Canadian Journal of Zoology
Revue canadienne de zoologie

Phylogeography of the longnose dace (*Rhinichthys cataractae*) species group in northwestern North America – the Nooksack dace problem

Journal:	<i>Canadian Journal of Zoology</i>
Manuscript ID:	cjz-2015-0014.R1
Manuscript Type:	Article
Date Submitted by the Author:	19-May-2015
Complete List of Authors:	Taylor, Eric; University of British Columbia, Zoology McPhail, John; University of British Columbia, Zoology Ruskey, Jennifer; University of British Columbia, Zoology
Keyword:	BIODIVERSITY < Discipline, BIOGEOGRAPHY < Discipline, SPECIATION < Discipline, TAXONOMY-SYSTEMATICS < Discipline, FRESHWATER < Habitat, CYPRINIDAE < Taxon

SCHOLARONE™
Manuscripts

Phylogeography of the longnose dace (*Rhinichthys cataractae*) species group in northwestern North America - the Nooksack dace problem

Eric B. Taylor*

J.D. McPhail

and

J.A. Ruskey

Department of Zoology, Biodiversity Research Centre
and Beaty Biodiversity Museum
University of British Columbia
6270 University Blvd.,
Vancouver, BC, Canada V6T 1Z4
etaylor@zoology.ubc.ca

*Corresponding author

Phylogeography of the longnose dace (*Rhinichthys cataractae*) species group in northwestern North America – the Nooksack dace problem**E.B. Taylor, J.D. McPhail, and J.A. Ruskey**

Abstract: The longnose dace (Cyprinidae: *Rhinichthys cataractae* (Valenciennes 1842)) is one of the most widespread freshwater fishes in North America and across this range there have been several divergent forms described that are of uncertain taxonomic status. One of these forms, the Nooksack dace, is found in southwestern British Columbia and adjacent portions of western Washington State, and is distinguished from longnose dace by a lower number of lateral line scales. We sequenced a total of approximately 1,400 base pairs of mitochondrial DNA and found that the longnose dace found west of the continental divide and Nooksack dace constituted reciprocally-monophyletic clades that differed from each other by between 2 and 3% sequence divergence. Sequence analysis at two nuclear loci (the S7 ribosomal protein and recombination activation gene 1 introns, ~1,600 base pairs), however, showed no consistent difference between longnose dace and Nooksack dace and several alleles were shared between them. By contrast, consistent differences both at mitochondrial and nuclear DNA loci were resolved between *R. cataractae* samples from east and west of the Continental Divide. The Nooksack dace does not appear to warrant separate taxonomic status from the longnose dace, but the mitochondrial DNA differences support its recognition as an important component of the evolutionary and biogeographic legacy of *R. cataractae*.

Keywords : phylogeography, cryptic species, Chehalis fauna, *Rhinichthys*, DNA sequencing

Introduction

Cryptic species are defined as those that are difficult or impossible to distinguish by morphology, and which may have been incorrectly lumped as a single species (Sáez and Lozano 2005). This does not mean that they are literally identical; cryptic species may be morphologically distinguishable with the use of tools such as scanning electron microscopy and statistical morphometrics, and may rely upon non-visual cues to identify conspecifics (e.g. Feulner et al. 2006). Cryptic species may signal general morphological stasis: species diversification, decoupled from ecological and/or morphological change (Bond et al. 2001). The rate of identification of cryptic species has increased enormously since the advent of the polymerase chain reaction and affordable DNA sequencing and they are an important component of taxonomically un-described biodiversity (Bickford et al. 2007). The discovery of cryptic species can have important conservation implications: in many cases, a species thought to have a wide distribution has been shown to be a cryptic species complex, with each species having a very small distribution and therefore being at much greater risk of extinction (e.g. Ravaoarimanana et al. 2004). Detection of cryptic species has increased diversity estimates within some groups by over 100% (e.g. Funk et al. 2012) and can, therefore, have a huge effect on biodiversity estimates, making understanding the distributions of cryptic species important for identifying global biodiversity hotspots (Trontelj and Fišer 2008).

The genus *Rhinichthys* or “rifle daces” (Pisces: Cyprinidae) contains seven to nine species endemic to North America. The genus has a curious distribution with one widespread species, the longnose dace (*R. cataractae* (Valenciennes 1842)) found from northern Mexico, north to Labrador in the east and the lower Mackenzie River in the

northwest, and south to the middle Columbia River in Oregon (Lee et al., 1980). The other species have much more limited distributions: the blacknose dace (*R. atratulus* (Hermann 1804)) from the Atlantic coastal drainages in Georgia northwest to the Nelson River system in Manitoba, the speckled dace (*R. osculus* (Girard 1846)) from western coastal and interior drainages in Sonora, Mexico, north to the Kettle River (Columbia River) in Canada, the leopard dace (*R. falcatus* (Eigenmann and Eigenmann 1893)) from the Columbia and Snake River systems in the United States to the upper Fraser River system in Canada, the Umatilla dace (*R. umatilla* (Gilbert and Evermann 1894)) endemic to the Columbia River system, in Canada and the US, the Umpqua dace (*R. evermanni* Snyder 1908) has the smallest range and is endemic to the Umpqua River and tributaries in coastal Oregon, and the loach minnow (*R. cobitis* (Girard 1856)) is found in streams and small rivers throughout the Gila River and San Pedro River systems in Arizona, New Mexico, and Sonora (Lee et al. 1980). A potential eighth species, the western blacknose dace (*R. obtusus* Agassiz 1854) is of uncertain validity (see Fraser et al. 2005; Kraczkowski and Chernoff 2014) and a ninth taxon is extinct, the Las Vegas dace (*R. deaconi* Miller 1984). The species have been considered to comprise three species groups: the longnose dace group (longnose and Umpqua dace), the blacknose dace group (the blacknose dace), and the speckled dace group (speckled, leopard, and Las Vegas daces, e.g. see Hubbs et al. 1974 and Miller 1984). Given that the longnose dace group has such a wide distribution, it is surprising that only two species are recognized within this species group – the widely distributed longnose dace and the much more narrowly-distributed Umpqua dace.

