Molecular phylogeny of the hexagrammid fishes using a multi-locus approach

Karen D. Crow, a,* Ziyusei Kanamoto, b and Giacomo Bernardi a

a Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95060, USA
b Division of Aquatic Biology and Ecology, Center for Marine Environmental Studies, Ehime University, Matsuyama 790-8577, Ehime, Japan

Received 7 November 2003; revised 17 March 2004
Available online 30 May 2004

Abstract

Ideally, organisms are grouped into monophyletic assemblages reflecting their evolutionary histories. Single (molecular) markers can reflect the evolutionary history of the marker, rather than the species in question, therefore, phylogenetic relationships should be inferred from adequate sampling of characters. Because the use of multiple loci greatly improves the resolving power of the molecular assay, we constructed a molecular phylogeny of the family Hexagrammidae based on six loci, including two mitochondrial and four nuclear loci. The resulting molecular phylogeny, from the combined data, was significantly different from the morphological topology suggested by Shinohara [Memoirs of the Faculty of Fisheries, Hokkaido University 41 (1994) 1]. Our data support a monophyletic assemblage for the genera Hexagrammos and Pleurogrammus. However, other taxa traditionally included in the family Hexagrammidae did not form a monophyletic assemblage. The monotypic genus Ophiodon was more closely associated with cottids than with other hexagrammids. Our data concur with the morphological topology in that the genera Zaniolepis and Oxylebus formed a monophyletic clade, which was distinct and basal to the remaining hexagrammids, seven cottids and one agonid.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Hexagrammidae; Molecular phylogeny; Multiple loci; mtDNA; Intron

1. Introduction

Historically, organisms have been classified using morphological characters. Ideally, organisms are grouped into monophyletic assemblages (clades) that are highly supported from adequate sampling of characters, that reflect their evolutionary history through homology by descent. Recently, the use of molecular markers has provided powerful tools to resolve ambiguous phylogenetic relationships. However, the field has progressed primarily using single molecular markers. Phylogenies based on a single molecular marker may, in some cases, result in errors in phylogenetic inference, which reflect the evolution of the marker rather than the species (i.e., gene trees vs. species trees). When using molecular characters, the use of multiple loci reduces the probability of constructing gene trees, and increases the probability of recovering the correct phylogeny. In this study, we constructed a phylogeny of the family Hexagrammidae using six molecular loci. Each locus was evaluated independently before the combined dataset was analyzed. The resulting molecular phylogeny was then compared with the existing morphological data (Shinohara, 1994). We conclude by addressing the utility of using multiple loci for making phylogenetic inferences.

1.1. Background on hexagrammids

The family Hexagrammidae (greenling) is comprised of 3–5 genera (Hexagrammos, Pleurogrammus, Ophiodon, Oxylebus, and Zaniolepis) with 9–13 species depending on which systematic scheme is followed. For example, Quast (1964) and Hart and Clemens (1973) claim that H. lagocephalus and H. superciliosus are synonymous, and Nelson (1994) maintains that
*P. monopterygius* and *P. azonus* are synonymous. Finally, considerable doubt has been raised as to whether the genera *Oxyelebus* and *Zaniolepis* should be included in the family and it has been proposed that these two genera be reclassified as a separate family, the Zaniolepididae (Shinohara, 1994). The current systematics of this family are based on morphological characters. However, in some cases the variability of morphological characters within species is greater than among genera confounding their taxonomic status (Rutenberg, 1962). Estimates of genetic variation could help resolve taxonomic relationships within this family and lend support to the existing morphological phylogeny.

Fishes of the family Hexagrammidae (greenling) are endemic to the North Pacific Ocean. Generally, they occur in nearshore benthic communities (except *Pleuragrammus* species) around the North Pacific Ocean, from the Yellow Sea to Baja California, Mexico. *Hexagrammos agrammus* and *H. otakii* occur in the western Pacific. *Hexagrammos octogrammus, H. steller*, and the two *Pleuragrammus* species are more broadly distributed in the North Pacific. *Hexagrammos lagocephalus* occurs in the western North Pacific and *Hexagrammos superciliosus* occurs in the eastern North Pacific along with *Hexagrammos decagrammus* and *Ophiodon elongatus*. *Oxyelebus pictus, Zaniolepis latipinnis*, and *Z. frenata* have more restricted ranges off North America in the eastern Pacific.

Hexagrammid fishes are differentiated from other scorpaeniformes by the lack of spines on the upper part of the head (Mecklenburg and Eschmeyer, 2003). All species have fleshy cirri, or supra orbital flaps, except *Zaniolepis* species (Rutenberg, 1962). The genera *Oxyelebus* and *Zaniolepis* exhibit three to four prominent spines at the base of the anal fin, while the other species have rudimentary anal spines or none. All have etenoid scales except for *Ophiodon* which has cycloid scales (Mecklenburg and Eschmeyer, 2003). The scientific names for the genera *Hexagrammos* and *Pleuragrammus*, and many of the nominal species, refer to multiple lateral lines exhibited by members of this family. *Hexagrammos agrammus* is the easiest to characterize morphologically. Unlike other species in these genera, *H. agrammus* exhibit a single lateral line. All remaining members of the genera *Hexagrammos* and *Pleuragrammus* exhibit five paired lateral lines, numbered 1–5 from dorsal to ventral regions. *Pleuragrammus* species have a different lateral line configuration than *Hexagrammos* species. Species belonging to the genera *Ophiodon*, *Oxyelebus*, and *Zaniolepis* exhibit one paired lateral line.

