The Linnean Society of London BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2015, ••, ••–••. With 5 figures.

Morphological and genetic analysis of sympatric dace within the *Rhinichthys cataractae* species complex: a case of isolation lost

JENNIFER A. RUSKEY and ERIC B. TAYLOR*

Department of Zoology, Biodiversity Research Centre and Beaty Biodiversity Museum, University of British Columbia, #4200-6270 University Blvd, Vancouver, BC, V6T 1Z4, Canada

Received 21 June 2015; revised 5 August 2015; accepted for publication 5 August 2015

The Nooksack dace (Pisces: an undescribed putative taxon within *Rhinichthys*) and longnose dace (*Rhinichthys cataractae*) are two forms within the *R. cataractae* species complex that are distinguishable from one another by mitochondrial (mt) DNA divergence of 2–3%, as well as by subtle morphological differences. The two forms are found in allopatry in south-eastern British Columbia (BC), Canada, and adjacent areas of western Washington, USA, and are sympatric in three streams in the lower Fraser River valley, BC, and may represent cryptic species. We assayed 12 morphometric traits and two meristic characters (N = 582; 23 sampling locations) to test for diagnosability of the two dace, as well as to test for morphological differentiation by mtDNA type in sympatry. We then employed a 10-locus microsatellite DNA assay (N = 374; 12 sampling locations) to test for genetic distinction between Nooksack dace and longnose dace in sympatry. We found that the two dace could not be reliably differentiated morphologically: there was overlap in all characters measured, and sampling location had a much larger effect on morphology than mtDNA group. Microsatellite analysis showed no distinction by mtDNA type in localities with sympatric dace, indicating complete admixture between the sympatric Nooksack dace and longnose dace. The Nooksack dace and longnose dace provide an example of 'ephemeral speciation': two lineages that, despite an estimated 1.1 Myr of isolation, have developed no apparent barriers to reproduction and appear to have collapsed into a single interbreeding population where they come into secondary contact. Nonetheless, the zone of secondary contact and the diagnosability of the Nooksack dace in terms of mtDNA represent significant aspects of the evolutionary legacy within *R. cataractae* and support its conservation importance. © 2015 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2015, **00**, 000–000.

ADDITIONAL KEYWORDS: cryptic species – Cyprinidae – ephemeral speciation – introgression – reverse speciation – secondary contact.

INTRODUCTION

The path to becoming a new species is rarely straightforward; incipient species may remain in 'limbo' for millions of years (Avise & Walker, 1998), and may undergo parallel speciation, reverse speciation or re-speciation (Turner, 2002; Taylor *et al.*, 2006) potentially multiple times, before ultimately becoming two stable species or collapsing back into one (Rosenblum *et al.*, 2012). Episodes of recurrent vicariance and reconnection, such as those caused by the glacial cycles of the Pleistocene, can create complex patterns of divergence within species or between closely-related species (Bernatchez & Wilson, 1998). During the Pleistocene glaciations, temperate freshwater fishes experienced a changing landscape of intermittently connected and isolated waterways and glacial refugia, leading to complex phylogeography and many opportunities to study recent or incomplete speciation (Behnke, 1972; Bernatchez & Wilson, 1998; Gíslason *et al.*, 1999; Taylor, 1999; Ruzzante *et al.*, 2011).

Cyprinids are the largest family of vertebrates and the most diverse freshwater fish family in North America (Dowling *et al.*, 2002). The genus *Rhinichthys* (or 'riffle daces') consists of eight living species that are endemic to North America where they are broadly distributed in coolwater habitats

*Corresponding author. E-mail: etaylor@zoology.ubc.ca

(Matthews, Jenkins & Styron, 1982). Several studies have investigated systematics and biogeography within *Rhinichthys* (e.g. blacknose dace, *Rhinichthys atratulus*: Fraser, Mandrak & McLaughlin, 2005; Tipton *et al.*, 2011; speckled dace: *Rhinichthys osculus*, Pfrender, Hicks & Lynch, 2004; Kinziger *et al.*, 2011; Hoekzema & Sidlauskas, 2014). There has, however, been less study of the phylogeography and population structure of the longnose dace (LND) (*Rhinichthys cataractae*) which is the most widely distributed species in the genus and, indeed, perhaps the most widely distributed native North American minnow (McPhail & Taylor, 2009). An assemblage of closely related *Rhinichthys* is considered to constitute the LND 'species group', consisting of: LND, Umpqua dace (*Rhinichthys evermanni*), and the as yet undescribed Millicoma dace and Nooksack dace (NSD), with all but the LND restricted to west of the Continental Divide (McPhail, 1967; Bisson & Reimers, 1977; McPhail & Taylor, 2009; Taylor, McPhail & Ruskey, 2015).

The Nooksack dace (NSD) forms part of the 'Chehalis fauna', a group of animals proposed to have been isolated in the Chehalis River glacial refugium (south of Puget Sound and north of the Columbia River drainage) during the Pleistocene glaciations (McPhail, 1967). Although the NSD can be distinguished from the LND by a slightly more slender caudal peduncle and fewer scales on average along the lateral line and around the caudal peduncle, the two forms are otherwise morphologically similar to each other although they can be diagnosed using mitochondrial (mt) DNA (McPhail, 1967; Bisson & Reimers, 1977; Ruskey, 2014; Taylor *et al.*, 2015). Mitochondrial DNA sequence analysis has shown that LND west of the Continental Divide and NSD form reciprocally monophyletic clades that differ from each other by approximately 2% (cytochrome *b* and cytochrome *c* oxidase subunit I) to 3% (ND2) mitochondrial DNA sequence divergence, with a fossil-calibrated separation time between the two lineages estimated to be 1.1 Myr (Taylor *et al.*, 2015).

Our current understanding suggests that when the last glaciers receded, the NSD dispersed northwards from the Chehalis River refugium through postglacial lakes at the edge of the receding ice to recolonize several areas of western Washington as well as tributaries of the lower Fraser River in British Columbia (BC; McPhail, 1967; McPhail & Lindsey, 1986). The LND is considered to have survived in multiple Wisconsinan refugia and recolonized the lower Fraser River from the Pacific refugium using dispersal routes through the Columbia River valley, and also to other regions of BC from the Bering, Great Plains, and/or Pacific refugia (McPhail, 1967; McPhail & Lindsey, 1986; Taylor *et al.*, 2015). Consequently, although the LND is widespread across BC there are only four streams, all in south-western BC, that contain only fish bearing NSD mtDNA. Additionally, fish that bear NSD mtDNA and fish that bear LND mtDNA are known to be sympatric in only three streams that drain to the north side of the Fraser River, also in south-western BC (Taylor *et al.*, 2015). These streams with sympatric NSD and LND mtDNA lineages may therefore represent a zone of postglacial secondary contact and provide an opportunity to study the degree of reproductive isolation between the two mtDNA groups.

Although the two lineages are considered to represent isolation in different glacial refugia for at least 1 Myr (Taylor *et al.*, 2015), they remain morphologically very similar, any ecological differentiation between them is unknown, and their degree of mtDNA differentiation is similar to that between some fishes that are strongly reproductively isolated from one another (Broughton & Gold, 2000; Dowling *et al.*, 2002; April *et al.*, 2011, 2013) in which case the NSD and LND could represent cryptic biological species (Saez & Lozano, 2005; Bickford *et al.*, 2007; Hoekzema & Sidlauskas, 2014). By contrast, many lineages with deep levels of mtDNA divergence have not developed any significant reproductive barriers (Ward *et al.*, 2005; Zemlak *et al.*, 2009; Webb, Marzluff & Omland, 2011). Our understanding of the degree of concordance between mtDNA differentiation and biological species status is therefore incomplete. Consequently, testing the degree of genetic differentiation in the nuclear genome in sympatric dace presents an excellent opportunity to address the sometimes ambiguous nature of species identity among allopatric populations of differentiated forms (McPhail, 1984; Zamudio & Savage, 2003; Genner *et al.*, 2004; Hey *et al.*, 2004; Hausdorf, 2011). In the present study, we used a combination of morphological and microsatellite DNA analyses to test two hypotheses regarding the degree of distinctiveness between NSD and LND. First, we tested whether NSD and LND were morphologically and genetically distinct in allopatry, and thus consistently diagnosable from one another. Second, we tested whether sympatric NSD and LND co-occur without significant interbreeding between them such that they are genetically distinct from one another and therefore behave as distinct biological species. If, by contrast, the NSD and LND dace interbred completely upon secondary contact with extensive introgression such that there is no evidence of current reproductive isolation between types, the occurrence of two mtDNA groups in sympatry may better represent a vestige of past isolation.

MATERIAL AND METHODS

Sampling

Fish for morphological analysis consisted of historical collections drawn from the Beaty Biodiversity Museum's fish collection sampled between 1950 and 2008 (http://www.beatymuseum.ubc.ca/fish-collection) and fresh collections made during 2012. Altogether, 582 fish were sampled from 23 streams (see Supporting information, Table S1). Fish were collected using a combination of minnow trapping, seining, and electrofishing, initially stored in 10% formalin, and subsequently stored in 37.5% isopropyl alcohol. A subsample of the fish used in the morphological analyses were assayed for mtDNA (332 fish from 12 locations) and microsatellite DNA variation (374 fish from 12 locations) (Fig. 1, Table 1) (Taylor *et al.*, 2015). The samples used in the molecular analyses consisted of small fin clips stored in 95% ethanol before whole specimens were fixed in formalin.

