Are *Hypomesus chishimaensis* and *H. nipponensis* (Osmeridae) Distinct Species? A Molecular Assessment Using Comparative Sequence Data from Five Genes

KATRINA L. ILVES AND ERIC B. TAYLOR

We used mitochondrial (cytb, 16S) and nuclear (ITS2, S71, RAG1) sequence data to assess the degree of genetic differentiation between the recently described lacustrine *Hypomesus chishimaensis* from the Kuril Islands and its anadromous congener *H. nipponensis* from Hokkaido, Japan. The results showed that the two species are not genetically distinct at these loci. Phylogenetic trees depicted an unstructured clade consisting of individuals of both species and sequence data showed haplotypes that are shared between putative species for all markers. We conclude that *H. chishimaensis* is an alternative life-history type of *H. nipponensis* that should not be recognized at either the species or the sub-species level.

THE phenomenon of multiple ecotypes within species is well documented in many north temperate fishes (Behnke, 1972; Bernatchez and Wilson, 1998; Taylor, 1999). Although molecular analyses have provided great insight into the number of origins of ecotypes in many groups (Bernatchez and Dodson, 1990; Taylor and Bentzen, 1993; Douglas et al., 2005), there is ongoing debate as to whether life-history variants of even well studied fishes should be classified as distinct taxa (Behnke, 1972; Taylor, 1999). For those groups where relatively few data are available, it is even more likely that current taxonomic designations incorrectly or incompletely classify diversity (Wheeler et al., 2004; Wilson, 2003). Here we present an analysis of a recently described lacustrine smelt, *Hypomesus chishimaensis* (Osmeridae), and its anadromous congener, *H. nipponensis*. Saruwatari et al. (1997) identified *H. chishimaensis* as a new species from lakes on Kunashir and Iturup, the southernmost main-chain Kuril Islands of Japan. The new species was described as differing from *H. nipponensis* by having small teeth in the middle of the posterior portion of the glossohyal, heavier pigmentation of the body, and larger eye diameter. Subsequent studies at four Kuril Island lakes (Sidorov and Pichugin, 2004) and in Peter the Great Bay (Chereshnev et al., 2001), however, failed to detect morphological differences between the two putative species, leading the authors to conclude that *H. chishimaensis* is best regarded as an ecotype of *H. nipponensis*. Given the common occurrence of cryptic species that are reproductively or genetically distinct but indistinguishable morphologically (Mayr, 1948; Martin and Bermingham, 2000; Arnegard et al., 2005), we used mitochondrial (mtDNA) and nuclear (nDNA) gene regions to test the hypothesis that *H. chishimaensis* and *H. nipponensis* are separate species. Following the flow-chart proposed by Wiens and Penkrot (2002) for assessing species boundaries with molecular phylogenetic data, species designation of *H. chishimaensis* would be supported if the individuals assigned to this species based on morphology form a well-supported monophyletic group separate from the individuals of *H. nipponensis*.

**MATERIALS AND METHODS**

**Taxon sampling and molecular analysis.**—Morphologically defined samples representing *Hypomesus nipponensis*, *H. chishimaensis*, *H. olidus*, and the outgroup *Mallotus villosus* were analyzed for two mitochondrial (cytochrome *b* [cytb] and 16S rRNA [16S]) and three nuclear (internal transcribed spacer 2 [ITS2], S7 ribosomal protein, intron 1 [S71], and recombination-activating gene 1 [RAG1]) gene regions. Thirteen samples of *H. chishimaensis* were from the two southern-most main-chain Kuril Islands, Kunashir (3) and Iturup (4), and one of the Habomai group islands, Zelionyi (3), of Japan, and Sakhalin Island (3), Russia. Six *H. nipponensis* samples came from Hokkaido, Japan, and four *H. olidus* individuals from Kamchatka, Russia (Fig. 1). The outgroup *M. villosus* sample was obtained from Tribune Channel, British Columbia, Canada. *Hypomesus olidus*, the most closely related species to *H. chishimaensis* and *H. nipponensis* (K. Ilves, unpubl. data), was included in the study to ensure the markers used are appropriate for distinguishing well-recognized species. The outgroup, *Mallotus villosus*, is the basal taxon most closely related to *Hypomesus* (K. Ilves, unpubl. data).