The presence of cryptic variation and its taxonomic implications have long been recognized as issues within the *R. cataractae* species group (e.g., Hubbs 1926; McPhail

1967; Bartnik 1972; Kim and Conway 2014). For instance, McPhail (1967) first noted the occurrence of a divergent form of the longnose dace, commonly referred to as the “Nooksack dace”, from western Washington State and southwestern British Columbia which consistently differed from the longnose dace in scale count traits (see also Bisson and Reimers 1977; Ruskey and Taylor 2015). The Nooksack dace has a horseshoe-shaped geographic distribution that is centered around the Chehalis River of western Washington State (Fig. 1). It extends from northwestern edge of the Olympic Peninsula (Quillayute River) south to the Chehalis and Willapa rivers, north up the east side of Puget Sound to the Nooksack River and the lower Fraser River in BC (McPhail 1967). The Chehalis River Valley was a Pleistocene glacial refuge that was independent from the much larger lower Columbia/Pacific refuge (McPhail and Lindsey 1986). McPhail (1967) hypothesized that the Nooksack dace diverged from the longnose dace at some point in the Pleistocene as a result of its isolation in the Chehalis Refuge. Further, Bisson and Reimers (1977) described another morphologically divergent form of longnose dace, the Umpqua dace, from the Umpqua River on the Oregon coast. McPhail and Taylor (2009) used molecular sequence data to argue the Umpqua dace and the undescribed “Millicoma dace” (found in the Coos River, Oregon) are sister species that together form an Oregon coastal clade embedded within the *R. cataractae* species group. Further, Kim and Conway (2014) used a multigene approach to suggest that a primary division within the *R. cataractae* group should be made between those that are found east of the North American Continental Divide and those west of the Continental Divide (and including the longnose, Umpqua, and Millicoma dace). Interestingly, Bartnik (1972) demonstrated that longnose dace east and west of the Continental Divide, although morphologically similar to one another, had divergent

reproductive behaviour and ecology supporting their designation as distinct subspecies, *R. c. cataractae* and *R. c. dulcis*, respectively. Further, the speckled dace in Oregon contains several divergent lineages some of which may represent cryptic species (Pfrender et al. 2004; Ardren et al. 2011; Hoekzema and Sidlauskas 2014).

In this study, we assessed the distinctiveness of the Nooksack dace as another component of potentially cryptic species diversity within the western lineage of longnose dace. Specifically, we tested the hypothesis that the Nooksack dace, as suggested by its morphological distinctiveness and presumptive origin in a glacial refuge distinct from those used by the longnose dace, would form a monophyletic clade within the western group of *R. cataractae* both at mtDNA and nDNA loci. We further predicted that the Nooksack and longnose dace would be divergent from one another at a level consistent with the timing of one or more of the Pleistocene glaciations (i.e., diverged at least 100,000 to 150,000 years ago, see April et al. 2013). Resolution of the relationship between the longnose dace and the Nooksack dace is important for a more complete understanding of the nature and origins of *R. cataractae* in North America (Kim and Conway 2014), especially because the Nooksack dace is recognized as a wildlife species that is distinct from the longnose dace and is listed as Endangered under Canada's *Species at Risk Act* (COSEWIC 2007).

Materials and methods

Sample acquisition

Various species and populations of dace were sampled between 1993 and 2012 and comprised a mixture of field collections and sequences obtained from the GenBank database (Table 1, Fig. 1, <http://www.ncbi.nlm.nih.gov/genbank/>). The areas sampled

focused on British Columbia and Washington State, but included samples from across North America from *R. cataractae*, and outgroup taxa *R. evermanni*, *R. cataractae* sp. “*Millicoma dace*” (hereafter referred to as *Millicoma dace*), *R. atratulus*, *R. falcatus*, and *R. osculus*. Sequences obtained from GenBank were selected such that our results would be directly compared with the previous studies of McPhail and Taylor (2009) and Kim and Conway (2014).

Sequencing

Genomic DNA was extracted using QiaQuik spin columns, eluted in 150 μ L of AE buffer supplied by the manufacturer, and stored at -20°C until analysis. We used the polymerase chain reaction (PCR) to amplify an approximately fragments of five genes: the mitochondrial DNA cytochrome *b* (*cytb*), and NADH subunit 2 (ND2) genes, and the single copy nuclear DNA encoded S7 ribosomal protein gene intron 1 (S7) and the recombination activation gene 1 (RAG1). A total of 52 individuals were sequenced for *cytb*, after which we sequenced at least one individual of each unique *cytb* haplotype: 17 for ND2, 26 for S7, and 18 for RAG1. Amplifications were carried out in 50 μ L total volumes in the presence of (final concentrations): $1\times$ New England Biolabs ThermoPol Buffer (20 mM Tris-HCl; 10 mM $(\text{NH}_4)_2\text{SO}_4$; 10 mM KCl; 2 mM MgSO_4 ; 0.1% Triton X-100), 100 μ M of each dNTP, 0.2 μ M each primer, 1.5 units of *Taq* polymerase and between 200 and 1000 ng of template DNA. The *cytb* gene amplification comprised a 630 base pair long fragment using the primers: HD and GluDG (Palumbi 1996; Dowling et al. 2002). Amplification cycling conditions consisted of an initial denaturation at 95°C for 3 min, five cycles of 95°C denaturation, 55°C annealing, and 72°C extension for 45 s, 30 s, and

30 s, respectively, 30 cycles of 92 °C denaturation, 52 °C annealing, and 72 °C extension for 45 s, 30 s, and 30 s, respectively, and a final extension step at 72 °C for 10 min. The ND2 gene PCR comprised a ~1,050 bp fragment using the primers nd2-H and nd2-L of Tipton et al. (2011). The S7 (~850 bp) and RAG1 (~1,400 bp) introns were amplified using the primers S7RPEX1F and S7RPEX2R (Chow and Hazama 1998) and R1 2533F and R1 4078R (Lopez et al. 2004), respectively. Amplified fragments were checked for quality (e.g., single clear fragment) and quantity on 1.5% agarose gels stained using SYBR[®] Safe DNA gel stain and viewed under ultraviolet light. Amplified fragments were purified using Qiaquick PCR Purification columns and cycle-sequenced on an Applied Biosystems Inc. 37390S sequencer using dGTP BigDye[®] Terminator v3.0 chemistry using the HD (cytb), nd2-H (ND2), COF (COI), S7RPEX1F (S7), and R1 2533F (RAG) primers.