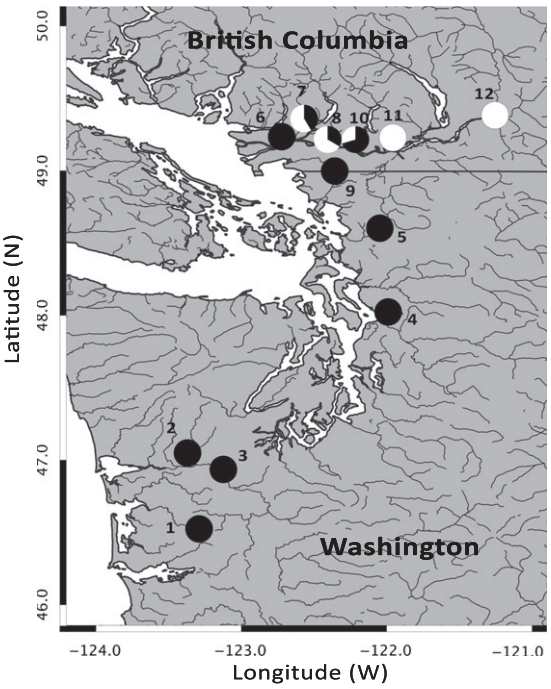

Figure 1. Map of Nooksack dace (black) and longnose dace (white) (*Rhinichthys cataractae*) cytochrome *b* mitochondrial DNA haplotypes. Each pie shows the proportional composition of Nooksack dace and longnose dace haplotypes (*sensu* Taylor *et al.*, 2015). 1, Willipa River (N = 5); 2, Satsop River (N = 25); 3, Porter Creek (N = 5); 4, Stilliguamish River (N = 5); 5, Nooksack River (N = 5); 6, Brunette River (N = 25); 7, Coquitlam; River (N = 41); 8, Alouette River (N = 86); 9, Bertrand Creek (N = 32); 10, Kanaka Creek (N = 44); 11, Norrish Creek (N = 40); 12, Coquihalla River (N = 19).

Morphometric and meristic measurements

Twelve morphometric characters were measured (mm) on the left side of each specimen and two meristic counts were made: lateral line scale count and pectoral fin ray count (Table 2) following Hubbs, Lagler & Smith (1958). Measurements were made using Vernier dial calipers accurate to 0.10 mm, and a dissecting microscope when necessary. We also counted vertebrae by eye from images generated using a microcomputed tomography scanner and visualization using MICROVIEW (GE Healthcare) in an initial sample of 53 fish (17 from allopatric NSD populations; 36 from allopatric LND populations). An analysis of variance (ANOVA) conducted on differences in the vertebral counts between NSD and LND, however, was nonsignificant (F = 0.11, P > 0.7) and there was almost complete overlap of vertebral counts in fish from NSD and LND mtDNA groups (Ruskey, 2014). The mean (SD) vertebral count was 34.7 (0.93) and 35.7 (0.66) for NSD and LND, respectively. Given the apparent lack of major differences in vertebral count in these initial samples, we did not make further counts.

Statistical analysis of morphological and meristic data

Morphometric measurements were first standardized to a common standard length to remove the effect of body size variation among samples. We performed allometric scaling to a standard size following Reist (1985) using the equation:

$$M_s = M_o(L_s/L_o)^b$$

where M_S is the standardized measurement, M_O is the measured character length, L_S is the overall mean standard length for all fish, L_O is the standard length of specimen, and b was estimated for each character using the allometric growth equation:

$$M = aL^b$$

We estimated parameter b as the slope of the regression of log M_O on log L_O. All samples were used to estimate b, although the intercept was allowed to vary between groups (streams allopatric for NSD, allopatric for LND, and sympatric for both NSD and LND mtDNA groups). No meristic counts

Table 1. Summary of genetic data for the 12 locations of allopatric Nooksack dace (N), allopatric longnose dace (L) or sympatric longnose and Nooksack dace (S) used in mitochondrial DNA and microsatellite analysis

Population number	Stream	Latitude	Longitude	N_1	N_2	F_{IS}	H_O	H_E	N_A	A_R
1-N	Porter Creek, WA	46.9465	123.2954	5	5	0.143	0.78	0.68	52	3.94
2-N	West Fork Satsop River, WA	47.0598	123.5410	25	25	–0.030	0.81	0.83	135	4.27
3-N	Bertrand Creek, BC	49.0043	122.5322	32	22	–0.004	0.77	0.78	90	3.78
4-N	Brunette River, BC	49.2414	122.8959	25	11	0.034	0.47	0.45	36	2.41
5-S	Coquitlam River, BC	49.0416	122.7711	41	66	0.013	0.83	0.82	177	4.21
6-S	Alouette River, BC	49.2392	122.5793	86	118	0.023	0.82	0.80	193	4.16
7-S	Kanaka Creek, BC	49.2022	122.5413	44	92	0.013	0.75	0.74	159	3.76
8-L	Norrish Creek, BC	49.2349	122.1336	40	11	0.050	0.72	0.68	67	3.55
9-L	Fraser River, BC	49.3831	122.4522	NA	6	0.041	0.83	0.80	71	4.45
10-L	Coquihalla River, BC	49.3885	121.4334	19	13	–0.016	0.86	0.88	99	4.39
11-L	Beaver Creek, BC	49.1009	117.5572	NA	16	–0.003	0.77	0.78	90	3.88
12-L	Beaver River, ON	44.3089	79.0405	NA	8	0.180	0.81	0.71	65	3.98

N_1, number of individuals sampled for mitochondrial DNA; N_2, number of individuals sampled for microsatellites; F_{IS}, inbreeding coefficient; H_O, observed heterozygosity; H_E, expected heterozygosity; N_A, total number of alleles across 10 loci; A_R, mean allelic richness across 10 loci. NA, not assayed; WA, Washington State; BC, British Columbia; ON, Ontario (N, L or S, as determined by Taylor *et al.*, 2015). All localities other than the Beaver River, ON, sample are from west of the Continental Divide.

Table 2. Description of morphometric (M1–M11) and meristic (Count 1–2) measurements examined in populations of Nooksack and longnose dace (*Rhinichthys cataractae*)

Variable	View	Description of measurement
M1	Dorsal	Distance between the eyes
M2	Dorsal	Width at dorsal fin origin
M3	Dorsal	Width at origin of caudal fin
M4	Left side	Distance from dorsal fin insertion to back side of eye
M5	Left side	Distance from dorsal fin insertion to pectoral fin insertion
M6	Left side	Distance from dorsal fin insertion to pelvic fin insertion
M7	Left side	Distance from dorsal fin insertion to anal fin insertion
M8	Left side	Distance from dorsal fin insertion to bottom of caudal fin insertion
M9	Left side	Distance from snout to front side of eye
M10	Left side	Distance from pectoral fin insertion to anal fin insertion
M11	Left side	Width of caudal peduncle at caudal fin insertion
Count 1	Left side	Lateral line scale count
Count 2	Left side	Pectoral fin ray count

were size-transformed; we tested for relationships between counts and standard length but found none (r between 0.02 and 0.1; all $P > 0.1$). These calculations were conducted using the R statistical environment (R Core, 2013).

We performed principal components analysis (PCA) on the correlation matrix of eleven size-transformed morphometric characters and two meristic traits using FactoMineR for R, version 1.24 (Husson *et al.*, 2013). The Jolliffe cut-off criterion was used to decide how many principal components (PCs) to retain for subsequent analyses (Jolliffe, 2002).

We performed nested ANOVA on PC scores with sampling location (random factor) nested within mtDNA group (fixed factor) to compare the effect of each factor. There were three categories for mtDNA group: NSD or LND (for fish from allopatric populations) and SYM (for fish from sympatric populations, where sympatric refers to the presence both of NSD and LND mtDNA groups in the locality sample). First, we conducted ANOVAs using allopatric populations to test for the morphological diagnosability of NSD and LND. We then analyzed allopatric and sympatric populations together to determine whether sympatric populations were morphologically intermediate compared to allopatric populations, as might be expected in a hybrid zone, or whether they comprised morphologically discrete groups, as might be expected if the two forms were reproductively isolated from one another in sympatry. Fish from all these localities had been characterized previously as either NSD and LND using mtDNA haplotype group (Taylor *et al.*, 2015), but in the present analysis haplotype identity was not recorded for individual fish

that were examined morphologically except in the case of one sympatric population, Kanaka Creek ($N = 44$). For the Kanaka Creek samples, individual fish were matched for mtDNA and morphological data to test for a difference in morphotype between mtDNA types in sympatry using single-factor ANOVAs on the PCs scores.

Lateral line scale count has been identified as perhaps the most distinguishing character between NSD and LND (McPhail, 1967). Consequently, we performed a nested ANOVA on this character separately with locality nested within either NSD or LND mtDNA group. Again, a single-factor ANOVA was analyzed separately for samples from Kanaka Creek where individual fish lateral line scale counts were matched with mtDNA group.

We also used model-based, unsupervised analyses to test for morphological separation of NSD and LND without a priori assignment to mtDNA group (Fraley & Raftery, 2007). Here, we employed the R package MCLUST, version 4 (Fraley *et al.*, 2012), which fits a series of Gaussian mixture models to the data and tests the fit of each model with different numbers of clusters, then selects the model and number of clusters that maximizes the Bayesian information criterion (BIC). Separate MCLUST analyses were conducted on (1) allopatric populations only, (2) allopatric and sympatric populations together, and (3) each sympatric stream separately. In all cases, we expected that if NSD and LND were morphologically diagnosable the MCLUST analysis would return two as the most likely number of morphological groups, each assignable to either NSD or LND. If more than two morphological groups were recovered, we expected that at least NSD and LND would be partitioned into largely separate subsets within these groups.

We also used the morphological data to conduct a model-based discriminant analysis in MCLUST. Discriminant analysis is used when groups are known, or assumed a priori, to exist and data characterizing each grouping are used to 'train' the discriminant analysis by finding the combination of the predictor variables that maximizes the difference between predefined groups: the discriminate function. We trained the discriminant analysis on half of the samples from allopatric NSD and LND localities and then tested its discriminatory ability on the other half. We also applied this discriminant function to sympatric samples from Kanaka Creek where each individual fish had matching mtDNA and morphological data.