Hexagrammid fishes exhibit several interesting reproductive natural history traits affecting their evolution and distribution including high dispersal potential, parental care, sexual selection, and hybridization. Multiple females deposit egg clutches in nests guarded by single males (Crow et al., 1997). The eggs incubate for approximately 20–22 days, but each species exhibits projected spawning seasons ranging from 6 to 12 weeks. Therefore the number of clutches may vary throughout the spawning season. The mating system exhibited by hexagrammids is conducive to the evolution of various sexual strategies centered on mate recognition. Females exhibit sexual selection by preferring males guarding established nests with one or more clutches (Crow et al., 1997), and males exhibit sneak spawning (Munehara et al., 2000; Munehara and Takenaka, 2000). These behaviors could lead to non-assortative mating and, indeed, three species of *Hexagrammos* fishes exhibit hybridization (Balanov and Antonenko, 1999; Balanov et al., 2001). Hybridization indicates incomplete reproductive isolation between these taxa. Thus, the mating system exhibited by hexagrammids ultimately affects mechanisms of speciation within the group (Crow, 2003).

Our specific objectives in this paper were to reconstruct a phylogeny of the family Hexagrammidae using multiple molecular loci, look for concordance with the morphological phylogeny, verify species status of hybridizing *Hexagrammos* species, evaluate the monophyly of the family, and understand the evolutionary history of these species and their distribution throughout the North Pacific.

## 2. Materials and methods

### 2.1. Sampling

Sampling of hexagrammid species for construction of a molecular phylogeny was completed in summer 1998, using trap, hook and line, or scuba. Fin clips, muscle or liver tissue was collected and preserved for DNA extraction. Specimens were not archived into museum collections. All hexagrammid species were sampled, except one (*Zaniolepis frenata*). Several individuals of western Pacific *Hexagrammos* species (*H. agrammus*, *H. otakii*, *H. octogrammus*, *H. stelleri*, and *H. lagocephalus*) and *Pleuragrammus* species were collected from various marine stations throughout Japan (Table 1). One to five samples of all remaining hexagrammid species were taken from Alaska (*H. octogrammus, H. stelleri, H. superciliosus*, and *P. monopterygius*) or sites near Monterey, California (*H. decagrammus, Ophiodon elongatus, Oxyelebus pictus*, and *Zaniolepis latipinnis*). Because hybridization occurs between three *Hexagrammos* species, individuals chosen for this study exhibited unambiguous morphological characteristics consistent with species descriptions. In the case of two hexagrammid species that closely resemble each other, *H. lagocephalus* and *H. superciliosus*, species identity was inferred from sampling location.

One to two individuals of seven cottids and one agonid (Table 1) were included in the phylogeny for comparison with the morphological phylogeny, and to test the
Table 1
Taxa represented in the phylogenetic analyses, abbreviation, number of individuals sampled, and corresponding sampling locations

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Abbreviation</th>
<th>Japan</th>
<th>Alaska</th>
<th>Oregon</th>
<th>California</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sebastes atrovirens</td>
<td>Sat</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Hexagrammos agrammus</td>
<td>Hag</td>
<td>1&lt;sup&gt;a&lt;/sup&gt;, 1&lt;sup&gt;b&lt;/sup&gt;, 4&lt;sup&gt;d&lt;/sup&gt;, 4&lt;sup&gt;d&lt;/sup&gt;</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hexagrammos otakii</td>
<td>Hot</td>
<td>1&lt;sup&gt;b&lt;/sup&gt;, 8&lt;sup&gt;c&lt;/sup&gt;, 3&lt;sup&gt;d&lt;/sup&gt;</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hexagrammos octogrammus</td>
<td>Hoc</td>
<td>1&lt;sup&gt;e&lt;/sup&gt;, 15&lt;sup&gt;d&lt;/sup&gt;</td>
<td>1&lt;sup&gt;f&lt;/sup&gt;</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hexagrammos stelleri</td>
<td>Hst</td>
<td>2&lt;sup&gt;a&lt;/sup&gt;, 1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>2&lt;sup&gt;g&lt;/sup&gt;</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hexagrammos superciliosus</td>
<td>Hsu</td>
<td>2&lt;sup&gt;e&lt;/sup&gt;, 3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hexagrammos lagocephalus</td>
<td>Hla</td>
<td>5&lt;sup&gt;a&lt;/sup&gt;, 1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hexagrammos decagrammus</td>
<td>Hde</td>
<td>3&lt;sup&gt;e&lt;/sup&gt;</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pleurogrammus azonus</td>
<td>Pza</td>
<td>2&lt;sup&gt;a&lt;/sup&gt;, 1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pleurogrammus monopterygius</td>
<td>Pmo</td>
<td>3&lt;sup&gt;e&lt;/sup&gt;</td>
<td>1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ophiodon elongatus</td>
<td>Oel</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Oxylebius pictus</td>
<td>Opi</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Zaniolepis latipinnis</td>
<td>Zla</td>
<td>3&lt;sup&gt;e&lt;/sup&gt;</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scorpaenichthys marmoratus</td>
<td>Sma</td>
<td>2&lt;sup&gt;e&lt;/sup&gt;</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arctedus carrallus</td>
<td>Aco</td>
<td>1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinocottus recalvus</td>
<td>Cre</td>
<td>2&lt;sup&gt;e&lt;/sup&gt;</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinocottus analis</td>
<td>Can</td>
<td>1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chitonotus pugetensis</td>
<td>Cpu</td>
<td>1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leptocottus armatus</td>
<td>Lar</td>
<td>1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhamphocottus richardsonii</td>
<td>Rvi</td>
<td>1&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jordania zonope</td>
<td>Jzo</td>
<td>1&lt;sup&gt;e&lt;/sup&gt;</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stellerina xyosterna</td>
<td>Sxy</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