Data analyses

All sequences were aligned using Bioedit (Vers. 7.0.5.3, updated from Hall 1999) and have been deposited in GenBank under submission numbers KP694016-29 (S7), KP694030-47 (cytb), KP744099-114 (ND2), and KP744115-122 (RAG1). Following sequence alignment, jModeltest2 v 2.1.2 was used to estimate the best fit model among 88 models of sequence evolution for each set of sequences using the corrected Akaike Information Criterion (AIC_c, Darriba et al. 2012). Recombination in the S7 and RAG1 sequences was tested using the Phi test implemented in Splitstree (vers 4.12.8, Huson and Bryant 2006). Phylogenetic analyses using unique sequences for each DNA segment were subsequently conducted on haplotypes or alleles using maximum-likelihood employing the

appropriate substitution models accompanied by 1,000 bootstrap replicates using MEGA 5.1 (Tamura et al. 2012).

We estimated divergence time for longnose dace and Nooksack dace using the Bayesian framework implemented in BEAST (vers. 2.1.3.0, updated from Drummond et al. 2012) and concatenated sequences for cytb and ND2 (total of 1,453 base pairs). To allow for different evolutionary rates amongst lineages we used the log-normal relaxed molecular clock model under a calibrated Yule model of speciation. We used the HYK substitution model with a gamma category count of 4. We used a single, fossil-based calibration point for a specimen of *R. osculus* with an estimated age of about 4.5 million years (Smith and Dowling 2008; see also Kim and Conway 2014). Consequently, the origin of the *R. osculus* lineage was treated as a minimum age, using a lognormal distribution with a mean of 1.95, SD of 1.2, and an offset of 4.5. The BEAST analysis was replicated four times with each MCMC analysis running for 40,000,000 generations, sampled every 1,000 generations and disregarded the first 4,000,000 steps as a burn-in period which resulted in acceptable effective sample sizes (>200) and acceptance probabilities for all parameters of >0.25.

Results

Phylogenetic relationships

Our cytb sequencing covered 433 base pairs (bp) and resolved 17 unique haplotypes amongst the 52 *R. cataractae* individuals examined. Average uncorrected p-distances between *R. cataractae* and *R. falcatus*, *R. osculus*, *R. atratulus*, *R. evermanni*, and Millicoma dace were 0.081, 0.095, 0.090, 0.040, and 0.041, respectively. The longnose dace haplotypes sampled from west of the Continental Divide differed from those sampled east of the Continental Divide by an average 0.040 (0.025 net) p-distance.

Phylogenetic analysis for the cytb sequences suggested that *R. cataractae* was paraphyletic with respect to *R. evermanni* and the Millicoma dace (Fig. 2). The cytb tree resolved a clade consisting of *R. evermanni* and the Millicoma dace which was sister to a clade consisting of all *R. cataractae* from west of the Continental Divide. The samples from east of the Continental Divide formed a clade (99% support) that was sister to all “western” *R. cataractae* group plus *R. evermanni* and the Millicoma dace (Fig. 2). Within the clade consisting of all western *R. cataractae* group, there was one subclade (97% bootstrap support) containing all *R. cataractae* sampled from areas tributary to the upper and lower Fraser River and the upper Columbia River (all considered to be localities with fish resembling *R. cataractae* morphologically). Another subclade consisted of *R. evermanni* and the Millicoma dace (100% bootstrap support), and a third subclade (65% bootstrap support) comprised of fish sampled from tributaries of the Chehalis, Nooksack, and lower Fraser rivers, and from the Willapa River (all considered to be localities containing only the Nooksack dace morphologically - the Nooksack dace clade, Fig. 2). Three tributaries of the lower Fraser River (Kanaka Creek, and the Coquitlam and Alouette rivers) were the only ones to contain haplotypes both from of the western *R. cataractae* and Nooksack dace clades (Table 1). The western longnose dace and the Nooksack dace differed from each other by 0.026 (0.021 net) p-distance.

We sequenced each cytb haplotype at the ND2 gene (1,020 bp, $N = 1$ fish per haplotype) and found that p-distances between *R. cataractae* and *R. falcatus*, *R. osculus*, *R. atratulus*, *R. evermanni*, and the Millicoma dace were 0.142, 0.140, 0.168, 0.070, and 0.063, respectively. Among the longnose dace there were 13 distinct haplotypes (i.e., seven of the distinct cytb haplotypes were the same when sequenced at ND2, two fish with identical

cytb sequences had different ND2 sequences, Table 1). Longnose dace sampled from west of the Continental Divide differed from those sampled east of the Continental Divide by 0.072 (0.050 net) p-distance and western longnose dace and Nooksack dace differed from one another by 0.030 (0.023 net) p-distance. Phylogenetic analysis of the ND2 sequences also indicated that *R. cataractae* was paraphyletic with respect to *R. evermanni* and the Millicoma dace because the eastern longnose dace clade (100% bootstrap support) was sister to a lineage consisting of western longnose dace and *R. evermanni*/Millicoma dace (Fig. 2). The ND2 sequences resolved the same distinction between western longnose dace and Nooksack dace as with cytb, but with much higher levels of bootstrap support (96% - 100%, Fig. 2).

We resolved 629 bp of the S7 sequence and found no evidence for recombination ($P > 0.5$). Phylogenetic analysis of the S7 sequences also suggested that *R. cataractae* was paraphyletic with respect to *R. evermanni* and the Millicoma dace, and also resolved a major split between western and eastern groups of longnose dace (92% bootstrap support, Fig. 3). Two individuals of *R. evermanni* shared allele 1 with a total of 21 longnose and Nooksack dace. Further, there was no evidence for a phylogenetic distinction between western longnose dace and Nooksack dace (Fig. 3). In fact, two alleles were shared between western longnose dace and Nooksack dace and, as with *R. evermanni*, allele 1 was widely shared between fish characterized as longnose dace and Nootsack dace (Table 1). Alleles of S7 in *R. cataractae* (including *R. evermanni* and Millicoma dace) differed from those in *R. falcatus*, *R. osculus*, and *R. atratulus* by 0.021, 0.024, 0.034 p-distance, respectively. The average p-distance between *R. cataractae* sampled east and west of the Continental Divide was 0.013 (0.005 net p-distance).