Finally, we examined an alternative explanation for the previously reported morphological differentiation between NSD and LND (i.e. ecotypic variation in response to environmental differences) by testing for significant associations between morphological variation and environmental variables characterizing each watershed. Our choice of variables to test was determined by a combination of suggested relevance in previous studies (Barlow, 1961; Bisson, Sullivan & Nielsen, 1988; Langerhans *et al.*, 2003; McDowall, 2003), as well as the availability of environmental data for each watershed. We ran linear regressions of morphologically-derived PC scores for all dace against the environmental variables: watershed area, annual mean temperature, maximum temperature of the warmest month, minimum temperature of the coolest month, annual temperature range, annual precipitation, precipitation of wettest quarter, and precipitation of driest quarter. We also included the date of fish sample collection to assess possible effects of sample storage or degradation on morphological differentiation. We used raster GIS data from WorldClim (http://www.worldclim.corg; Hijmans *et al.*, 2005) and shapefiles from the United States Geological Survey (2014) and DataBC (British Columbia Provincial Government 2014). These data were joined to sampling location using QGIS (QGIS Development Team 2009), and linear regressions were performed in R.

Mitochondrial and microsatellite DNA analyses

Extractions of genomic DNA from fin samples were performed using Qiagen DNeasy Blood & Tissue Kit and Qiagen QIAamp DNA Investigator Kit spin columns and the DNA was stored at −20 °C. To assign new samples of dace to either the NSD or LND mtDNA group, we examined a 630-bp portion of the cytochrome *b* gene diagnostic for the two groups. The fragment was amplified using the polymerase chain reaction (PCR) using the material and procedures outlined in Taylor *et al.* (2015).

After PCR, the 630-bp product was subject to restriction fragment length polymorphism (RFLP) analysis using the restriction enzyme *Ava*II, which produces RFLP profiles diagnostic of the NSD and LND identified from fixed differences in the DNA sequences (Taylor *et al.*, 2015) (Fig. 1). Under these conditions, RFLP analysis of the 630-bp fragment produced two fragments of approximately 530 bp and 100 bp in NSD and a single (uncut) fragment of approximately 630 bp in LND. Some DNAs were of a quality insufficient to permit amplification of the larger 630-bp fragment. Consequently, a smaller 250-bp fragment was amplified using the primers: forward [5′-TGCCCCGTTAGCATGTATATT-3′] and reverse [5′-ACGAAAAACCCACCCACTAA-3′].

An annealing temperature of 54 °C was used, and all other conditions were the same as for the 630-bp fragment. The RFLP analysis of the approximately 250-bp product produced fragments of 157 bp and

93 bp in NSD, and a single (uncut) 250-bp fragment in LND. The resulting DNA fragments were separated on 2.5% agarose gels stained with CyberGreen (Molecular Probes, Inc.) and visualized under ultraviolet light.

Variation in the nuclear genome was analyzed by surveying allelic variation at 10 microsatellite DNA loci: *Rhca*15b, *Rhca*16, and *Rhca*23 (Girard & Angers 2006), and *Rhca*4, *Rhca*5, *Rhca*7, *Rhca*36, *Rhca*42, *Rhca*43, and *Rhca*45 (Beasley *et al.*, 2014) (see Supporting information, Table S2). PCRs were performed in 20-μL total volumes using the Qiagen Multiplex PCR Kit in accordance with the manufacturer's instructions with three multiplexes (MP): MP1 (*Rhca*5, *Rhca*36, *Rhca*42), MP2 (*Rhca*7, *Rhca*16, *Rhca*43), and MP3 (*Rhca*4, *Rhca*15b, *Rhca*45). *Rhca*23 was run in a separate PCR with annealing temperature of 55 °C and conditions otherwise as above for 250-bp cytochrome *b* fragment. For the multiplex PCRs, modifications were made in accordance with the manufacturer's instructions. In MP1, initial activation was followed by two 'touchdown' cycles in which the annealing temperature was 58 °C. This was followed by 35 cycles with an annealing temperature of 55 °C. Multiplex 2 was the same as MP1. Multiplex 3 had two 'touchdown' cycles at 60 °C followed by 35 cycles at 58 °C. All multiplexes used a final extension step of 30 min at 60 °C. The forward primer of each locus pair was fluorescently labeled to facilitate detection and allele identification using a Beckman-Coulter CEQ 8000 automated genotyper.

Statistical analysis of genetic data

Microsatellite data were first checked for the presence of null alleles and PCR or scoring artefacts using MICRO-CHECKER (Van Oosterhout *et al.*, 2004). We used FSTAT, version 2.9.3.2 (Goudet, 1995), to compile descriptive population genetic statistics: number of alleles, allelic richness, and observed and expected heterozygosity. Using GENEPOP, version 4.2.2 (Rousset, 2012), we tested for deviations from Hardy–Weinberg equilibrium for each combination of locus and population using an exact test in which probability values were estimated using a Markov chain method. We tested for linkage disequilibrium for all combinations of locus pairs within each population with a Markov chain method using GENEPOP default values.

We calculated pairwise F_{ST}, as estimated by Weir & Cockerham's (1984) θ, to estimate genetic divergence amongst each pair of locality samples using GENETIX, version 4.05 (Belkhir *et al.*, 2004). First, we compared mean pairwise F_{ST} among allopatric NSD with mean pairwise F_{ST} among allopatric LND to determine whether the overall level of divergence was different between NSD and LND and tested for a difference in pairwise F_{ST} using the permutation utility in FSTAT (Goudet, 1995). Next, we estimated mean pairwise F_{ST} for different groupings to determine whether localities with sympatric dace were more similar to localities with only NSD or only LND mtDNA to assess bias in introgression in favour of one form over the other. We tested for genetic distinction between sympatric dace bearing NSD and LND mtDNA by calculating pairwise F_{ST} between NSD and LND for each locality with sympatric dace using FSTAT. We also used analysis of molecular variance (AMOVA) to test for the degree of distinction in allele frequencies between localities with only LND mtDNA (five localities) or with only NSD mtDNA (four localities), and also within each locality with sympatric NSD and LND mtDNA groups using ARLEQUIN, version 3.5.1.3 (Excoffier, Laval & Schneider, 2005).

The model-based Bayesian clustering software STRUCTURE, version 2.3.4 (Pritchard, Stephens & Donnelly, 2000), was used to estimate the number of genetic populations (K) across all samples. STRUCTURE uses a Bayesian algorithm to identify clusters of individuals based on their genotypes at multiple loci by finding the arrangement that minimizes Hardy–Weinberg and linkage disequilibria within clusters (Pritchard *et al.*, 2000). We first analyzed the full dataset of 12 localities (Table 1) and then examined each locality with sympatric samples independently. Here, we expected that, if allopatric NSD and LND were genetically distinct lineages with respect to the nuclear genome, they would fall into two distinct groupings. For sympatric samples, we expected that if the NSD and LND were largely reproductively isolated from one another, each sympatric locality would contain two distinct genetic groups that closely matched the mtDNA group (i.e. NSD or LND). We used the admixture model and ran STRUCTURE five times for each model of K = 1–12, for 250 000 Markov chain Monte Carlo replications after a burn-in period of 50 000. We then used STRUCTURE HARVESTER, version 0.6.93 (Earl & vonHoldt, 2012), to evaluate different values of K in accordance with the ΔK method of Evanno, Regnaut & Goudet (2005). All tests involving multiple comparisons used P values adjusted following Narum (2006). All morphological and genetic data have been deposited in Dryad: doi:10.5061/dryad.g4237.

RESULTS

Variation in morphology and meristics in LND and NSD

For the analysis of all localities with only one mtDNA group (NSD or LND), the first five PCs were

retained, and accounted for 59.3% of the morphological/meristic variation (see Supporting information, Table S3). Based on the relative loadings of each morphological character, PC1 separated fish based on body robustness (thickness and depth) and lateral line scale count, PC2 separated fish based on body size posterior to the dorsal fin, PC3 separated fish with wide bodies from those that had elongate body parts anterior to the dorsal fin, and PC4 and PC5 separated fish largely with respect to pectoral fin ray counts (see Supporting information, Table S4). Two of the measurements (M6 and perhaps M2) had loadings that were sufficient to be considered 'high association' (loadings > 0.71 or < -0.71) (i.e. those characters making large contributions to the PC) (Tabachnick & Fidell, 2001).

There was little evidence of consistent morphological differences between NSD and LND after accounting for variation among localities. Nested ANOVAs showed that both mtDNA group (NSD or LND) and sampling location had significant effects for all PCs (all $P < 0.01$), except for the effect of mtDNA group on PC4 ($P = 0.103$). Typically, however, location accounted for a much greater percentage of the total morphological variation; for PC1, mtDNA group accounted for 11.6% of the total variation and location accounted for 39.3%. For PC2, mtDNA group accounted for 21.9% of the variation and location accounted for 25.1%. For PC3 to PC5, mtDNA group accounted for 0.5–1.8% of the variation and location accounted for 20.0–35.3%.

For the PC analysis on all allopatric and sympatric populations, the first five PCs were also retained and explained a total of 58.5% of the morphological/meristic variation (see Supporting information, Table S5). These axes also separated fish based on measures of body robustness (widths or depths), lateral line scale, and pectoral fin rays counts (see Supporting information, Table S6).

Nested ANOVAs conducted on PCs for allopatric and sympatric populations also showed that mtDNA group (LND, NSD or SYM) and sampling location had significant effects for all PCs (all $P < 0.028$), but again, location, again, typically accounted for a greater proportion of the total variation (e.g. for PC 1, mtDNA group accounted for 11.8% of the variation and location accounted for 35.9%). The only example of mtDNA group accounting for a greater proportion of the variation than location was for PC2; mtDNA group accounted for 29.9% and location accounted for 22.3%. For PC3 to PC5, mtDNA group accounted for 1.0–3.8% of the variation and location accounted for 17.5–28.6%.

Our analysis of sympatric dace in Kanaka Creek dace also resolved no consistent morphological differences between fish carrying NSD and LND mtDNA; ANOVAs were nonsignificant (all $P > 0.50$) for all comparisons except for along PC3 ($F = 4.07$, $P = 0.024$) where fish bearing LND mtDNA had wider caudal areas compared to those bearing NSD mtDNA.