* Nakajima Marine Biological Station (Ehime University).
  + Nota Marine Laboratory (Kanazawa University).
  + Tsukuri Marine Research Center (University of Tokyo).
  + Usui Marine Biological Station (Hokkaido University).
  + Hakodate fish market.
  +Kachemak Bay in the lower Cook Inlet.
  + Shemya Island, Aleutian archipelago.
  + Oregon Coast Aquarium.
  + Monterey Bay area.
  + Monterey Bay Aquarium.
  + Indicates individuals were sampled from an unknown location offshore.

hypothesis of monophyly for the family Hexagrammidae. **Sebastes atrovirens** was chosen as an outgroup because it represents a basal scorpaeniform (sensu Shinohara, 1994), ancestral to all families included in this study, namely the Hexagrammidae, Cottidae, and Agonidae.

2.2. DNA extraction and polymerase chain reaction amplification

Muscle or liver tissue was collected and preserved in 95% ethanol. Tissues were digested overnight at 55 °C in 500 µl salt extraction buffer (400 mM NaCl, 10 mM Tris, 2 mM EDTA, 1% SDS, and 20 µg/ml Proteinase K). DNA was purified by standard phenol:chloroform extraction and isopropanol precipitation (Sambrook et al., 1989).

Six genes including both mitochondrial and nuclear loci were amplified and sequenced for each species. Two mtDNA loci were amplified using the following primers: (1) cytochrome b, GLUDG-L and CB2H (Kocher et al., 1989); and (2) 16S ribosomal subunit, 16SAR and 16SBR (Kocher et al., 1989). Amplification of the four nuclear loci was accomplished with the following: (3) the first intron of the S7 ribosomal protein, S7RPEX1F and S7RPEX2R (Chow and Hazama, 1998); (4) the 4th intron of the Calmodulin gene, CALMex4F and CALMex5R (Chow, 1998); (5) an unspecified intron from the creatine kinase gene, ckF1, ckR1, and ckR2 (Quattro and Jones, 1999); and (6) an unspecified intron from the lactate dehydrogenase gene, LdhF1, LdhF2, and LdhR (Quattro and Jones, 1999). Each 50 µl reaction contained 10–100 ng of DNA, 10 mM Tris–HCl, pH 8.3, 50 mM KCl, 5 mM MgCl<sub>2</sub>, 1.5 U of *Taq* DNA Polymerase (Perkin–Elmer, Norwalk, CT), 0.25 mM dNTPs, and 0.3 µM each primer. Amplification cycling profiles were as follows:

1. 1st PCR: 1 min at 94 °C, 1 min at 48 °C, and 1 min at 72 °C for 35 cycles;
2. 2nd PCR: 1 min at 94 °C, 1 min at 48 °C, and 1 min at 72 °C for 35 cycles;
3. 1st PCR: 1 min at 94 °C, 1 min at 60 °C, and 2 min at 72 °C for 35 cycles;
4. 2nd PCR: 1 min at 94 °C, 1 min at 54 °C, and 1 min at 72 °C for 35 cycles.
directions with the primers used in the polymerase chain reaction (PCR) amplification on an ABI 373 automated sequencer (Applied Biosystems, Foster City, CA). Sequences for all loci were deposited in GenBank under the following accession numbers (AY582088-AY582108, cyt b; AY583110-AY583130, 16S; AY583169-AY583186, CaM; AY583187-AY583206, S7RP; AY583131-AY583148, ck; and AY583149-AY583168, Ldh).

2.3. Sequence analysis

Sequences were aligned using the computer program Clustal V (Higgins et al., 1992), implemented by Sequence Navigator (Applied Biosystems). Phylogenetic relationships were assessed by maximum parsimony (MP), neighbor joining (NJ), and Bayesian inference methods implemented by the software packages PAUP (Phylogenetic Analyses Using Parsimony version 4.0, Swofford, 1998) and MrBayes (version 2.1, Huelsenbeck and Ronquist, 2001). The most parsimonious trees were obtained using a heuristic search. Statistical confidence in nodes was evaluated using 2000 bootstrap replicates (Felsenstein, 1985; Hedges, 1992; Hillis and Bull, 1993) and homogeneity tests were performed using 1000 replicates. MrBayes default settings for the likelihood analysis were adopted including the GTR model (unequal base frequencies and six substitution rates), and statistical confidence in nodes was evaluated by Bayesian posterior probabilities. Stationarity of tree likelihood, sampled every 100 cycles, was consistently achieved after 3000 generations and all sampled trees preceding stationarity were discarded. Congruence of data sets was assessed using partition homogeneity tests. Alternate phylogenetic topologies were compared using permutation tests, and Kishino–Hasegawa tests (Kishino and Hasegawa, 1989; pairwise comparisons) or Shimodaira–Hasegawa tests (Shimodaira and Hasegawa, 1999; multiple comparisons), implemented by PAUP.