DNA was extracted from ethanol-preserved or frozen tissue using either the Gentra PUREGENE® DNA Purification Kit or the Qiagen DNeasy®...
Tissue Kit following the manufacturer’s instructions. Amplification reactions for all markers contained 50–300 ng of genomic DNA template. Primers used for the amplification of cyt* (cyt*b, GluDG), 16S (16Sar, 16Sbr), ITS2 (5.8sr, 28s), S71 (S7RPEX1F, S7RPEX2R), and RAG1 (RAG1F, RAG1R) were obtained from Kocher et al. (1989), Waters et al. (2002), Presa et al. (2002), Chow and Hazama (1998), and Quenouille et al. (2004), respectively. Cyt*b was amplified in 50-μl reactions containing final concentrations of 800 μM of dNTPs, 800 nm of each primer, 1 unit of Invitrogen™ Taq DNA polymerase, 20 mM Tris-HCl (pH 8.4), 50 mM KCl, and 2.5 mM of MgCl₂ under the following conditions: 95 C for 3 min, 52 C for 1 min, and 1 min at 72 C, followed by 5 cycles of 1 min each of 94 C, 52 C, and 72 C, and 30 30-sec cycles of 92 C, 55 C, and 72 C. A final extension step was performed at 72 C for 5 min. A protocol modified from Waters and Cambray (1997) was used to amplify 16S: 5 cycles of 94 C (30 sec), 61 C (1 min), 72 C (2 min) followed by 35 cycles of 30 sec each at the same temperatures. The procedure for ITS2 was modified from Presa et al. (2002): 95 C for 5 min, 5 cycles of 95 C (1 min 30 sec), 60 C (2 min), 72 C (3 min) followed by 35 cycles of 30 sec each at the same temperatures. For some samples a 55 C annealing temperature was used. Procedures for amplifying S71 and RAG1 were unmodified from Chow and Hazama (1998) and Quenouille et al. (2004), respectively. PCR products were purified using the Qiagen QIAquick™ PCR Purification Kit and sequenced at the NAPS Unit at the University of British Columbia on an ABI Prism 377 automated sequencer. All sequences have been deposited in GenBank (see Material Examined for accession numbers).

Phylogenetic analysis.—Sequences were aligned using ClustalX (Thompson et al., 1997) or manually with MacClade (v4.06, D. R. Maddison and W. P. Maddison, Sinauer Associates, Inc., Sunderland, MA, 2003) and edited with Se-Al (v2.0a11 A. Rambaut, 1996, http://evolve.zoo.ox.ac.uk/) or MacClade. The alignments for cyt*b, 16S, and RAG1 were unambiguous. For extra confidence in the 16S rRNA alignment, a secondary structure model presented by Waters et al. (2002) was followed, showing that the few indels occurred within loops. There were many indels in the alignments of ITS2 and S71, making positional homology uncertain in several locations; therefore, 80 and 21 characters were excluded from the two alignments, respectively. PAUP* (v.4.0b10 D. L. Swofford, PAUP*: phylogenetic analysis using parsimony [*and other methods], Sinauer Associates, Inc., Sunderland, MA, 2002) was used to calculate pair-wise distances and perform maximum likelihood (ML) analyses, and MrBayes v3.1.1 (Huelsenbeck and Ronquist, 2001) was used for the Bayesian estimates of phylogeny.

Separate analyses were conducted by marker as well as by combined mitochondrial, combined nuclear, and all markers combined. Modeltest v3.6 (Posada and Crandall, 1998) was run for all data partitions (individual and combined) to select a model of sequence evolution for use in the ML and Bayesian analyses. Posada and Buckley (2004) suggest that the hierarchical Likelihood Ratio Tests (hLRT) implemented in previous versions of Modeltest are inferior to the Akaike Information Criterion (AIC) methods of model testing. We therefore implemented the model chosen by AIC in our analyses. For the combined analyses this model corresponded to TVM + I (mtDNA), GTR + G (nDNA), and GTR + I (all markers).