By contrast to the PCA, there was a strong difference in lateral line scale counts between allopatric NSD and allopatric LND ($F = 186.5$, $P < 0.0001$) (Fig. 2). In a nested ANOVA, both mtDNA group and location had significant effects ($P < 0.0001$) and mtDNA group accounted for 28.9% of the variation, whereas location accounted for 15.5%. When the analysis was performed on samples from Kanaka Creek and grouped by NSD and LND mtDNA group, however, the distribution of lateral line scale counts was very similar between the NSD and LND (ANOVA: $F = 0.018$, $P > 0.98$) (Fig. 2). Overall, lateral line scale counts for sympatric populations were intermediate to those from allopatric NSD and LND: mean lateral line scale counts were 59.9, 62.5, and 66.3 for allopatric NSD, sympatric dace, and

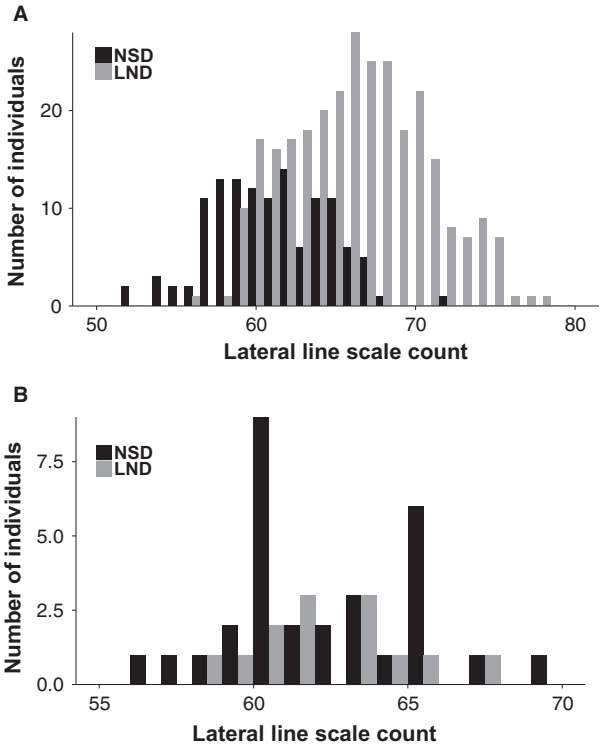

Figure 2. Histogram showing lateral line scale count by mitochondrial mtDNA haplotype in (A) samples from 19 allopatric Nooksack and longnose dace (*Rhinichthys cataractae*) sampling locations ($N = 414$) and (B) Kanaka Creek where both mtDNA haplotypes are found in the same stream ($N = 44$). Light grey bars signify longnose dace mtDNA, black bars signify Nooksack dace mtDNA. LND, longnose dace; NSD, Nooksack dace.

allopatric LND, respectively (ANOVA: $F = 117.6$, $P < 0.0001$; all Tukey's pairwise $P < 0.001$).

If NSD and LND are highly morphologically differentiated from one another we expected our MCLUST analysis to resolve two distinct groups of dace. Rather, the MCLUST analysis resolved no clear distinction between NSD and LND when only allopatric populations were examined, or when sympatric and allopatric samples were examined together. For example, the results for MCLUST analysis on allopatric NSD and LND populations indicated that the top model according to the BIC criterion was six morphological clusters (BIC = −6199.9 vs. −6224.9 for the next best model of three clusters). Under the six-cluster model, a strong association was detected between morphocluster (1–6) and mtDNA type ($G = 139.2$, $P < 0.0001$). Although each cluster was dominated by either NSD or LND mtDNA, all clusters contained a mixture of NSD or LND mtDNA (see Supporting information, Fig. S1). The only exception was cluster six, a particularly morphologically divergent group consisting of samples from the Wynoochee River (a NSD locality) in western Washington (see Supporting information, Fig. S1).

In the analysis combining allopatric and sympatric dace populations, the top model according to the BIC criterion was one consisting of five clusters (BIC = −8732.1 vs. −8742.6 for six clusters) (Fig. 3). Under the five-cluster model, a strong association was detected between morphocluster (1–5) and mtDNA type (NSD, LND or both in sympatric populations) across all populations ($G = 228.7$, $P < 0.0001$); however, each cluster again contained a mixture of samples from allopatric NSD sites, allopatric LND sites, and sympatric sites.

The results of MCLUST analysis conducted using only sympatric samples indicated that the top model was three clusters (BIC = −2742.2 vs. −2743.5 for four clusters). When each sympatric locality was analyzed separately, the best model was two clusters (BIC = −392.8) for the Coquitlam River, two clusters for the Alouette River (−773.4), and three clusters (BIC = −1029.7) for Kanaka Creek. For Kanaka Creek samples that had mtDNA type matched with morphology for individual fish, however, there was no significant association between mtDNA group (NSD or LND) and morphocluster membership (1–3; $G = 2.54$, $P = 0.28$) (see Supporting information, Fig. S1).

The discriminant analysis performed following MCLUST analysis on allopatric localities correctly assigned a mean of 88.4% of dace to the correct mtDNA group: 80.6% of NSD and 91.7% of LND were correctly assigned. When the discriminant analysis was applied to sympatric samples from Kanaka Creek, however, it correctly assigned a mean of only 47.7% of the dace to the correct mtDNA group: 38.7% of dace with NSD mtDNA and 69.0% of dace with LND mtDNA were correctly assigned.

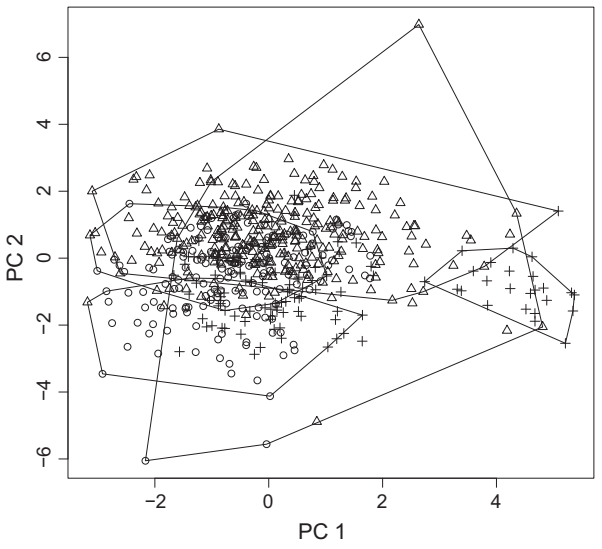

Figure 3. Convex hulls enclose each of five MCLUST-assigned cluster of individual Nooksack and longnose dace (*Rhinichthys cataractae*) based on variation in 11 morphometric and two meristic traits, measured in samples from all populations ($N = 582$). The mitochondrial DNA group is indicated by symbol: open triangles indicate allopatric longnose dace, plus signs indicate allopatric Nooksack dace, and open circles indicate sympatric samples. PC, principal component.

Geographical variation in morphology and environmental variables

We did not detect any strong associations between PC scores and environmental variables. Although some linear models were significant, r^2 values were all < 0.10, with the exception of the associations of PC2 with annual precipitation ($r^2 = 0.14$), highest precipitation quarter ($r^2 = 0.14$), collection date ($r^2 = 0.16$), and watershed area ($r^2 = 0.25$, all $P < 0.011$).

Mitochondrial DNA analysis

We combined our mtDNA results with previous mtDNA-based assays of *Rhinichthys* from across North America (Taylor *et al.*, 2015) (Fig. 1). The results showed that NSD and LND mtDNA both are present in the Coquitlam River, Alouette River, and Kanaka Creek in south-western BC, and that all other streams contain mtDNA from only one of the two mtDNA groups (Fig. 1). Taylor *et al.* (2015) showed that all samples of *R. cataractae* from Suicide Creek, BC (just 30 km east of Kanaka Creek)

to north-eastern BC and east to Québec (including samples from Alberta, Manitoba, Ontario, and Québec) contain only longnose dace mtDNA (N = 56 from 30 localities, analyzed for cytochrome *b* and ND2 regions).

Microsatellite DNA analysis

MICRO-CHECKER identified no artefacts as a result of stuttering or large allele dropout in our dataset, although null alleles were suggested to occur at *Rhca*7 in three localities (Coquitlam River, Alouette River, and Suicide Creek). Indeed, estimates of null allele frequencies provided by MICRO-CHECKER indicated that, in at least the Coquitlam River, null-null homozygotes were observed (3) at approximately the expected frequency (2.25) suggesting that the deficiency of heterozygotes at this locus in the Coquitlam River may be the result of a null allele. Nevertheless, removal of *Rhca*7 from our analyses did not influence the general results and so we retained this locus.

The highest allelic richness (averaged over loci) was 4.5 in the Fraser River sample and the lowest was 2.4 in the Brunette River. Mean allelic richness was 4.1 among allopatric LND populations, 3.6 among allopatric NSD populations, and 4.1 among sympatric populations (Table 1). The highest mean heterozygosity (averaged over populations) was found at locus *Rhca*43 (0.92), with the lowest at *Rhca*5 (0.55). Averaged across loci, the observed and expected heterozygosities were highest in the Coquihalla River (0.86, 0.88) and lowest in Brunette River (0.47, 0.45) (see Supporting information, Table S7).

Nine out of 120 tests for departure from Hardy–Weinberg equilibrium were significant at the adjusted critical P-value of 0.017 (15 at a nominal P-value of 0.05) and none were concentrated on a particular locus or population with the possible exception of the Alouette River sample which recorded three significant departures from Hardy–Weinberg equilibrium (see Supporting information, Table S7). Most of the departures were attributable to modest deficits of heterozygotes, although deficits at *Rhca*7 in two populations were more pronounced. Nine out of 450 tests of linkage disequilibrium were significant at the adjusted P-value of 0.011 (and 39 at P of 0.05), although no specific locus pairings were consistently out of equilibrium (Ruskey, 2014). Localities with both NSD and LND mtDNA accounted for a similar number of the deviations from linkage disequilibrium as allopatric localities: Alouette River (4 of 45 within-locality tests at P of 0.05), Coquitlam River (5), and Kanaka Creek (9 vs. between 0 and 6 for allopatric samples). The departures from Hardy–Weinberg equilibrium and especially linkage disequilibrium may be driven by the presence of some population structure within localities (see below) because samples were typically obtained from different reaches of each stream. Below, we address specifically whether or not such substructure is associated with the presence of LND and NSD within some streams.