3. Results

3.1. Sequences

A total of 2013 bp were sequenced for the 21 species under investigation (1150 bp nuclear and 863 mtDNA). Of these, 720 were variable and 437 were parsimony informative (Table 2).

Few or no differences (i.e., 0–4 nucleotides), for all loci, were found between several individuals of the same species 3–12 individuals each for Hexagrammos otakii, H. agrannos, and H. octogrammos; one to five individuals each for H. laocephalus and H. superciliosus; and one to three individuals each of all remaining hexagrammids, cottids, the agonid Stellerina xyostoma, and the outgroup Sebastes atrovirens (Table 1).

Sequencing attempts were unsuccessful for Clinocottus recalvus in lactate dehydrogenase (Ldh); Sebastes atrovirens, Leptocottus armatus, and Stellerina xyostoma in Calmodulin (CaM); and Chitonotus pugentensis in S7 ribosomal protein (S7RP). A 313 bp portion of the S7RP intron was removed from the analysis because this portion of the sequence could not be unambiguously aligned in several taxa. Therefore, only the remaining 236 bp were retained for analysis. Removal of these sequences did not have a significant effect on S7RP tree topology (p = 0.108). The creatine kinase intron (ck) of Pleurogrammus azonus and P. monopterygius was not included in the analysis because it was much larger (750 bp as opposed to 249) and did not align with other taxa.

Mitochondrial loci exhibited no insertions or deletions (indels), but nuclear loci exhibited 6 (ck and Ldh), 7 (S7RP), and 13 (CaM) indels. These indels were included in the analyses and counted as one single mutation each, regardless of size. Their removal or inclusion did not result in significant differences in tree topologies for any independent locus, nor when data were

<table>
<thead>
<tr>
<th>Locus</th>
<th>Cyt b</th>
<th>16S</th>
<th>ck</th>
<th>Ldh</th>
<th>CaM</th>
<th>S-7 exon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragment length (bp)</td>
<td>396</td>
<td>467</td>
<td>254</td>
<td>268</td>
<td>392</td>
<td>236</td>
</tr>
<tr>
<td>Variable sites</td>
<td>156</td>
<td>105</td>
<td>73</td>
<td>99</td>
<td>159</td>
<td>128</td>
</tr>
<tr>
<td>Parsimony informative sites</td>
<td>129</td>
<td>67</td>
<td>43</td>
<td>45</td>
<td>78</td>
<td>75</td>
</tr>
<tr>
<td># Indels</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>6</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td># Base pairs in indels</td>
<td>0</td>
<td>0</td>
<td>46</td>
<td>142</td>
<td>125</td>
<td>19</td>
</tr>
<tr>
<td># Most Parsimonious trees</td>
<td>2</td>
<td>231</td>
<td>6</td>
<td>18</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>TS/TV Ratio</td>
<td>2.9</td>
<td>2.72</td>
<td>1.53</td>
<td>1.65</td>
<td>1.12</td>
<td>0.97</td>
</tr>
<tr>
<td>Weighting scheme</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1*</td>
<td>1*</td>
</tr>
<tr>
<td>p (weighted vs. not weighted)</td>
<td>0.8948</td>
<td>0.4928</td>
<td>0.318</td>
<td>0.8092</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The symbol "*" indicates weighting scheme not appropriate. The symbol "~" indicates the relationship was not upheld in a monophyletic clade.
combined. Plots of transitions and transversions versus genetic distance for independent loci indicated that saturation was not reached within ingroups (not shown). Saturation was only beginning to be apparent in outgroup comparisons for cyt b, 16S, and ck. Observed transition to transversion ratios varied from 0.97 to 2.9 (Table 2), and when appropriate weighting schemes were attempted, the resulting tree topologies remained unchanged. Therefore, all characters were considered equal in weight for the remainder of the analyses.

3.2. Tree topologies

Tree topologies from individual loci were similar, but not identical. To test the hypothesis that there were no differences in tree topologies produced from individual
loci, the Shimodaira–Hasegawa test was performed to compare all loci simultaneously. The results indicated significant differences for alternate topologies inferred from four of six loci. However, inclusion of cottids in the overall analysis was found to be the source of significant differences in tree topologies. When only hexagrammids were included in the analysis (i.e., species from the genera Hexagrammos, Pleurogrammus, Ophidon, Oxylebius, and Zaniolepis) and the cottids and agonid were excluded from this analysis, tree topologies from individual loci did not exhibit significant differences. Finally, when data sets from all loci were combined, the number of most parsimonious trees was reduced to one, more nodes were supported statistically, and bootstrap values for supported nodes increased. Even when data sets are incongruent, the best estimate of hierarchical relationships is still derived from parsimonious interpretation of all the data, analyzed simultaneously (Brower et al., 1996). For these reasons, all data sets were combined and the final overall topology is hereafter assumed to be the best supported. However, topologies for individual loci will be described and discussed separately. The relevance of using multiple loci for phylogenetic analysis in general, the loci particularly chosen for this study, and the significance of gene trees vs. species trees also will be addressed.