We resolved 956 bp of RAG 1 and there was no evidence for recombination in these sequences ($P > 0.2$). Phylogenetic analysis of the sequences also indicated paraphyly of *R. cataractae* with respect to *R. evermanni* and the Millicoma dace (Fig. 3). A western lineage of *R. cataractae* (including *R. evermanni* and the Millicoma dace) was only weakly supported (50% bootstrap support), and one of the eastern alleles (from Québec, allele 5) was quite distinct from all other *R. cataractae* and different from them by a mean p-distance of 0.007 versus between 0.007 and 0.010 compared to the three other species of *Rhinichthys* (Fig. 3). Further, there was no phylogenetic distinction between longnose dace and Nooksack dace (Fig. 3). Alleles of RAG1 in *R. cataractae* (including *R. evermanni* and Millicoma dace) differed from those in *R. falcatus*, *R. osculus*, and *R. atratulus*, by 0.012, 0.010, and 0.012 p-distance, respectively. The mean difference between *R. cataractae* sampled east and west of the Continental Divide was 0.005 and two alleles were shared between western longnose dace and Nooksack dace (Fig. 3).

Divergence time estimation

The fossil-calibrated estimate of divergence between the longnose dace and Nooksack dace mtDNA clades using the combined cytb and ND2 data suggested that the two lineages diverged about 1.1 million years ago (mya), but confidence intervals (95% highest posterior density, HPD) ranged from 300,000 to 2,600,000 mya. The estimated divergence times between *R. evermanni* and the Millicoma dace mtDNA clades was 0.65 million years ago (95% HPD = 100,000 to 1.8 million years ago) and divergence time between the *R. cataractae* clade east of the Continental Divide and *R. cataractae*/*R. evermanni*/Millicoma

dace west of the Continental Divide was estimated as 2.5 million years ago (95% HPD = 500,000 to 6.75 million years ago).

Discussion

Our analysis resolved a relatively deep mitochondrial DNA (mtDNA) divergence (~2-3%) between the longnose dace and what has been referred to, from previous biogeographic and morphological analyses, as the Nooksack dace (cf. McPhail 1967; Ruskey and Taylor 2015). Ruskey and Taylor (2015) examined a larger sample of *R. cataractae*-like dace from across Canada and from Michigan, Oregon, and Washington State in the US using morphology and restriction fragment polymorphism assays of mtDNA ($N = 322$ fish from 12 localities using morphology and 269 fish from 12 localities using mtDNA) and found that the Nooksack dace has only been detected in portions of the lower Fraser River valley in southwestern BC, south along the eastern portion of Puget Sound and west to the Chehalis River area of western Washington State (see also McPhail 1967). The longnose dace – Nooksack dace divergence was nested within a broader scale divergence between longnose dace east and west of the North American Continental Divide (~4-7%), and is similar in magnitude to divergences between other western regional isolates within the *R. cataractae* group (e.g., *R. evermanni* and Millicoma dace). Consequently, our results support the idea that considerable cryptic diversity exists within the *R. cataractae* group and that such diversity occurs at different spatial scales (e.g., across North America and both east and west of the Continental Divide, cf. McPhail and Taylor 2009; Kim and Conway 2014). Notwithstanding cryptic diversity revealed by mtDNA, our data also indicated that while divergence at mtDNA was sometimes accompanied by

divergence at nuclear loci (e.g., eastern versus western *R. cataractae*), there was no consistent nDNA divergence observed between longnose dace and Nooksack dace, or between *R. cataractae*, *R. evermanni*, or the Millicoma dace within the western assemblage of the *R. cataractae* species group.

North American phylogeography

Kim and Conway (2014) conducted an extensive analysis within the *R. cataractae* group (but did not include the Nooksack dace) and our results are broadly consistent with theirs; paraphyly of *R. cataractae* with respect to *R. evermanni* and the Millicoma dace, and a consistent distinction between *R. cataractae* east and west of the Continental Divide. The analysis by Kim and Conway (2014) sampled many more areas east of the Continental Divide and they recovered a major distinction between fish from southeastern North America (Atlantic slope, southern Mississippi River drainage, and Rio Grande River) and those from northeastern North America (Arctic slope and northern tributaries of the Mississippi River). Our estimates of divergence time between western and eastern *R. cataractae* based on combined cytb and ND2 data were similar to that estimated by Kim and Conway (about 2.5 million years ago), but had wider confidence intervals. Kim and Conway (2014) suggested that the major west-east phylogeographic split within the *R. cataractae* group was associated with vicariant and re-connection events driven by repeated repositioning of the Continental Divide, and resultant watershed exchanges between the Miocene and Pleistocene. Further, the repeated glaciations during the Pleistocene and isolation of *R. cataractae* in separate refugia east and west of the Continental Divide may well have enhanced divergence initiated in earlier times (April et al. 2013; Taylor et al.

2013). Certainly, however, the west-east divergence greatly post-dates the major periods of Rocky Mountain orogeny (140 – 50 million years ago, English et al. 2004). Further, these lineages differ by 2.5% net p-distance at mtDNA cytochrome oxidase (Taylor unpubl. data) which is at or slightly above the level associated with significant reproductive isolation inferred from nuclear DNA analyses and with possible recognition as distinct taxa in several North American freshwater fishes (April et al. 2013). Interestingly, Bartnik (1972) reported significant differences in the spawning time, behaviour, and in male nuptial colouration between *R. cataractae* sampled from either side of the Continental Divide. While longnose dace on either side of the Continental Divide have designated by some as two subspecies (*R. c. cataractae* and *R. c. dulcis*, see Jordan and Evermann 1896; Bartnik 1972), this subspecific distinction is not widely recognized (McPhail and Lindsey 1970) and the magnitude of the molecular differences resolved here and by Kim and Conway (2014), coupled with the phenotypic measures of differentiation, support the idea that *R. cataractae* east and west of the Continental Divide may warrant separate taxonomic status at the species level (Kim and Conway 2014).