Divergence among populations

Across all pairwise comparisons, mean F_{ST} was 0.130 (all $P < 0.05$) and ranged from 0.023 to 0.346 (see Supporting information, Table S8). The mean F_{ST} between localities with only NSD mtDNA was 0.159, and the mean F_{ST} between those with only LND mtDNA was 0.129. Localities containing both mtDNA groups of dace were equally differentiated from allopatric NSD and LND. For example, the mean pairwise F_{ST} between the three localities with sympatric dace mtDNA groups and allopatric NSD localities was 0.101 (range 0.097–0.106) and between sympatric sites and allopatric LND localities was 0.104 (range 0.096–0.128). Furthermore, allopatric NSD and LND were more similar to fish from localities sympatric for both mtDNA groups than they were to each other; the mean pairwise F_{ST} between allopatric NSD localities and allopatric LND localities was 0.169 (vs. 0.101 and 0.104 as described above). Mean pairwise F_{ST} between the three localities with sympatric dace was 0.041 (see Supporting information, Table S8).

An AMOVA on microsatellite allele frequencies indicated that the extent of variation accountable by interpopulation variation within NSD or LND mtDNA groups (14.7%, $P < 0.001$) was six times greater than that accounted for by NSD or LND mtDNA group alone (2.4%, $P = 0.012$). The greatest extent of the variation was accounted for by variation within individual localities (82.9%; $P < 0.001$).

Fish grouped by mtDNA type within localities with sympatric dace were not differentiated from each other with respect to microsatellite allele frequencies; values of pairwise F_{ST} between mtDNA groups were −0.001 for the Coquitlam River ($P = 0.509$), 0.003 for the Alouette River ($P = 0.152$), and 0.001 for Kanaka Creek ($P = 0.421$). Similarly, AMOVAs conducted on each locality with sympatric NSD and LND mtDNA indicated that variation attributable to NSD or LND mtDNA groupings was nonsignificant (all $P > 0.1$).

Bayesian analysis of population structure

There was no clear separation among localities based on whether they contained either NSD or LND mtDNA based on STRUCTURE analysis of all 12

localities (Figs 4, 5; see also Supporting Information, Table S9). For example, the model of population structure with the highest likelihood score was $K = 11$ (Fig. 4; see also Supporting Information, Table S9) and geography appeared to be important to clustering [e.g. a single genetic group predominated in the two Washington State localities (both NSD) and a second genetic group predominated in the localities tributary to the lower Fraser River (Norrish Creek, Coquihalla River, and the Fraser River mainstem; all LND)]. Furthermore, distinctive genetic groups were found both in the fish from the interior Columbia River population (Beaver Creek) and the fish from the Ontario locality (Beaver River), the most geographically distant localities in the present study (Fig. 4). Distinctive genetic groups also tended to predominate in each of the three localities with sympatric dace (Coquitlam and Alouette rivers, Kanaka Creek) and both of the localities from southwestern BC with fish bearing only NSD mtDNA (Fig. 4).

The ΔK criterion supported a reduced model of $K = 3$ (see Supporting information, Table S9), although there was no clear separation between localities with NSD or LND mtDNA (Fig. 5). For example, with $K = 3$, a single genetic group predominated in both of two of the three localities with sympatric dace, a second genetic group predominated in the other locality with sympatric dace, and the third genetic group predominated in two of the allopatric NSD and two of the allopatric LND (Fig. 5). Interestingly, some allopatric NSD localities (Porter Creek, Satsop River) had fish that were largely composed of the same genetic group as some allopatric LND localities (both Beaver creeks, Coquihalla River, grey shading in Fig. 5).

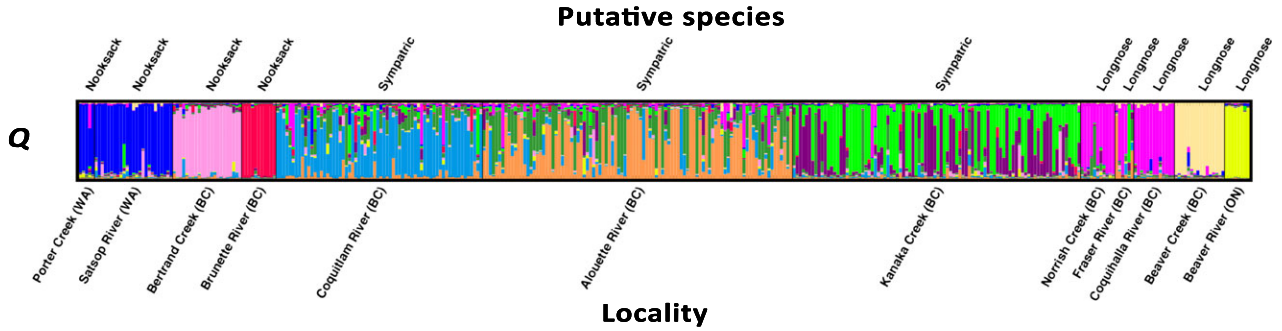

Figure 4. Output of STRUCTURE analysis showing admixture analysis from 10 microsatellite mtDNA loci in 374 Nooksack and longnose dace (*Rhinichthys cataractae*) ($K = 11$). Each fish is represented by a thin vertical line showing the proportional contribution [Q, which varies from 0.0 (bottom) to 1.0 (top)] to the genome of each fish by one of eleven genetic groups, each symbolized by a different colour. Population name is indicated below the graph and the mtDNA group is indicated above.

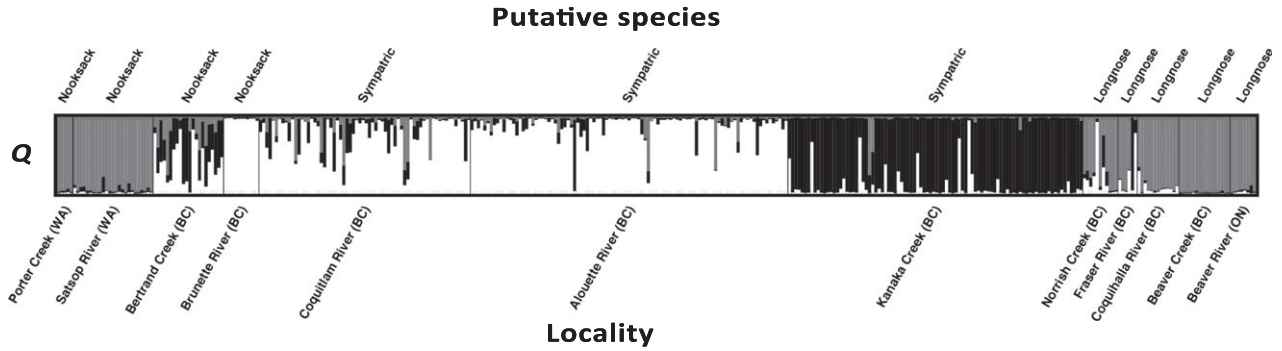

Figure 5. Output of STRUCTURE analysis showing admixture analysis from 10 microsatellite (mt)DNA loci in 374 Nooksack and longnose dace (*Rhinichthys cataractae*) ($K = 3$). Each fish is represented by a thin vertical line showing the proportional contribution [Q, which varies from 0.0 (bottom) to 1.0 (top)] to the genome of each fish by one of three genetic groups, each symbolized by black, grey, or white shading. Population name is indicated below the graph and the mtDNA group is indicated above.

The STRUCTURE analysis conducted within each locality containing fish bearing NSD and LND mtDNA revealed no evidence for more than one genetic group based on microsatellite DNA. Kanaka Creek and Coquitlam River both had $K = 1$ with the highest log likelihood, and bar plots for each of these two localities for any $K > 1$ showed equal contributions of each genetic cluster to each individual in the population (i.e. each fish showed the same proportional contribution of each hypothetical genetic group) (Ruskey, 2014) (see Supporting information, Table S10). Analysis of Alouette River also had $K = 1$ with the highest mean log likelihood (−4048.22), with $K = 2$ as the next most likely (−4289.52) (see Supporting information, Table S10). The ΔK criterion, however, suggested that $K = 3$ was the best model of population structure in the Alouette River, although here differences in admixture proportions amongst the three putative genetic groups were not distributed clearly by mtDNA type (see Supporting information, Fig. S2) (Ruskey, 2014).

DISCUSSION

Incomplete speciation in dace

The path to speciation is rarely straightforward; for every neat dichotomous branching, there are many other cases of incomplete speciation, or 'ephemeral speciation', in which new lineages form but do not persist (Rosenblum *et al.*, 2012). Dynesius & Jansson (2013) suggested that the rate of speciation should be considered as having three separate components: rate of lineage splitting, level of persistence of within-species lineages, and length of 'speciation duration': the time required to complete speciation. They suggested that the level of persistence as a factor of speciation rate has been understudied and that low persistence, as suggested by Rosenblum *et al.* (2012), is widespread and an important factor in determining overall rates of speciation. Additionally, incomplete speciation and the breakdown of reproductive isolation between lineages in contact may arise in situations where ecological conditions may have changed relatively recently after some period of reproductive isolation in sympatry (Taylor *et al.*, 2006; Nosil, Harmon & Seehausen, 2009; Behm, Ives & Boughman, 2010), or where the consequences of incomplete speciation played out after secondary contact (Webb *et al.*, 2011). Studies in zones of secondary contact are therefore an important component of investigating these questions and further understanding the rate of allopatric speciation (Haffer, 1969; Grant & Grant, 2009; April *et al.*, 2013).