3.3. Phylogenetic analysis

A single most-parsimonious tree, identical to the neighbor-joining tree and to the maximum-likelihood tree, was obtained (tree length = 1475 steps, Fig. 1). The molecular data indicated that Hexagrammos agrammus and H. otakii were closest relatives with a sequence divergence of 1.78 and 100% bootstrap replicate support. Hexagrammos octogrammus was the sister taxon of this clade, with sequence divergences of 3.02 and 2.43%, respectively. The next closest relative was H. stelleri, followed by H. lagocephalus. Hexagrammos decagrammus was most ancestral within the genus (Fig. 1). The genus Pleurogrammus formed the sister clade of Hexagrammos. Ophidon was associated with cottids, and the zaniolepidids were distinct and basal to both hexagrammids and cottids.

The molecular tree was compared to the morphological tree presented by Shinohara (1994, Figure 2). Differences between the morphological and molecular topologies include the following: a clade with Hexagrammos octogrammus and H. lagocephalus was not supported by the molecular data, and H. stelleri was not supported as the most ancestral species within the genus (Fig. 2). The genus Ophidon was more closely associated with cottids than the genera Hexagrammos and Pleurogrammus. Permutation tests indicated that differences between the morphological and molecular tree topologies were highly significant ($p = 0.001$), and the Kishino–Hasegawa test indicated significant differences as well ($p = 0.0407$).

Several cottid species and one agonid (Stellerina xyosterna), putative sister families to hexagrammids, were included in this analysis to test the hypothesis of monophyly in the family Hexagrammidae. The genera

![Fig. 2. Comparison of morphological and molecular cladograms for the family Hexagrammidae and consensus cottid sequence. Morphological topology from Shinohara (1994). Sebastes atrovirens is used as an outgroup. Results of bootstrapping (2000 replicates) for MP/NJ are shown at nodes. Only values over 50% are shown.](image-url)
Hexagrammos and Pleurogrammus did form a monophyletic assemblage. However, the historical assemblage of hexagrammids including the monotypic genus Ophiodon and the genera Zaniolepis and Oxylebias did not form a monophyletic group. The closest relative of O. elongatus, from these data, was the cottid Scopraenichthys marmoratus, and these taxa exhibited 0% sequence divergence in two loci (ek and S7RP). An additional individual of each these two species was independently sampled and sequenced to rule out PCR contamination, confirming these results. Regardless of its closest relative, Ophiodon was not associated with the genera Hexagrammos and Pleurogrammus in a monophyletic assemblage. Oxylebias pictus and Z. latipinnis formed a monophyletic clade with 100% bootstrap support, outside the hexagrammids, cottids and the agonid with >95% bootstrap support. These data support a separate family Zaniolepididae, as suggested by Shinohara (1994).

4. Discussion

4.1. Systematics, comparison with morphological data, and monophyly of the family Hexagrammidae

Shinohara (1994) and Imamura and Shinohara (1998) presented a very thorough morphological phylogeny of the Scopraeniformes based on detailed descriptions of comparative osteology, myology, and cranial nerves including the suborders Hexagrammoidei and Cottoidei. Estimates of relationships within the genus Hexagrammos also included meristics (Shinohara, 1994). Because there were no cottid genera common to both studies, consensus data for cottids from each individual study were used for comparison (Fig. 2). The inclusion of cottids in both the morphological and molecular studies was necessary to test the hypothesis of monophyly in the family Hexagrammidae.

Historically the suborder Hexagrammoidei, including the genera Hexagrammos, Pleurogrammus, Ophiodon, Oxylebias, and Zaniolepis, had been considered a natural group. However both the morphological data (Imamura and Shinohara, 1998; Shinohara, 1994) and our molecular data support the exclusion of the genera Oxylebias and Zaniolepis, forming the basal suborder Zaniolepidoidei. The remaining hexagrammids were more closely related to cottids and together formed a monophyletic group.

4.2. Families

There were significant differences between the morphological and molecular data with respect to hexagrammid taxa. Our data diverged from the morphological data on two points: they did not support inclusion of the monotypic genus Ophiodon in the family Hexagrammidae, and there were two rearrangements in topology within the genus Hexagrammos. First, the molecular data placed Ophiodon within the Cottoidei with bootstrap values of 58, 78, for neighbor joining and maximum parsimony, respectively, and posterior probabilities of 100% for Bayesian inference. Therefore, even if one accepts the exclusion of the Zaniolepidae from the family Hexagrammidae, the remaining taxa still are not monophyletic. Interestingly, Ophiodon was very closely associated with the cottid Scopraenichthys marmoratus, and in fact they were identical at two loci. It is beyond the scope of this paper to reclassify these two genera, but these data suggest that Ophiodon should not be considered a hexagrammid. The fact that the molecular data grouped Ophiodon with cottids was surprising because such a relationship has never before been proposed in a phylogenetic hypothesis. However, early morphological descriptions of Ophiodon, Hexagrammos, and Scopraenichthys noted that some osteological (Gutberlet, 1915) and internal morphological (Allen, 1905) traits were more similar between Ophiodon and Scopraenichthys than Ophiodon and Hexagrammos.