Bayesian analysis was conducted on all sequences for each marker separately and for the combined datasets with a reduced number of individuals (see below) using the general model chosen by the AIC method in Modeltest, allowing MrBayes to calculate the exact parameter values. Two parallel analyses were run for 5 × 10⁶ generations with four MCMC chains, a sample frequency of 100 and a burnin of 5,000. For ML analyses on the combined data sets, heuristic searches were conducted with five random replicates of stepwise taxon addition. Confidence in groupings was assessed using bootstrap pseudo-replicates (1,000 for all data combined and 300 for the mtDNA and nDNA partitions), retaining only groupings that appeared with a frequency of at least 50%. We considered Bayesian posterior probabilities of at least 95%
(Huelsenbeck and Ronquist, 2001) and bootstrap values of at least 70% (Hillis and Bull, 1993) to indicate well-supported nodes.

**RESULTS**

**Sequences.**—The individual marker Bayesian analyses with all sequences showed that all *H. solidus* individuals either had identical sequences or clustered in 99–100% of the trees, therefore a single individual was used for subsequent combined ML and Bayesian analyses. Similarly, because preliminary Bayesian analysis of each marker with all sequences resulted in a single unstructured clade of *H. chishimaensis* and *H. nipponensis* individuals in each case and little variation was found among the *H. chishimaensis* individuals (with one variant for cyt* b* and ITS2, two for 16S and RAG1, and four for S71), six were included for the final combined analyses. These six individuals had sequences for all five markers and encompassed all of the haplotype diversity apart from one 16S variant. The other seven individuals had sequences for at least one mitochondrial and two nuclear loci. Since they shared haplotypes at these three loci with the other *H. chishimaensis* individuals, obtaining sequences for the remaining two would not have provided any additional information and was therefore deemed unnecessary. All six individuals of *H. nipponensis* were included for the final analyses, with two missing a RAG1 sequence. These final analyses included 15 sequences (14 for RAG1) ranging from 286–425 bp (cyt* b*), 529–548 bp (16S), 339–441 bp (ITS2), 338–678 bp (S71), and 1254–1431 bp (RAG1). Combined mitochondrial, nuclear, and all data analyses contained 973, 2,590, and 3,563 characters, respectively, including indels.

Uncorrected pair-wise differences between the samples show that the markers used in this study have varying degrees of divergence, with greater differences within cyt* b*, ITS2, and S71 than the more slowly evolving 16S and RAG1. Furthermore, for each marker at least one individual assigned to *H. chishimaensis* has an identical sequence to an individual assigned to *H. nipponensis*.

**Phylogenetic analyses.**—The ML and Bayesian analyses of all markers yielded a phylogeny where all *H. chishimaensis* and *H. nipponensis* fell into a single, unstructured monophyletic group (Fig. 2). The individual marker Bayesian analyses produced the same topology. The only slight variations on this topology were with the ML analysis of mtDNA and the ML and Bayesian analyses of nDNA, which added some poorly supported structure to the *H. chishimaensis/H. nipponensis* grouping. In the mtDNA ML phylogeny, one individual assigned to *H. nipponensis* fell outside the rest of the group with a bootstrap support of 69%, while in the nDNA analyses two individuals assigned to *H. chishimaensis* grouped outside the other sequences with bootstrap support and Bayesian posterior probabilities of 54 and 51%, respectively (data not shown). These values fall below accepted standards for supported clades (Hillis and Bull, 1993; Huelsenbeck and Ronquist, 2001), and it is clear that even with some structure in the group, there is no differentiation between the individuals named *H. chishimaensis* and those named *H. nipponensis* (Fig. 2).

**DISCUSSION**

Analyses of five molecular gene regions, both individually and in combination, show no genetic divergence between the newly described smelt, *Hypomesus chishimaensis*, and its congener, *H. nipponensis*. These results together with Wiens
and Penkrot’s (2002) framework for delimiting species, and the apparent lack of morphological divergence (Chereshnev et al., 2001; Sidorov and Pichugin, 2004), strongly suggest that these two named species are, in fact, conspecific.

Relevance of geologic history of Kuril Islands.—After the formation of the southernmost main-chain islands, Kunashir and Iturup, in the Pliocene and Pleistocene, respectively (Pietsch et al., 2001), major sea-level regressions in the late Pleistocene connected the terrestrial habitats of the Kuril archipelago, Hokkaido, and Kamchatka, when Hokkaido and Sakhalin were attached to mainland Asia (Briggs, 1974; Pietsch et al., 2001). The most recent regression occurred between 10,000 and 30,000 years ago (Pietsch et al., 2001). Assuming *H. nipponensis* invaded the freshwaters of some of these islands prior to this last sea-level regression, the period of isolation resulting from this drop in sea-level may have been too short for detectable morphological and molecular divergence to develop (Taylor, 1972).