Our data suggest that NSD and LND might represent a case of incomplete speciation. First, although allopatric samples of NSD and LND are divergent in mtDNA and somewhat divergent in morphology, they are not diagnosable in terms of microsatellite variation, nor are they clearly diagnosable in morphology. Second, there are cases where differentiated phenotypes that are not clearly diagnosable genetically in allopatry are nevertheless genetically distinct in sympatry and behave as distinct biological species. For example, there are several cases in postglacial fishes where genetic diagnosability of allopatric phenotypes at neutral loci is problematic, yet, when distinct phenotypes occur in sympatry, they are highly genetically distinct from one another (Bernatchez & Dodson, 1990; Schluter, 1996; Turgeon & Bernatchez, 2003). By contrast, our analyses did not reveal morphological or nuclear DNA distinctions in sympatric dace bearing NSD and LND mtDNA and, for all measures, locality accounted for more variation than mtDNA group. Taken together, our data suggest that where these dace have come into secondary contact they appear to have freely interbred such that there is no distinction in morphology or the nuclear DNA genome as assessed with microsatellites. In general, our study shows that despite an estimated 1.1 Myr of isolation (Taylor *et al.*, 2015), there has been minimal morphological differentiation between the NSD and LND and there are no reproductive barriers evident between NSD and LND after they came into secondary contact.

Subtle morphological differentiation between NSD and LND

Nooksack dace and LND showed subtle morphological differentiation but could not be reliably discriminated from one another and the effect of sampling location was stronger than the effect of mtDNA group. In particular, one morphological cluster was composed solely of dace from the Wynoochee River, WA, possibly as a result of drift, a small and distinctive founding population, and/or unique but unknown environmental features of that system (Clegg *et al.*, 2002; Langerhans *et al.*, 2003). This points towards the effect of mtDNA group being slight, whereas sampling location had a much larger effect.

There has been little morphological divergence observed between other dace within the *R. cataractae* group or across the geographical range of *R. cataractae* generally (Bisson & Reimers, 1977). By contrast, the speckled dace *R. osculus* exhibits considerable morphological variation across its range (Oakey, Douglas & Douglas, 2004) as well as extensive genetic diversity within Oregon's Great Basin (Hoekzema & Sidlauskas, 2014). Similar to *R. cataractae*, *R. osculus* experienced repeated range

fragmentation during the Pleistocene glaciations. This type of phylogeographical history, with species of limited dispersal ability being distributed across a fragmented landscape, has been suggested as one process by which genetic differentiation is generated without accompanying morphological differentiation. That *R. osculus* displays greater phenotypic variation across its range may be reflective of its more generalist nature: though the two fish have similar diets, *R. cataractae* is adapted for specialization in fast-flowing stream riffles whereas *R. osculus* is found in a wide range of habitats, from small streams to deep lakes. Specialization is considered to constrain phenotypic variation; for example, the African butterfly fish (*Pantodon buchholzi*) is highly specialized morphologically and provides one of the strongest examples in extant vertebrates of morphological stasis despite the evolution of deep intraspecific genetic divergence (Lavoué *et al.*, 2011). Although *R. cataractae* is not as extreme a specialist as the African butterfly fish, ecological specialization may still explain why its morphology appears less geographically variable than that of other *Rhinichthys* species.

Admixture of NSD and LND in sympatry

Microsatellite genetic data analyzed using F_{ST} and STRUCTURE indicated that there was no population structure within any of the streams with sympatric dace that could be associated with mtDNA type or morphology. The observation of divergent sympatric mtDNA groups, but lack of evidence for direct or indirect measures of reproductive isolation, could be the result of the presence of nuclear pseudogenes, natural selection, hybridization with a closely-related species, or cryptic biological species. That the NSD and LND mtDNA represent pseudogenes is very unlikely given that the divergences are apparent at cytochrome *b*, ND2, and cytochrome *c* oxidase subunit I and that, in all sequences, there is no sign of double peaks or stop codons (Song *et al.*, 2008; Taylor *et al.*, 2015). Although selection acting on mtDNA is certainly a possibility (Toews & Brelsford, 2012; Pavlova *et al.*, 2013; Toews *et al.*, 2014), it is highly unlikely in the case of NSD and LND given that divergent mtDNA types are often sampled within the same riffles within 0.5 m^2 of each other (E. B. Taylor, pers. observ.). Finally, the mtDNAs both of NSD and LND are clearly monophyletic with respect to other *Rhinichthys* and analysis of nuclear sequences at two loci indicates no evidence of hybridization with other members of the genus (Taylor *et al.*, 2015).

Our analyses refute the idea that LND and NSD represent cryptic biological species and point to another scenario: complete nuclear DNA admixture accompanied by retention of divergent mtDNA in sympatric populations. Such admixture was suggested by the lack of association of mtDNA type with the most divergent morphological character between LND and NSD (lateral line scale count) and no evidence for divergence in microsatellite DNA between mtDNA types. Furthermore, lateral line scale counts of sympatric samples were intermediate to those of allopatric NSD and LND, which is exactly what typically occurs if morphologically differentiated groups interbred upon secondary contact (April & Turgeon, 2006; Hermida *et al.*, 2009; Gompert *et al.*, 2010). The continued presence of NSD and LND mtDNA within the three sympatric localities suggests that there has been no bias in the interbreeding that might lead to the differential introgression of one mtDNA type over the other, which contrasts with some observations from other contact zones (Lamb & Avise, 1986; Redenbach & Taylor, 2002; Toews & Brelsford, 2012; Zieliński *et al.*, 2013).

Why NSD and LND appear to have interbred completely upon secondary contact is unknown. Webb *et al.* (2011) speculated that the interbreeding of once allopatric lineages (4% mtDNA divergence) of common ravens (*Corvus corvax*) may be related to their wide range of ecological tolerances, and/or the conservatism of signal traits in the genus *Corvus*. These explanations appear unlikely for the NSD and LND; males from LND populations east of the Rocky Mountains have been documented to have bright red nuptial coloration and breed during the day, whereas males from populations west of the Rocky Mountains typically do not have such red coloration and breed during the night (Bartnik, 1970, 1972), indicating that signal traits and mating behaviours may not be conserved across the range of *R. cataractae*. *Rhinichthys cataractae* consume a wide range of prey, but they are habitat specialists, adapted in both internal and external morphology for life at the bottom of swift-moving streams. Although ecological generalism may explain why long-separated species undergo neutral sequence divergence without functional divergence, ecological specialization may also be associated with morphological conservatism/stasis (Trontelj & Fišer, 2008). Such conservatism may have presented limited opportunities for ecological segregation to promote reproductive isolation in NSD and LND upon secondary contact (Schluter, 2001; Nosil, Crespi & Sandoval, 2003). Although apparently not the case for NSD and LND, there may still be undiscovered diversity in the *R. cataractae* group, particularly given morphological differences, deep mtDNA and nDNA divergences, and the aforementioned behavioural differences found between LND on either side of the Continental Divide, (Bartnik, 1972; Girard & Angers, 2011; Kim & Conway, 2014; Ruskey, 2014; Taylor *et al.*, 2015).

Conservation implications

Despite its uncertain taxonomic status, the Nooksack dace has been listed as Endangered under Canada's *Species at Risk Act* (SARA). The results of the present study are inconsistent with distinct biological species status for the Nooksack dace and illustrate that caution is required when relying upon mtDNA percentage sequence divergence thresholds for declaring new taxa (Hogner *et al.*, 2012). Despite 2–3% divergence in mtDNA, which puts it just within the estimated 'species cut-off' for cyprinid fishes (April *et al.*, 2011), the Nooksack dace exhibits no reproductive isolation from its sister mtDNA lineage, nor does it display any apparent functional divergence. Furthermore, Bailey, Winn & Smith (1954) suggested that a morphology-based definition of subspecies should be that at least 93% of the individuals from each allopatric population of one putative subspecies differ from allopatric populations of the other. Our discriminant analysis indicted that NSD and LND fell well below this proposed standard as well. Taxonomic considerations aside, however, the NSD still merits recognition and protection as a conservation unit separate from the LND. For example, under Canada's *Species at Risk Act* (SARA) a 'designatable unit' is a legally listable wildlife species and is defined as a population, or assemblage of populations, that is distinct (e.g. in terms of genetics or biogeography) from other such populations, and where this distinction is evolutionarily significant (i.e. in terms of being locally adaptive or of phylogeographical importance; COSEWIC 2014). Notwithstanding the lack of current reproductive isolation between the NSD and LND, the NSD is part of the distinctive Chehalis fauna and represents a separate mtDNA evolutionary lineage with slight, but discernible, morphological differences from its sister lineage, and thus qualifies as a significant component of the evolutionary legacy within *R. cataractae*.

ACKNOWLEDGEMENTS

We thank J. D. McPhail and M. Pearson for inspiration, valuable assistance, and discussion, as well as R. Seifert for field and laboratory assistance. This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC, discovery and equipment grants programs) and by the Canadian Wildlife Federation (Endangered Species Research program) in the form of grants awarded to EBT. We appreciate the comments on our research made by Kevin Omland and several anonymous reviewers.

REFERENCES

April J, Turgeon J. 2006. Phylogeography of the banded killifish (*Fundulus diaphanus*): glacial races and secondary contact. *Journal of Fish Biology* **69:** 212–228.

April J, Mayden RL, Hanner RH, Bernatchez L. 2011. Genetic calibration of species diversity among North America's freshwater fishes. *Proceedings of the National Academy of Sciences of the United States of America* **108:** 10602–10607.

April J, Hanner RH, Dion-Côté AM, Bernatchez L. 2013. Glacial cycles as an allopatric speciation pump in north-eastern American freshwater fishes. *Molecular Ecology* **22:** 409–422.

Avise JC, Walker D. 1998. Pleistocene phylogeographic effects on avian populations and the speciation process. *Proceedings of the Royal Society of London Series B, Biological Sciences* **265:** 457–463.

Bailey RM, Winn HE, Smith CL. 1954. Fishes from the Escambia River, Alabama and Florida, with ecologic and taxonomic notes. *Proceedings of the Academy of Natural Sciences of Philadelphia* **1954:** 109–164.

Barlow GW. 1961. Causes and significance of morphological variation in fishes. *Systematic Biology* **10:** 105–117.