Our data agree with Shinohara (1994) in that the Hexagrammidae (hereafter in this text referring to the genera Hexagrammos and Pleurogrammus) are closely related to cottids, and the zaniolepids are distinct. Perhaps a new suborder should be erected including Hexagrammos and Pleurogrammus, and another including Ophiodon, cottids, agonids, and anoplapomids (sablefish, a related cottoid included in Shinohara (1994)). However, the family Zaniolepididae, consisting of the genera Zaniolepis and Oxylebias, is entirely distinct, basal, and would be placed outside these putative suborders.

4.3. Genera

The topology of the genus Hexagrammos supports a phylogeographic model of evolution around the North Pacific along a longitudinal gradient, with the most ancestral Hexagrammos species distributed in the eastern North Pacific (H. decagrammus) and the most derived species occurring in the western North Pacific (H. otakii and H. agrammus). It appears that there may have been two vicariant events resulting in (1) the divergence between the genera Pleurogrammus and Hexagrammos, and (2) the divergence between H. agrammus and H. otakii from other Hexagrammos species. Hexagrammos agrammus was one time classified as a separate genus (Agrammus) because this species exhibits a single lateral line, unlike all other members of the genera Hexagrammos and Pleurogrammus. However, this appears to be due to a character reversal. Interestingly, lateral lines 1, 2, 4, and 5 are non-functional in H. decagrammus and H. stelleri (Wonseller and Webb, 1997).
4.4. Species

With respect to species within the genus *Hexagrammos*, hybridizing species *H. otakii*, *H. agrammurus*, and *H. octogrammatus* were genetically distinct and exhibited fixed differences at all mitochondrial and nuclear loci (i.e., no shared genotypes or haplotypes), therefore species status was upheld for these taxa. We found no evidence of introgression due to hybridization in the individuals we examined, but we purposely chose individuals that exhibited typical morphologies for each species. The potential for introgression and consequences of hybridization have not been characterized in this system.

The molecular data do not support the retention of *Hexagrammos superciliosus* as a separate species from *H. lagocephalus*. These species resemble each other closely, but were originally described as distinct species by Pallas (1810) based on differences in the size of the supraorbital flap, or cirri. The cirri of *H. lagocephalus* is described as a small, coarsely fimbriated lobe located posteriorly above the eye, whereas the cirri of *H. superciliosus* is long, flat, densely fimbriate, and with a length approximately equal to the eye diameter (Rutenberg, 1962). In addition, these species are described as having distinct geographical ranges, with *H. lagocephalus* considered an Asian species and *H. superciliosus* an American species. *Hexagrammos lagocephalus* is reportedly distributed off Northeastern Hokkaido, the Kurile Islands, and the Kamchatka peninsula. *Hexagrammos lagocephalus* is unknown off Alaska (Rutenberg, 1962) but a few individuals were reported off the Commander Islands and Attu Island (western tip of the Aleutians). However Rutenberg (1962) considers the species identity of these specimens questionable, and states that their species identity require further confirmation. *Hexagrammos superciliosus* is reportedly distributed off North America to the Aleutians. The individuals sampled in this study were collected from three distinct areas in the purported ranges of these species-*H. lagocephalus* was sampled off Northern Hokkaido, well within their reported range; and samples of *H. superciliosus* were obtained off Shemya Island, Alaska (near the end of their range) and Kachemak Bay (well within their reported range). Because these species do not overlap in geographical distribution, species identification was inferred from sampling location. Sequences of one to five individuals of each species were identical at three loci (*cytb*, *16S*, and *S7RP*), or exhibited 1 or 2 bp differences (*Ldh*, *CaM*, and *ck*). Therefore, these data support a single species (*H. lagocephalus*) with a broader distribution across the North Pacific from northern Hokkaido to California.

Nelson (1994) classified *Pleurogrammus monopterygius* and *P. azonus* as synonymous. However, our data support these taxa as separate species with as much, or greater, sequence divergence (7.93%) as other congeneric species comparisons. Together they formed a monophyletic clade with very high bootstrap support for the genus.

4.5. Multiple loci and gene trees

We constructed a phylogeny of the family *Hexagrammidae* based on six molecular loci including both nuclear and mitochondrial loci. The use of multiple loci improves the resolving power of the molecular assay (Avise, 2000) because the phylogenetic history revealed by a single locus may not reflect the true species phylogeny (i.e., gene trees vs. species trees, Brower et al., 1996; Maddison, 1997; Nichols, 2001; Pamilo and Nei, 1988; Zardoya and Meyer, 1996). To increase the probability of recovering the true species phylogeny, DNA sequences from several independent loci should be used (Avise, 1994; Brower et al., 1996; Harrison, 1998; Pamilo and Nei, 1988). However, topologies inferred from different individual loci may be incongruent due to a number of factors including differences in rates of evolution, number of variable and parsimony informative sites, base compositional bias, (Rokas et al., 2003), horizontal gene transfer, genome location (Machado and Hey, 2003; Rokas et al., 2003), and lineage sorting (Page and Charleston, 1997). It has become increasingly clear that using combined data for recovering species trees is far superior to using data from a single locus, and in certain circumstances concatenated data sets have a higher probability of recovering the true species phylogeny than a consensus tree (Dequeiroz, 1993; Dequeiroz et al., 1995; Rokas et al., 2003; Wiens, 1998). While this is not intended to be a review of relevant factors associated with the latter debate (combined vs. consensus), we base our argument that a concatenated data set is superior for our data set for several reasons. First, topologies were not incongruent with respect to hexagrammids (i.e., when cottsids and the agonid were removed from the analysis), therefore combining the data is the best way to extract the most information contained in those data. Second, information is lost in data sets combined into a consensus tree and the signal to noise ratio is expected to decrease. Indeed, several polytomies where introduced when we constructed a consensus tree (with cottsids included), and the resulting consensus tree was therefore less informative and less interesting.