**Taxonomic status of Hypomesus chishimaensis.**—Although there has been relatively little evolutionary time for divergence to occur between the populations assigned to *H. chishimaensis* and *H. nipponensis*, this explanation for why we see no genetic (or morphological) differences is incomplete. Differentiation into separate ecotypes on comparable timescales is a fairly widespread phenomenon in north temperate fish faunas (Bernatchez and Wilson, 1998; Taylor, 1999), with examples from smelts (Taylor and Bentzen, 1993), sticklebacks (Taylor and McPhail, 1999), salmonids (Bernatchez and Dodson, 1990), and lampreys (Sulewski, 2003). In each of these examples there is a considerable body of evidence for quantifiable morphological, ecological, behavioral and/or genetic differences, all of which are lacking for *H. chishimaensis*. A further complication, relevant to many of these other cases, is the possibility that the freshwater populations reflect independent evolution of each lake-resident population (Sidorov and Pichugin, 2004). The problem of how to deal with a similar ecological form that has evolved in separate geographic locations has caused taxonomic difficulties in many groups. In Osmeridae in particular, genetic data showed that assigning a single taxonomic name (*Osmerus spectrum*; Lanteigne and McAllister, 1983) to a ‘dwarf’ version of the rainbow smelt *O. mordax* did not accurately reflect its multiple-origin evolutionary history (Taylor and Bentzen, 1993).

The assignment of a Linnaean binomial to a taxon has significant theoretical and practical implications in terms of cataloging, understanding, and protecting biodiversity (Mayr, 1963; Agapow et al., 2004; Isaac et al., 2004); therefore, the process of assigning and retracting such designations is not a trivial matter. Given the lack of morphological and genetic divergence between the populations assigned to *Hypomesus chishimaensis* and *H. nipponensis*, we feel that Chereshnev et al. (2001) and Sidorov and Pichugin (2004) were correct in concluding that *H. chishimaensis* is a lake-resident life-history type of *H. nipponensis* that does not merit recognition at the species level. Further, assigning the freshwater populations as a subspecies of *H. nipponensis* is not justified because there are no morphological, genetic, or ecological traits that distinguish them, and further, each lake population may have had an independent origin. Although the freshwater populations are not a separate species from the anadromous *H. nipponensis*, there remains a clear need for basic ecological research on these lacustrine populations. In particular, investigations to distinguish between single- and multiple-origin hypotheses would further clarify their evolutionary history.

**Material Examined**

Institutional abbreviations follow Leviton et al. (1985). GenBank accession numbers for cyt *b*, 16S, ITS2, S71, and RAG1 are in parentheses.

ACKNOWLEDGMENTS

For sending specimens or tissue samples we thank B. Urbain and T. Pietsch for *H. chishi-*maensis and *H. olidus*, with much consideration to K. Maslenikov, S. Mori for *H. nipponensis*, and A. Morton for *Mallotus villosus*, *Hypomesus chishi-*maensis and *H. olidus* tissue samples from the University of Washington were collected with support from the Biological Sciences Directorate (Biotic Surveys and Inventories Program) and the International Program Division of the U.S. National Science Foundation (grants DEB-9400821 and 9505031, T. Pietsch, principal investigator) and by the Japan Society for the Promotion of Science (grant no. BSAR-401, Kunio Amaoka, principal investigator). M. Younger deserves credit for some initial DNA extractions and cyt* b* sequences and J. McLean for helping secure some of the samples. For sequencing, M. van der Geest and L. O’Hara were particularly helpful as was J. Witt for insightful discussions about phylogenetic analysis. Drafts of the manuscript were greatly improved by comments from R. de Graaf, J. Gow, J. McLean, P. Tamke, W. Tymchuk, J. Weir, and J. Witt. This study was made possible by funding from a UGF from the University of British Columbia awarded to KLI and Natural Sciences and Engineering Research Council of Canada Discovery Grant awarded to EBT.

LITERATURE CITED