Bartnik VG. 1970. Reproductive isolation between two sympatric dace, *Rhinichthys atratulus* and *R. cataractae*, in Manitoba. *Journal of the Fisheries Research Board of Canada* **27:** 2125–2141.

Bartnik VG. 1972. Comparison of the breeding habits of two subspecies of longnose dace, *Rhinichthys cataractae*. *Canadian Journal of Zoology* **50:** 83–86.

Beasley R, Lance SL, Ruskey JA, Taylor EB. 2014. Development and characterization of twenty-five microsatellite markers for the longnose dace (Cyprinidae: *Rhinichthys cataractae*) using paired-end Illumina shotgun sequencing. *Conservation Genetics Resources* **6:** 1011–1013.

Behm JE, Ives AR, Boughman JW. 2010. Breakdown in postmating isolation and the collapse of a species pair through hybridization. *American Naturalist* **175:** 11–26.

Behnke RJ. 1972. The systematics of salmonid fishes of recently glaciated lakes. *Journal of the Fisheries Board of Canada* **29:** 639–671.

Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F 2004. Genetix (Version 4.05) Logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).

Bernatchez L, Dodson JJ. 1990. Allopatric origin of sympatric populations of lake whitefish (*Coregonus clupeaformis*) as revealed by mitochondrial-DNA restriction analysis. *Evolution* **44:** 1263–1271.

Bernatchez L, Wilson C. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology* **7:** 431–452.

Bickford DL, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I. 2007. Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* **22:** 148–155.

Bisson PA, Reimers PE. 1977. Geographic variation among Pacific Northwest populations of longnose dace, *Rhinichthys cataractae*. *Copeia* **3:** 518–522.

Bisson PA, Sullivan K, Nielsen JL. 1988. Channel hydraulics, habitat use, and body form of juvenile coho salmon, steelhead, and cutthroat trout in streams. *Transactions of the American Fisheries Society* **117:** 262–273.

British Columbia Provincial Government. 2014. DataBC Geographic Services. Available at: http://www.data.gov.bc.ca/dbc/geographic/

Broughton R, Gold J. 2000. Phylogenetic relationships in the North American cyprinid genus *Cyprinella* (Actinopterygii: Cyprinidae) based on sequences of the mitochondrial ND2 and ND4L Genes. *Copeia* **2000:** 1–10.

Clegg SM, Degnan SM, Moritz C, Estoup A, Kikkawa J, Owens IP. 2002. Microevolution in island forms: the roles of drift and directional selection in morphological divergence of a passerine bird. *Evolution* **56:** 2090–2099.

COSEWIC. 2014. Guidelines for recognizing designatable units. Committee on the Status of Endangered Wildlife in Canada (COSEWIC) Operations and Procedures manual. Appendix F5. Available from Environment Canada, Canadian Wildlife Service, Gatineau, QC, Canada.

Dowling T, Tibbets C, Minckley W, Smith GR. 2002. Evolutionary relationships of the plagopterins (Teleostei: Cyprinidae) from cytochrome b sequences. *Copeia* **2002:** 665–678.

Dynesius M, Jansson R. 2013. Persistence of within-species lineages: a neglected control of speciation rates. *Evolution* **68:** 923–934.

Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4:** 359–361.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14:** 2611–2620.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1:** 47–50.

Fraley C, Raftery AE. 2007. Model-based methods of classification: using the mclust software in chemometrics. *Journal of Statistical Software* **18:** 1–13.

Fraley C, Raftery AE, Murphy TB, Scrucca L 2012. mclust Version 4 for R: Normal mixture modeling for model-based clustering, classification, and density estimation. Available at: https://www.stat.washington.edu/research/reports/2012/tr597.pdf

Fraser B, Mandrak N, McLaughlin R. 2005. Lack of morphological differentiation in eastern (*Rhinichthys atratulus*) and western (*Rhinichthys obtusus*) blacknose dace in Canada. *Canadian Journal of Zoology* **83:** 1502–1509.

Genner MJ, Seehausen O, Cleary DF, Knight ME, Michel E, Turner GF 2004. How does the taxonomic status of allopatric populations influence species richness within African cichlid fish assemblages? *Journal of Biogeography* **31:** 93–102.

Girard P, Angers B. 2006. Characterization of microsatellite loci in longnose dace (*Rhinichthys cataractae*) and interspecific amplification in five other Leuciscinae species. *Molecular Ecology Notes* **6:** 69–71.

Girard P, 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. *Journal of Molecular Evolution* **73:** 45–57.

Gíslason D, Ferguson MM, Skúlason S, Snorrason SS. 1999. Rapid and coupled phenotypic and genetic divergence in Icelandic Arctic char (*Salvelinus alpinus*). *Canadian Journal of Fisheries and Aquatic Sciences* **56:** 2229–2234.

Gompert Z, Lucas LK, Fordyce JA, Forister ML, Nice CC. 2010. Secondary contact between *Lycaeides idas* and *L. melissa* in the Rocky Mountains: extensive admixture and a patchy hybrid zone. *Molecular Ecology* **19:** 3171–3192.

Goudet J. 1995. FSTAT (Version 1.2). A computer program to calculate F-statistics. *Journal of Heredity* **86:** 485–486.

Grant PR, Grant BR. 2009. The secondary contact phase of allopatric speciation in Darwin's finches. *Proceedings of the National Academy of Sciences of the United States of America* **106:** 20141–20148.

Haffer J. 1969. Speciation in Amazonian forest birds. *Science* **165:** 131–137.

Hausdorf B. 2011. Progress toward a general species concept. *Evolution* **65:** 923–931.

Hermida M, San Miguel E, Bouza C, Castro J, Martínez P. 2009. Morphological variation in a secondary contact between divergent lineages of brown trout (*Salmo trutta*) from the Iberian Peninsula. *Genetics and Molecular Biology* **32:** 42–50.

Hey J, Won Y-J, Sivasundar A, Nielsen R, Markert JA. 2004. Using nuclear haplotypes with microsatellites to study gene flow between recently separated cichlid species. *Molecular Ecology* **13:** 909–919.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25:** 1965–1978.

Hoekzema K, Sidlauskas BL. 2014. Molecular phylogenetics and microsatellite analysis reveal cryptic species of speckled dace (Cyprinidae: *Rhinichthys osculus*) in Oregon's Great Basin. *Molecular Phylogenetics and Evolution* **77:** 238–250.

Hogner S, Laskemoen T, Lifjeld JT, Porkert J, Kleven O, Albayrak T, Kabasakal B, Johnsen A. 2012. Deep sympatric mitochondrial divergence without reproductive isolation in the common redstart *Phoenicurus phoenicurus*. *Ecology and Evolution* **2:** 2974–2988.

Hubbs CL, Lagler KF, Smith GR. 1958. *Fishes of the great lakes region*. Ann Arbor, MI: University of Michigan Press.

Husson F, Josse J, Le S, Mazet J. 2013. FactoMineR: multivariate exploratory data analysis and data mining with R. R package version **1:** 102–123.

Jolliffe IT. 2002. *Principal components analysis*. New York, NY: Springer-Verlag.

Kim D, Conway KW. 2014. Phylogeography of *Rhinichthys cataractae* (Teleostei: Cyprinidae): pre-glacial colonization across the Continental Divide and Pleistocene diversification within the Rio Grande drainage. *Biological Journal of the Linnean Society* **111:** 317–333.

Kinziger AP, Nakamoto RJ, Anderson EC, Harvey BC. 2011. Small founding number and low genetic diversity in an introduced species exhibiting limited invasion success (speckled dace, *Rhinichthys osculus*). *Ecology and Evolution* **1:** 73–84.

Lamb T, Avise JC. 1986. Directional introgression of mitochondrial DNA in a hybrid population of tree frogs: the influence of mating behavior. *Proceedings of the National Academy of Sciences of the United States of America* **83:** 2526–2530.

Langerhans RB, Layman CA, Langerhans AK, Dewitt TJ. 2003. Habitat-associated morphological divergence in two Neotropical fish species. *Biological Journal of the Linnean Society* **80:** 689–698.

Lavoué S, Miya M, Arnegard ME, McIntyre PB, Mamonekene V, Nishida M. 2011. Remarkable morphological stasis in an extant vertebrate despite tens of millions of years of divergence. *Proceedings of the Royal Society of London Series B, Biological Sciences* **278:** 1003–1008.

Matthews W, Jenkins R, Styron JTJ. 1982. Systematics of two forms of blacknose dace, *Rhinichthys atratulus* (Pisces: Cyprinidae) in a zone of syntopy, with a review of the species group. *Copeia* **1982:** 902–920.

McDowall RM. 2003. Variation in vertebral number in galaxiid fishes (Teleostei: Galaxiidae): A legacy of life history, latitude and length. *Environmental Biology of Fishes* **66:** 361–381.

McPhail JD. 1967. Distribution of freshwater fishes in western Washington. *Northwest Science* **41:** 1–11.

McPhail JD. 1984. Ecology and evolution of sympatric sticklebacks (*Gasterosteus*): morphological and genetic evidence for a species pair in Enos Lake, British Columbia. *Canadian Journal of Zoology* **62:** 1402–1408.

McPhail JD, Lindsey CC. 1986. Zoogeography of the freshwater fishes of Cascadia (the Columbia system and rivers north to the Stikine). In: Hocutt CH, Wiley EO, eds. *The zoogeography of North American freshwater fishes*. New York, NY: John Wiley & Sons, 615–637.

McPhail JD, Taylor EB. 2009. Phylogeography of the longnose dace (*Rhinichthys cataractae*) species group in northwestern North America — the origin and evolution of the Umpqua and Millicoma dace. *Canadian Journal of Zoology* **87:** 491–497.

Narum SR. 2006. Beyond Bonferroni: less conservative analyses for conservation genetics. *Conservation Genetics* **7:** 783–787.

Nosil P, Crespi BJ, Sandoval CP. 2003. Reproductive isolation driven by the combined effects of ecological adaptation and reinforcement. *Proceedings of the Royal Society of London B: Biological Sciences* **270:** 1911–1918.

Nosil P, Harmon LJ, Seehausen O. 2009. Ecological explanations for (incomplete) speciation. *Trends in Ecology and Evolution* **24:** 145–156.

