisen through distinct ecological and evolutionary processes over space and time. Whether or not surfperch meet the three criteria of an adaptive radiation as outlined in Glor (2010) (i.e., multiplication of species and common descent, adaptation, and extraordinary diversification) remains uncertain. Our data and previous phylogenetic analyses support multiplication of species and common descent, while other work has shown potential ecomorphological adaptions (De Martini, 1969), however more definitive work is needed. The 24 species of surfperches involved in the surfperch radiation.

Acknowledgments

We would like to thank Michael R. Miller for his generosity and insight on RAD-seq techniques. We would like to thank Karen Crow, Michael Westphal, and Steve Moyer (team surfperch) for insight on RAD-seq techniques. We would like to thank Karen Crow, Michael Westphal, and Steve Moyer (team surfperch) for their assistance with the exploration stages of analysis pipelines for phylogenetic RAD data. We would also like to thank Stephen Hauskins for his ongoing IT support. This project was funded by the Lerner-Gray fund for Marine research, Meyers Trust, and the UCSC Ecology and Evolutionary biology research and travel fund. This work used the Vincent J. Coates Genomics Sequencing Laboratory at UC Berkeley, supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at: http://dx.doi.org/10.1016/j.jympev.2015.03.027.