Western lineages of *R. cataractae*-like dace

The strong differentiation of the eastern samples of *R. cataractae* also complicates interpretations of the taxonomic relationships of western lineages within the longnose dace because our nDNA analyses indicated that *R. cataractae* was paraphyletic with respect to *R. evermanni* and the Millicoma dace (see also Kim and Conway 2014). This suggests that, at least based on molecular characterization, the three taxa (the Millicoma dace is still awaiting taxonomic description – see Bisson and Reimers 1977) may be better considered

as a single polytypic species, *R. cataractae*. If, however, only the western lineages of longnose dace like samples are examined, then western *R. cataractae* are monophyletic with respect both to *R. evermanni* and the Millicoma dace at least for mtDNA. The lack of resolution amongst western *R. cataractae*, *R. evermanni*, and the Millicoma dace at the two nuclear loci could be attributable to the generally longer coalescence time for these nuclear loci especially if the divergence of all three took place recently during the Pleistocene (Bisson and Reimers 1977; Zink and Barrowclough 2008; McPhail and Taylor 2009).

Similarly, we found no evidence of phyletic distinction between *R. cataractae* and the Nooksack dace in the two nuclear genes that we sequenced. Further, Girard and Angers (2011) documented modest variation among eastern populations of *R. cataractae* at single exons of growth hormone and trypsin loci, but we found no distinctions at these loci between small samples (five each) of longnose and Nooksack dace (E.B. Taylor, unpubl. data). As discussed above for *R. evermanni* and the Millicoma dace, the lack of nuclear sequence differentiation between longnose and Nooksack dace, despite their strong mtDNA divergence, does not necessarily mean that they might not be, in fact, distinct taxa because of a lack of resolution at these loci or owing to incomplete lineage sorting. Ruskey (2014) and Ruskey and Taylor (2015), however, studied longnose dace and Nooksack dace at three sites where they occur in sympatry in the lower Fraser River area of southwestern British Columbia. At each of these sites, fish identified as longnose dace and Nooksack dace using mtDNA were not differentiated when assayed at 10 microsatellite DNA loci and individual fish were not diagnosable using a key character, lateral line scale count, used in the original identification of the Nooksack dace (McPhail 1967). Consequently, the longnose dace and Nooksack dace were members of the same genetic populations in sympatry and the

presence of both mtDNA lineages in single localities is best interpreted as the result of the retention of an historical divergence owing to the haploid, clonal nature of mtDNA in fish that appear to have freely interbred upon secondary contact in the lower Fraser Valley (Ruskey and Taylor 2015). Consequently, the longnose and Nooksack dace appear not to warrant separate taxonomic recognition.

Our work on *R. cataractae* and that of Kim and Conway (2014) and the work of others on related *Rhinichthys* (e.g., Oakey et al. 2004; Pfrender et al. 2004; Smith and Dowling 2008; Hoekzema and Sidlauskas 2014) have indicated that the genus contains a number of cryptic lineages, especially within the mountainous landscape of western North America. The dynamic tectonic history of the area coupled with general features of the life-history of *Rhinichthys* (freshwater-resident, lithophilic, relatively short generation time, small-bodied and likely of limited dispersal abilities) have probably interacted to promote the evolution of divergent intraspecific lineages and perhaps some cryptic species. Despite some continuing taxonomic uncertainties (potential cryptic species in eastern and western lineages, ambiguity of status of *R. evermanni*) conservation of divergent lineages within *Rhinichthys* may be accomplished through their legal recognition as “distinct population segments” and “designatable units” under the *Endangered Species Act* and the *Species at Risk Act* in the US and Canada, respectively (USFWS and NMFS 1996; COSEWIC 2007).

Acknowledgements

This work was supported by Natural Sciences and Engineering Research Council of Canada Discovery and Equipment grants awarded to EBT. We thank M. Pearson, S. Avery-

Gomm, J. Witt, J. Mee, and L. Bernatchez for providing a number of samples and two two anonymous reviews for their input.

Draft

References

- April, J., Hanner, R.H., Dion-Cote, A-M., and Bernatchez, L. 2013. Glacial cycles as an allopatric speciation pump in northeastern American freshwater fishes. *Mol. Ecol.* **22**: 409-422.
- Ardren, W.R., Baumsteiger, J., and Allen, C.S. 2010. Genetic analysis and uncertain taxonomic status of threatened Fossket Spring speckled dace. *Conserv. Genet.* **11**: 1299-1315.
- Bartnik, V.G. 1972. Comparison of the breeding habits of two subspecies of longnose dace, *Rhinichthys cataractae*. *Can. J. Zool.* **50**: 83-86.
- Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K., and Das, I. 2007. Cryptic species as a window on diversity and conservation. *Trends Ecol. Evol.* **22**: 148-155.
- Bisson, P.A., and Reimers, P.E. 1977. Geographic variation among Pacific Northwest populations of longnose dace, *Rhinichthys cataractae*. *Copeia*, 1977: 518-522.
- Bond, J.E., Hedin, M.C., Ramirez, M.G., and Opell, B.D. 2001. Deep molecular divergence in the absence of morphological and ecological change in the Californian coastal dune endemic trapdoor spider *Aptostichus simus*. *Mol. Ecol.* **10**: 899-910.
- Chow, S., and Hazama, K. 1998. Universal PCR primers for S7 ribosomal protein gene introns in fish. *Mol. Ecol.* **7**: 1255-1256.
- COSEWIC. 2007. COSEWIC assessment and update status report on the Nooksack dace *Rhinichthys cataractae* ssp. in Canada. Ottawa: Committee on the Status of Endangered Wildlife in Canada (COSEWIC), vii + 27 pp.
- Darriba, D., Taboada G.L., Doallo, R., and Posada, D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**: 772.
- Dowling, T.E., Tibbets, C.A., Minckley, W.L., and Smith, G.R. 2002. Evolutionary relationships of the plagiocerins (Teleostei: Cyprinidae) from cytochrome b sequences. *Copeia*, 2002: 665-678.
- Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**: 1969-1973.
- English, J.M., and Johnston, S. T. 2004. The Laramide Orogeny: what were the driving forces? *International Geol. Rev.* **46**: 833-838.