Our results indicated that if the hexagrammid phylogenetic reconstruction were based on a single locus, slightly different topologies would have been recovered, with reduced bootstrap support. For example, *16S* and *CaM* inferred topologies somewhat similar to the morphological data alone. The *Ldh* locus inferred an alternative topology with a polytomy for five *Hexagrammos* species, and a clade with *H. decagrammatus* associated
with the genus *Pleurogrammos*. The remaining loci sequenced in this study—Cyt b, S7RP, and ck-inferred topologies similar to the combined data, with varying degrees of bootstrap support. However, the most surprising result from these data was the exclusion of the genus *Ophiodon* from any association with the genera *Hexagrammos* and *Pleurogrammus*, and this result was consistent in the topologies inferred from each individual locus. Furthermore, the close association of *Ophiodon* with the cotti *Scorpaenichthys* was maximally supported in the combined data.

Finally, gene trees could differ from a true species phylogeny due to horizontal gene transfer via hybridization (Machado and Hey, 2003; Maddison, 1997; Rokas et al., 2003). While hybridization is known to occur between *Hexagrammos octogrammus* and *H. agrammus* (Balanov and Antonenko, 1999) or *H. otakii* (Balanov et al., 2001), we found no evidence of introgression or horizontal gene transfer in sequences from 10 to 17 individuals included in this study. Therefore, horizontal gene transfer did not appear to be a source of gene–gene discordance, and in fact tree topologies were not incongruent with respect to hexagrammids.

Overall, these data illustrate the utility of using a multi-locus approach when making phylogenetic inferences based on molecular data. When data from multiple loci were combined, the number of most parsimonious trees was reduced to one, more nodes were supported statistically, and bootstrap values for supported nodes increased to highly significant levels.

### 4.6. Loci chosen for this study

There are advantages to using sequences from both mitochondrial and nuclear DNA for recovering species phylogenies. Mitochondrial DNA has the benefit of containing orthologous, single-copy genes (Page, 2000). It evolves faster, lacks recombination and segregates independently, so topologies based on mtDNA have a higher probability of accurately tracking short internodes in species phylogenies (Moore, 1995). Therefore, mitochondrial and nuclear loci offer different windows of resolution for phylogenetic tree reconstruction (Creer et al., 2003; Page, 2000; Weibl and Moore, 2002). In this study, mtDNA sequences were better at resolving species relationships within genera, but deeper nodes were not as well supported in analyses from mtDNA alone. Finally, combining sequences from both nuclear and mitochondrial sequences has been shown to increase topological support in constructing phylogenies (e.g., Creer et al., 2003).

### 4.7. mtDNA

Cytochrome *b* sequences are the most ubiquitous source of data for comparative studies of fish sibling species (McCune and Lovejoy, 1998) and for vertebrates in general (Avise and Johns, 1999). The topology suggested by cytochrome *b* sequences in this study was identical to the overall topology with respect to the hexagrammids. These data supported monophyly of the genera *Hexagrammos* and *Pleurogrammus*, but did not support the inclusion of *Ophiodon* or the zaniolepids in the family Hexagrammidae. Cytochrome *b* sequences resolved species relationships for the genera *Hexagrammos* and *Pleurogrammus* with high bootstrap support. Furthermore, cytochrome *b* has the utility of a calibrated molecular clock for fishes (approximately 1.5–2.5% sequence divergence/million years, Meyer et al., 1990). These sequences suggest that sister species *H. otakii* and *H. agrammus* diverged over one million years ago (2.85% sequence divergence), and that these Japanese species diverged from other *Hexagrammos* species approximately 2.2–3.6 million years ago. The genus diverged from *Pleurogrammus* approximately 6 million years ago (~15% sequence divergence).

Sequences of the structural gene 16S ribosomal subunit have been commonly used for phylogenetic reconstruction. We used 16S to increase the sample size of mtDNA, and to compare with cytochrome *b*. Interestingly, cytochrome *b* and 16S are inherited as a single linkage group, yet they suggest different rates of evolution in hexagrammids. The 16S locus provided the least support for any phylogenetic structure, supporting few nodes, and producing 231 most parsimonious trees (Table 2). Sequences of 16S were not as informative as cytochrome *b* and variation was not sufficient to resolve relationships within the genera *Hexagrammos* and *Pleurogrammus*. Monophyly of the genera *Hexagrammos* and *Pleurogrammus* was not supported, nor was the association of *Ophiodon* with the family Hexagrammidae. However, the zaniolepids were supported as monophyletic and distinct from the hexagrammids and cottids.