Oakey DD, Douglas ME, Douglas MR. 2004. Small fish in a large landscape: diversification of *Rhinichthys osculus* (Cyprinidae) in western North America. *Copeia* **2004:** 207–221.

Pavlova A, Amos JN, Joseph L, Loynes K, Austin JJ, Keogh JS, Stone GN, Nicholls JA, Sunnucks P. 2013. Perched at the mitoc-nuclear crossroads: divergent mitochondrial lineages correlate with environment in the face of ongoing nuclear gene flow in an Australian bird. *Evolution* **67:** 3412–3428.

Pfrender ME, Hicks J, Lynch M. 2004. Biogeographic patterns and current distribution of molecular-genetic variation among populations of speckled dace, *Rhinichthys osculus* (Girard). *Molecular Phylogenetics and Evolution* **30:** 490–502.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945–959.

QGIS Development Team. 2009**.** Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project. Available at: http://qgis.osgeo.org.

R Core Team 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org

Redenbach Z, Taylor EB. 2002. Evidence for historical introgression along a contact zone between two species of char (Pisces: Salmonidae) in northwestern North America. *Evolution* **56:** 1021–1035.

Reist JD. 1985. An empirical evaluation of several univariate methods that adjust for size variation in morphometric data. *Canadian Journal of Zoology* **63:** 1429–1439.

Rosenblum EB, Sarver BAJ, Brown JW, Des Roches S, Hardwick KM, Hether TD, Eastman JM, Pennell MW, Harman LJ. 2012. Goldilocks meets Santa Rosalia: an ephemeral speciation model explains patterns of diversification across time scales. *Evolutionary Biology* **39:** 255–261.

Rousset F. 2012. Genepop (Version 4.2.2). Updated from Raymond M, Rousset F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86:** 248–249.

Ruskey JA 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. University of British Columbia, Vancouver, Canada.

Ruskey JA, Taylor EB. 2015. Data from: Morphological and genetic analysis of sympatric dace within the *Rhinichthys cataractae* species complex: a case of isolation lost. *Dryad Digital Repository*. doi:10.5061/dryad.g4237.

Ruzzante DE, Walde SJ, Macchi PJ, Alonso M, Barriga JP. 2011. Phylogeography and phenotypic diversification in the Patagonian fish *Percichthys trucha*: the roles of Quaternary glacial cycles and natural selection. *Biological Journal of the Linnean Society* **103:** 514–529.

Saez AG, Lozano E. 2005. Body doubles. *Nature* **433:** 111.

Schluter D. 1996. Ecological speciation in postglacial fishes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **351:** 807–814.

Schluter D. 2001. Ecology and the origin of species. *Trends in Ecology and Evolution* **16:** 372–380.

Song H, Buhay JE, Whiting MF, Crandall KA. 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 13486–13491.

Tabachnick BG, Fidell LS. 2001. *Using multivariate statistics*. Boston, MA: Allyn and Bacon.

Taylor EB. 1999. Species pairs of north temperate freshwater fishes: evolution, taxonomy, and conservation. *Reviews in Fish Biology and Fisheries* **9:** 299–324.

Taylor EB, Boughman JW, Groenenboom M, Sniatynski M, Schluter D, Gow JL. 2006. Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology* **15:** 343–355.

Taylor EB, McPhail J, Ruskey J 2015. Phylogeography of the longnose dace (*Rhynichthys cataractae*) species group in northwestern North America – the Nooksack dace problem. *Canadian Journal of Zoology* **93**. DOI**:** 10.1139/cjz-2015-0014.

Tipton ML, Gignoux-Wolfsohn S, Stonebraker P, Chernoff B. 2011. Postglacial recolonization of eastern Blacknose Dace, *Rhinichthys atratulus* (Teleostei: Cyprinidae), through the gateway of New England. *Ecology and Evolution* **1:** 343–358.

Toews DPL, Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* **21:** 3907–3930.

Toews DP, Mandic M, Richards JG, Irwin DE. 2014. Migration, mitochondrial, and the yellow-rumpted warbler. *Evolution* **68:** 241–255.

Trontelj P, Fišer C. 2008. Cryptic species diversity should not be trivialised. *Systematics and Biodiversity* **7:** 1–3.

Turgeon J, Bernatchez L. 2003. Reticulate evolution and phenotypic diversity in North American ciscoes, Coregonus ssp. (Teleostei: Salmonidae): implications for the conservation of an evolutionary legacy. *Conservation Genetics* **4:** 67–81.

Turner GF. 2002. Parallel speciation, despeciation and respeciation: implications for species definition. *Fish and Fisheries* **3:** 225–229.

United States Geological Survey. 2014. National hydrography dataset. Available at: http://nhd.usgs.gov/

Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4:** 535–538.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. 2005. DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society Series B, Biological Sciences* **360:** 1847–1857.

Webb WC, Marzluff JM, Omland KE. 2011. Random interbreeding between cryptic lineages of the Common Raven: evidence for speciation in reverse. *Molecular Ecology* **20:** 1–13.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* **38:** 1358–1370.

Zamudio K, Savage W. 2003. Historical isolation, range expansion, and secondary contact of two highly divergent mitochondrial lineages in spotted salamanders (*Ambystoma maculatum*). *Evolution* **57:** 1631–1652.

Zemlak TS, Ward RD, Connell AD, Holmes BH, Hebert PDN. 2009. DNA barcoding reveals overlooked marine fishes. *Molecular Ecology Resources* **9**(Suppl. 1)**:** 237–242.

Zieliński P, Nadachowska-Brzyska K, Wielstra B, Szkotak R, Covaciu-Marcov SD, Cogălniceanu D, Babik W. 2013. No evidence for nuclear introgression despite complete mtDNA replacement in the Carpathian newt (*Lissotriton montandoni*). *Molecular Ecology* **22:** 1884–1903.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Figure S1. Stacked bar chart showing the mitochondrial DNA composition of each morphological cluster generated by the MCLUST analysis conducted on: (a) allopatric Nooksack and longnose dace (*Rhinichthys cataractae*) (N = 414) and (b) Nooksack and longnose dace (*Rhinichthys cataractae*) from Kanaka Creek (N = 44).

Figure S2. Output of STRUCTURE run showing admixture analysis with K = 3 from 10 microsatellite DNA loci in 99 Nooksack and longnose dace (*Rhinichthys cataractae*) from the Alouette River, a site sympatric for longnose and Nooksack mitochondrial DNA haplotypes.

Table S1. List of sampling locations for Nooksack dace and longnose dace (*Rhinichthys cataractae*) used in the morphological analyses.

Table S2. Details for 10 polymorphic microsatellite loci developed for *Rhinichthys cataractae*.

Table S3. Eigenvalues and percentage of variance explained by each axis of a principal components analysis conducted on morphological samples for allopatric populations of Nooksack dace and longnose dace (*Rhinichthys cataractae*).

Table S4. Character loadings for the first five principal components of the principal components analysis conducted on 11 size-transformed morphological traits and two meristic traits values of morphological samples for allopatric populations of Nooksack dace and longnose dace (*Rhinichthys cataractae*).
Table S5. Eigenvalues and percentage of variance explained by each axis of a principal components analysis conducted morphological samples for allopatric and sympatric populations of Nooksack dace and longnose dace (*Rhinichthys cataractae*).
Table S6. Character loadings for the first five axes of the principal components analysis of 11 size-transformed morphological traits and two meristic traits. Principal components analysis was conducted on 582 Nooksack and longnose dace (*Rhinichthys cataractae*).
Table S7. Observed (H_E) and expected (H_O) heterozygosities for each locus/population combination, for 10 microsatellite loci and 12 populations of Nooksack and longnose dace (*Rhinichthys cataractae*).
Table S8. Pairwise F_{ST} for all sampling locations, with each sympatric population broken down into subsets for those dace with Nooksack dace (NSD) and longnose dace (LND) mitochondrial DNA.
Table S9. Results from STRUCTURE HARVESTER (Earl & vonHoldt, 2012) for STRUCTURE analysis conducted on all 12 Nooksack and longnose dace (*Rhinichthys cataractae*) populations, testing assumed number of populations $K = 1–12$, with five repetitions, and a total $N = 374$. K = number of populations assumed; Reps = number of times the simulation was run for a given K; mean $LnP(K)$ = mean log likelihood of K over all reps for that K; Stdev $LnP(K)$ = SD for $LnP(K)$ over all reps for that K; $Lni(K)$ = first-order rate of change of mean $LnP(K)$, defined as $LnP(K) - LnP(K–1)$; $|Ln''(K)|$ = second-order rate of change, defined as $Ln^i(K + 1) - Ln'(K)$; $\Delta K = |Ln''(K)|$ divided by Stdev $LnP(K)$. The highest log likelihood is indicated in bold ($K = 11$) and the highest ΔK is highlighted in grey ($K = 3$).
Table S10. Results from STRUCTURE HARVESTER (Earl & vonHoldt, 2012) for Kanaka Creek (A) ($N = 118$), Coquitlam River (B) ($N = 66$), and Alouette River (C) ($N = 99$) samples of Nooksack and longnose dace (*Rhinichthys cataractae*). Each analysis was conducted for assumed number of populations $K = 1–6$, with five repetitions. K, number of genetic populations assumed; Reps, number of times the simulation was run for a given K; mean $LnP(K)$, mean log likelihood of K over all repetitions for that K; Stdev $LnP(K)$, SD for $LnP(K)$ over all repetitions for that K; $Ln'(K)$, first-order rate of change of mean $LnP(K)$, defined as $LnP(K) - LnP(K - 1)$; $|Ln''(K)|$, second-order rate of change, defined as $Ln'(K + 1) - Ln'(K)$; $\Delta K = |Ln''(K)|$ divided by Stdev $LnP(K)$. For each stream, the highest log likelihood is indicated in bold ($K = 1$) and the highest ΔK is highlighted in grey ($K = 5$).

SHARED DATA

Data deposited in the Dryad digital repository (Ruskey & Taylor, 2015).