- Feulner, P.G.D., Kirschbaum, F., Schugardt, C., Ketmaier, V., and Tiedemann, R. 2006. Electrophysiological and molecular genetic evidence for sympatrically occurring cryptic species in African weakly electricfishes (Teleostei: Mormyridae: *Campylomormyrus*). *Mol. Phylogenet. Evol.* **39**: 198-208.
- Fraser, B.A., Mandrak, N.E., and McLaughlin, R.L. 2005. Lack of morphological differentiation in eastern (*Rhinichthys atratulus*) and western (*Rhinichthys obtusus*) blacknose dace in Canada. *Can. J. Zool.* **83**: 1502-1509.
- Funk, W.C., Caminer, M., and Santiago, R.R. 2012. High levels of cryptic species diversity uncovered in Amazonian frogs. *R. Soc. Lond. B Biol. Sci.* **279**: 1806-1814.
- Girard, P., and Angers, B. 2011. The functional gene diversity in natural populations over postglacial areas: the shaping mechanisms behind genetic composition of longnose dace (*Rhinichthys cataractae*) in northeastern North America. *J. Mol. Evol.* **73**: 45-57.
- Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**: 95-98.
- Hoekzema, K., and Sidlauskas, B.L. 2014. Molecular phylogenetics and microsatellite analysis reveal cryptic species of speckled dace (Cyprinidae: *Rhinichthys osculus*) in Oregon's Great Basin. *Mol. Phylogenet. Evol.* **77**: 238-250.
- Hubbs, C. 1926. A check-list of the fishes of the Great Lakes and tributary waters, with nomenclatorial notes and analytical keys. University of Michigan Museum of Zoology Miscellaneous Publications No. 15. Ann Arbor, MI: University of Michigan.
- Hubbs, C.L., Miller, R.R., and Hubbs, L.C. 1974. Hydrographic history and relict fishes of the north-central Great Basin. *Mem. California Acad. Sci.* VII.
- Huson, D.H., and Bryant, D. 2006. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**: 254-267.
- Jordan, D.S., and Evermann, B.W. 1896. Fishes of North and Middle America. U.S. Nat. Mus. Bull. No. 47.
- Kim, D., and Conway, K.W. 2014. Phylogeography of *Rhinichthys cataractae* (Teleostei: Cyprinidae): pre-glacial colonization across the Continental Divide and Pleistocene diversification within the Rio Grande drainage. *Biol. J. Linn. Soc.* **111**: 317-333.

- Kraczkowski, M.L., and Chernoff, B. 2014. Molecular phylogenetics of the eastern and western blacknose dace, *Rhinichthys atratulus* and *R. obtusus* (Teleostei: Cyprinidae). *Copeia*, 2014: 325-338.
- Lee, D.S., Gilbert, C.R., Hocutt, C.H., Jenkins, R.E., McAllister, D.E., Stauffer Jr, J.R. 1980. Atlas of North American freshwater fishes. Raleigh: North Carolina State Museum of Natural History.
- Lopez, J.A., Chen, W-J., and Orti, G. 2004. Esociform phylogeny. *Copeia*, 2004: 449–464.
- McPhail, J.D. 1967. Distribution of freshwater fishes in western Washington. *Northwest Sci.* **40**: 1 - 11.
- McPhail, J.D., Lindsey, C.C. 1970. Freshwater fishes of northwestern Canada and Alaska. *Fish. Res. Board Can. Bull.* **173**.
- McPhail, J.D., and Lindsey, C.C. 1986. Zoogeography of the freshwater fishes of Cascadia (the Columbia system and rivers north to the Stikine). *In* The Zoogeography of North American freshwater fishes. *Edited by* C.H. Hocutt and E.O. Wiley. J. Wiley and Sons, New York, pp. 615-637.
- McPhail, J.D., and Taylor, E.B. 2009. Phylogeography of the longnose dace (*Rhinichthys cataractae*) species group in northwestern North America-the origin and evolution of the Umpqua and Millicoma dace. *Can. J. Zool.* **87**: 491-497.
- Miller, R.R. 1984. *Rhinichthys deaconi*, a new species of dace (Pisces: Cyprinidae) from southern Nevada. *Occ. Pap. Mus. Zool. Univ. Michigan.* **707**: 1-21.
- Oakey, D.D., Douglas, M.E., and Douglas, M.R. 2004. Small fish in a large landscape: diversification of *Rhinichthys osculus* (Cyprinidae) in western North America. *Copeia*, 2004: 207-221.
- Palumbi, S.R. 1996. Nucleic acids II: the polymerase chain reaction. *Mol. Syst.* **2**: 205-247.
- Pfrender, M.E., Hicks, J., and Lynch, M. 2004. Biogeographic patterns and current distribution of molecular-genetic variation among populations of speckled dace, *Rhinichthys osculus* (Girard). *Mol. Phylogenet. Evol.* **30**: 490-502.
- Ravaoarimanana, I.B., Tiedemann, R., Montagnon, D., and Rumpler, Y. 2004. Molecular and cytogenetic evidence for cryptic speciation within a rare endemic Malagasy lemur, the Northern Sportive Lemur (*Lepilemur septentrionalis*). *Mol. Phylog. Evol.* **31**: 440-448.

- Ruskey, J. A. 2014. Morphological stasis and genetic divergence without reproductive isolation in the *Rhinichthys cataractae* species complex: insights from a zone of secondary contact in the lower Fraser Valley, British Columbia. MSc thesis, Dept. of Zoology, University of BC, Vancouver, BC.
- Ruskey, J.A., and Taylor, E.B. 2015. A test of the cryptic species hypothesis in sympatric mitochondrial DNA lineages of the longnose dace (Cyprinidae: *Rhinichthys cataractae*). In review, Biol. J. Linn. Soc.
- Sáez, A.G., and Encarnación, L. 2005. Body doubles. *Nature*, **433**: 111.
- Smith, G.R., and Dowling, T.E. 2008. Correlating hydrographic events and divergence times of speckled dace (*Rhinichthys*: Teleostei: Cyprinidae) in the Colorado River drainage. In Late Cenozoic drainage history of the southwestern Great Basin and lower Colorado River region: Geologic and Biotic Perspectives. Edited by M.C. Reheis, R. Hershler, and D.M. Miller. Geol. Soc. Am. Spec. Pap. No. **439**: 301–317.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**: 2731-2739
- Taylor, E.B., Darveau, C-A., and Schulte, P.M. 2013. Setting conservation priorities in a widespread species: phylogeographic and physiological variation in the lake chub, *Couesius plumbeus* (Pisces: Cyprinidae). *Diversity*, **5**: 149-165.
- Tipton, M.L., Gignoux - Wolfsohn, S., Stonebraker, P., and Chernoff, B. 2011. Postglacial recolonization of eastern Blacknose dace, *Rhinichthys atratulus* (Teleostei: Cyprinidae), through the gateway of New England. *Ecol. Evol.* **1**: 343-358.
- Trontelj, P., and Fišer, C. 2008. Cryptic species diversity should not be trivialised. *Syst. Biodivers.* **7**: 1-3.
- USFWS and NMFS 1996. Policy regarding the recognition of distinct vertebrate population segments under the Endangered Species Act. Department of the Interior, US Fish and wildlife Service, and Department of Commerce and National Oceanic and Atmospheric Association, National Marine Fisheries Service. *Federal Register*, **61**: 4721-4725.
- Zink, R.M., and Barrowclough, G.F. 2008. Mitochondrial DNA under siege in avian phylogeography. *Mol. Ecol.* **17**: 2107-2121.