### 4.8. Nuclear DNA

Because introns are non-coding regions of DNA, they are presumably released from selective constraint and therefore are useful indicators of evolutionary processes consistent with the neutral model of evolution. The introns chosen for this study correspond to single-copy loci (Chow and Hazama, 1998; Quattro and Jones, 1999) and should not exhibit the confounding effects of paralogy, or gene duplication (however, the CaM locus corresponded to a single-copy locus in seven of eleven species, further investigation confirming the presence of paralogs in the remaining four taxa was not performed, Chow, 1998). In our data, sequences of nuclear introns were better at resolving deeper nodes and were useful in suggesting placement of the zaniolepids, *Ophiodon*, and for testing the hypothesis of monophyly for the
family Hexagrammidae. The percentages of phylogenetically informative sites in these introns ranged from 17 to 20%, and 31% for the retained portion of the S7RP locus. One limitation associated with intron sequences is high levels of indel variation (Beltran et al., 2002; Quattro et al., 2001). The introns used in this study did contain significant proportions of indels, ranging from 8 to 53% of the nucleotides analysed. Their inclusion made alignment difficult for some taxa, (e.g., the ch intron for Pleurogrammus azonus and P. monopterygius were much larger and could not be aligned with other taxa, and the CaM sequence for the outgroup Sebastes atrovirens was highly ambiguous). However, these differences among taxa were phylogenetically informative, and were therefore included in the analyses.

Sequence alignments for the S7RP locus in Hexagrammos and Pleurogrammus taxa were unambiguous. However, sequences for the outgroup Sebastes atrovirens and the zaniolepids (Oxylebius pictus and Z. latipinnis) were problematic after 236 bp at the 5′ end. Therefore, only this region was retained for further analyses. These sequences exhibited few indels representing 8% of the base pairs analyzed. There was strong support for the monophyletic association of the genera Hexagrammos and Pleurogrammus, a distinct Zaniolepididae, and a close association of Ophiodon with the Cottid Scophaenichthys marmoratus.

Although the calmodulin intron exhibited 32% indels, it did estimate hexagrammid relationships with >50% bootstrap support at most nodes. The creatine kinase locus exhibited 18 and 17% indels and phylogenetically informative sites, respectively. These data performed well in estimating deeper nodes with strong bootstrap support.

The lactate dehydrogenase intron was the poorest at resolving relationships within the hexagrammids, or cottids represented, and exhibited 53% indels with 17% parsimony informative sites (Table 2).

In summary, the inclusion of four nuclear loci increased the sample size, making it comparable with mtDNA, and greatly increased our confidence in recovering the correct topology. Combining mitochondrial and nuclear markers expands the potential for resolving deep nodes as well as more recently divergent taxa.

5. Conclusions

In conclusion, we constructed a molecular phylogeny of the family Hexagrammidae, and related taxa, using a multiple locus approach. With respect to the genus Hexagrammos, our phylogenetic hypothesis supports a phylogeographic model of evolution around the North Pacific along a longitudinal gradient, with the most ancestral Hexagrammos species distributed in the eastern Pacific (H. decagrammus) and the most derived species occurring in the western Pacific (H. otakii and H. agrammus). The genus Pleurogrammus was associated with Hexagrammos in a monophyletic clade. The genus Ophiodon grouped with cottids, and was very closely associated with the cottid Scophaenichthys marmoratus. The genera Oxylebius and Zaniolepis formed a monophyletic clade that was distinct and basal to all other hexagrammids, seven cottids, and one agonid.

A major difference in the molecular based topology, from historical hexagrammid systematics based on morphological data, was the exclusion of the genus Ophiodon. Other differences included rearrangements in the placement of H. stelleri and H. lagocephalus within the genus Hexagrammos. Our data concur with the morphological data with respect to exclusion of the zaniolepids from the Hexagrammidae.

Our data illustrate the utility of using a multi-locus approach when making phylogenetic inferences based on molecular data. In this study, the topologies inferred from individual loci differed, but when all data were combined the number of most parsimonious trees was reduced to one, more nodes were supported statistically, and bootstrap values for supported nodes increased to highly significant levels.

Acknowledgments

We thank the following agencies and programs for financial support: National Science Foundation Dissertation Enhancement Grant (INT0123805), National Science Foundation-Monbusho Summer Program, Sigma Xi-Grants in Aid of Research, Earl H. Myers and Ethel M. Myers Oceanographic and Marine Biology Trust. Several people helped obtain samples for this phylogeny. Dr. Hiro yuki Munehara provided samples of Hexagrammos stelleri, H. lagocephalus, and Pleurogrammus species from Usujiri Marine Biological Station; Brenda Konar and Jared Figurski obtained samples of H. octogrammus and H. superciliosus from Alaska. Pete Dal Ferro provided samples of Ophiodon elongatus and Scophaenichthys marmoratus from Monterey, California. We thank the Monterey Bay Aquarium and Oregon Coast Aquarium for providing samples of Jordania zonope, and Rhampocottus richardsonii, respectively. Salvador Sanchez and Machiko Kanamoto helped collect samples in Japan, and provided logistical support in the field.

References