Draft

Table 1. Taxon, locality, drainage, and geographic coordinates for samples sequenced at cytochrome b (Cytb), NADH-2 (ND2) mitochondrial DNA genes and S7 ribosomal protein (S7) and recombination activation gene 1 (RAG1) nuclear loci in various taxa of North American *Rhinichthys*. All are single samples unless indicated otherwise in parentheses. Shown are the haplotype/allele numbers as depicted in Figures 2 - 3. AB = Alberta, BC = British Columbia, CO = Colorado, ID = Idaho, MB = Manitoba, MN = Minnesota, NE = Nebraska, OR = Oregon, QC = Québec, VI = Virginia, VT = Vermont, WA = Washington State. NS = Not sequenced, Ratr = *R. atratulus*, Reve = *R. evermanni*, Mill = *Millicoma dace*, Rfal = *R. falcatus*, Rosc = *R. osculus*, Mcau = *Mylocheilus caurinus*, all other haploypes/alleles are *R. cataractae*. Accession numbers for sequences obtained from GenBank are given below that haplotype/allele. Underlined localities with *R. cataractae* are east of the Continental Divide. Superscripts accompanying locality names (where applicable) refer to localities in Figure 1.

Taxon	Locality	Drainage	Haplotype/allele number				
			Coordinates	Cytb	ND2	S7	RAG1
<i>Rhinichthys cataractae</i> (Nooksack dace)							
	Alouette R. ¹	Lower Fraser River, BC	49.208889 -122.601389	2,3,6(2)	1,4	1(2),3,6(2)	2(2)
	Brunette R. ²	Lower Fraser R., BC	49.201111 -122.870833	3(5)	NS	4	NS
	Bertrand Cr. ³	Nooksack R., WA	49.030556 -122.530278	2,3,12(2)	8	1,2,7	NS
	Kanaka Cr. ⁴	Lower Fraser R., BC	49.206111 -122.533611	3(2),7	NS	NS	NS
	Coquitlam R. ⁵	Lower Fraser R., BC	49.234167 -122.753889	3,7(3)	NS	1	NS

Rhinichthys cataractae
(Longnose dace)

Chehalis R ⁶	Chehalis R., WA	46.968056 -123.595556	3	NS	NS	NS
Willipa R ⁷	Chehalis R., WA	46.731111 -123.731944	1,2,4(2),5	1(2),2(2),3	1(2),4,5	1,6
Porter Cr. ⁸	Chehalis R., WA	46.967778 -123.275556	14,15(2)	10,11	1,2	6
West Fork Satsop R ⁹	Chehalis R., WA	47.043333 -123.524444	16(2),17	12,13	1(2)	6
Norrish Cr. ¹⁰	Lower Fraser R., BC	49.176111 -122.148333	7	NS	1	NS
Coquihalla R. ¹¹	Lower Fraser R., BC	49.369722 -121.396111	7	NS	NS	NS
Little Sand Cr. ¹²	Upper Columbia R., BC	49.350556 -115.296111	7(3)	5	NS	3
Kootenay L. ¹³	Upper Columbia R., BC	49.625350 -116.934167	8(2)	6	NS	6
Beaver Cr. ¹⁴	Upper Columbia R., BC	49.108056 -117.554167	8	NS	5	4
Blueberry Cr. ¹⁵	Upper Columbia R., BC	49.250278 -117.955833	8,9	6	7,8	2

Columbia L. ¹⁶	Upper Columbia R., BC	52.102222 -115.818889	11	4	4	6
Petitot R. ¹⁷	Laird R., BC	59.648611 -122.924167	7(2)	NS	NS	NS
Bridge Cr. ¹⁸	Lower Columbia R., OR	44.733333 -120.307510	NS	NS	6	NS
Coast Fork R. ¹⁹	Lower Columbia R., OR	43.615278 -123.081111	7	NS	4	NS
<u>Upper Kananaskis Reservoir</u> ²⁰	South Saskatchewan R., AB	50.613601 -115.122853	13(2)	NS	11(2)	NS
<u>Churchill R.</u>	Hudson Bay, MB	58.675833 -94.184722	13(3)	9	12	NS
<u>St. Marguerite R.</u>	St. Lawrence R., QC	48.251944 -69.880000	10(2)	7	10(2)	5
<u>Rock Cr.</u>	Mississippi R., MN	45.716000 -92.886389	NS	NS	13 (KF640194)	7 (KF640168)
<u>Big Stoney Cr.</u>	Mississippi R., VI	37.420833 - 80.599722	NS	NS	NS	8 (KF640173)
<u>Blue Cr.</u>	Missouri R., NE	41.340556 -102.201667	NS	NS	NS	9 (JX443315)

<i>R. evermanni</i> (Umpqua dace)	Umpqua R.	Umpqua R., OR	43.310833 -123.217778	Reve (EU780890)	Reve	1	Reve
<i>Rhinichthys</i> sp. (‘Millicoma’ dace)	Coos R.	Coos R., OR	43.443889 -123.993056	Mill (EU871709)	Mill	Mill	Mill (KF640204)
<i>R. osculus</i> (Speckled dace)	Cottonwood Cr.	Malheur R., OR	42.632944 -120.506944	Rosc (JX442983)	NS	NS	Rosc (GU136359)
	Rock Cr. Springs	Colorado R, CO	39.699722 -107.070833	NS	NS	Rosc (GU134265)	NS
	Hay Cr.	Goose L., OR/CA *	42.184444 -120.707222	NS	Rosc (EU158229)	NS	NS
<i>R. falcatus</i> (Leopard dace)	Lower Fraser R.	Lower Fraser R., BC	49.203880 -121.785556	Rfal	Rfal	Rfal	Rfal
<i>R. atratulus</i> (Blacknose dace)	Lewis Cr.	Richelieu R., VT/QC	44.276111 - 73.183611	Ratr (KF640096)	NS	Ratr (KF640185)	Ratr (KF640178)
	Blackberry R.	Housatonic R., CT	42.000819 -73.219711	NS	Ratr (JN569207)	NS	NS

Mylocheilus caurinus
(Peamouth chub)

Billingsley Cr.	Snake R., ID	42.839166 -114.904444	NS	NS	NS	Mcau (JX443336)
-----------------	--------------	--------------------------	----	----	----	--------------------

* An endorheic basin

Draft

Figures

Figure 1. Map of various localities sampled for Nooksack dace (filled circles) and longnose dace (open circles) in western Canada and the United States. Numerals refer to numbered localities shown in Table 1. Only a subset of localities is shown to represent the general geographic distribution of Nooksack dace and longnose dace and owing to close spatial proximity of some localities, not each numeral is accompanied by a symbol. Other sample localities, for longnose dace in eastern North America and outgroup taxa, are detailed in Table 1. AK = Alaska, AB = Alberta, BC = British Columbia, MB = Manitoba, MT = Montana, OR = Oregon, WA = Washington State. The dashed heavy line represents the approximate northern, eastern, and southern range limit of the Nooksack dace.

Figure 2. Maximum-likelihood derived trees for (a) cytochrome *b* and (b) ND2 mitochondrial DNA haplotypes of various taxa of *Rhinichthys*. Numbers at branch points represent bootstrap percentages from 1,000 pseudoreplications with those with less than 50% support not shown. Also included for each tree is the best-fit substitution model for each sequence. The locations and taxon for each haplotype are defined in Table 1. Longnose dace haplotypes are symbolized by open circles, Nooksack dace haplotypes by closed circles, haplotypes from dace east of the Continental Divide are accompanied by asterisks (*). Rcat = *Rhinichthys cataractae*, Ratr = *R. atratulus*, Reve = *R. evermanni*, Rfal = *R. falcatus*, Rosc = *R. osculus*, Mill = Millicoma dace.

Figure 3. Maximum-likelihood derived trees for (a) S7 ribosomal protein intron 1 and (b) the recombination activation gene 1 for various taxa of *Rhinichthys*. Numbers at branch points represent bootstrap percentages from 1,000 pseudoreplications with those with less than 50% support not shown. Also included for each tree is the best-fit substitution model for each sequence. Longnose dace alleles are symbolized by open circles, Nooksack dace alleles by closed circles, alleles that are shared between longnose dace and Nooksack dace by closed diamonds, and that shared amongst longnose dace, Umpqua dace and Nooksack dace by an open diamond. Alleles from dace east of the Continental Divide are accompanied by asterisks (*). The locations and taxon for each allele are defined in Table 1. Rcat = *Rhinichthys cataractae*, Ratr = *R. atratulus*, Reve = *R. evermanni*, Rfal = *R. falcatus*, Rosc = *R. osculus*, Mill = Millicoma dace, Mcau = *Mylocheilus caurinus*.

Map of various localities sampled for Nooksack dace (filled circles) and longnose dace (open circles) in western Canada and the United States. Numerals refer to numbered localities shown in Table 1. Only a subset of localities is shown to represent the general geographic distribution of Nooksack dace and longnose dace and owing to close spatial proximity of some localities, not each numeral is accompanied by a symbol. Other sample localities, for longnose dace in eastern North America and outgroup taxa, are detailed in Table 1. AK = Alaska, AB = Alberta, BC = British Columbia, MB = Manitoba, MT = Montana, OR = Oregon, WA = Washington State. The dashed heavy line represents the approximate northern, eastern, and southern range limit of the Nooksack dace.

254x190mm (300 x 300 DPI)

. Maximum-likelihood derived trees for (a) cytochrome b and (b) ND2 mitochondrial DNA haplotypes of various taxa of *Rhinichthys*. Numbers at branch points represent bootstrap percentages from 1,000 pseudoreplications with those with less than 50% support not shown. Also included for each tree is the best-fit substitution model for each sequence. The locations and taxon for each haplotype are defined in Table 1. Longnose dace haplotypes are symbolized by open circles, Nooksack dace haplotypes by closed circles, haplotypes from dace east of the Continental Divide are accompanied by asterisks (*). Rcat = *Rhinichthys cataractae*, Ratr = *R. atratulus*, Reve = *R. evermanni*, Rfal = *R. falcatus*, Rosc = *R. osculus*, Mill = *Millicoma dace*.
254x190mm (300 x 300 DPI)

Maximum-likelihood derived trees for (a) S7 ribosomal protein intron 1 and (b) the recombination activation gene 1 for various taxa of *Rhinichthys*. Numbers at branch points represent bootstrap percentages from 1,000 pseudoreplications with those with less than 50% support not shown. Also included for each tree is the best-fit substitution model for each sequence. Longnose dace alleles are symbolized by open circles, Nooksack dace alleles by closed circles, alleles that are shared between longnose dace and Nooksack dace by closed diamonds, and that shared amongst longnose dace, Umpqua dace and Nooksack dace by an open diamond. Alleles from dace east of the Continental Divide are accompanied by asterisks (*). The locations and taxon for each allele are defined in Table 1. Rcat = *Rhinichthys cataractae*, Ratr = *R. atratulus*, Reve = *R. evermanni*, Rfal = *R. falcatus*, Rosc = *R. osculus*, Mill = *Millicoma dace*, Mcou = *Mylocheilus caurinus*.

254x190mm (300 x 300 DPI)