Journal of FISH BIOLOGY

fsbi An International Society for Fish Biology

Journal of Fish Biology (2016) **88**, 1460–1474
doi:10.1111/jfb.12922, available online at wileyonlinelibrary.com

Real-time PCR identification of lake whitefish *Coregonus clupeaformis* in the Laurentian Great Lakes

L. M. Overdyk*†, H. E. Braid‡, A. M. Naaum*, S. S. Crawford*§ and R. H. Hanner*‖

**Department of Integrative Biology, University of Guelph, 50 Stone Road East, Guelph, ON, N1G 2W1 Canada, ‡Institute for Applied Ecology New Zealand, Auckland University of Technology, Private Bag 92006, Auckland, New Zealand, §Chippewas of Nawash, Unceeded First Nation, 135 Lakeshore Blvd, Nevaashiinigmiing, ON, N1G 2W1 Canada and ‖Biodiversity Institute of Ontario, University of Guelph, 50 Stone Road East, Guelph, ON, N1G 2W1 Canada*

(Received 30 April 2015, Accepted 23 January 2016)

The purpose of this study was to develop a real-time PCR assay to specifically identify lake whitefish *Coregonus clupeaformis* in larval fish assemblages based on a 122 bp amplicon from the mitochondrial genome. The efficiency of the reaction, as calculated from the standard curve, was 90·77% with the standard curve having an r^2 value of 0·998. Specificity of the assay provided single melt peak in a melt-curve analysis and amplification of only the target species. The assay successfully identified target DNA in as low as 0·1% proportion of a DNA mixture. This assay was designed on the portable Smart Cycler II platform and can be used in both field and laboratory settings to successfully identify *C. clupeaformis*.

Key words: Coregoninae; DNA barcoding; entrainment; species identification; TaqMan.

INTRODUCTION

Species-level identification of ichthyoplankton can be extremely challenging due to the small size of larvae and lack of distinguishing morphometric and meristic characteristics (Scott & Crossman, 1973; Teletchea, 2009). The proper identification of ichthyoplankton is extremely important in fish ecology, especially for understanding important spawning and nursery habitats, life-history dynamics and year-class forecasting in fishery management (Ko *et al.*, 2013). The classical method for identification of larval fish relies on morphology-based keys (Scott & Crossman, 1973; Ko *et al.*, 2013), which can be highly unreliable and inaccurate (Hare *et al.*, 1994; Kochzius *et al.*, 2008).

The morphological identification of larval fishes in the subfamily Coregoninae is particularly challenging because phenotypic characteristics often vary across environments, and there are few distinct morphological characteristics for species with

†Author to whom correspondence should be addressed. Tel. +1 519 824 4120 ext. 53594; email loverdyk@uoguelph.ca

overlapping ranges (Scott & Crossman, 1973; Auer, 1982; Schlei *et al.*, 2008). In the Laurentian Great Lakes, coregonines such as lake whitefish *Coregonus clupeaformis* (Mitchell 1818), lake herring *Coregonus artedi* Lesueur 1818, bloater *Coregonus hoyi* (Milner 1874), shortjaw cisco *Coregonus zenithicus* (Jordan & Evermann 1909) and round whitefish *Prosopium cylindraceum* (Pennant 1784) are important species socially, commercially and ecologically to both indigenous and non-indigenous human communities (S. Crawford, A. Muir & K. McCann, unpubl. data; Golder Associates, unpubl. data). Within the Great Lakes, other coregonids such as the deepwater ciscos *Coregonus kiyi* (Koelz 1921) and *Coregonus johannae* (Wagner 1910), the shortnose cisco *Coregonus reighardi* (Koelz1924), the blackfin cisco *Coregonus nigripinnis* (Milner 1874) and the longjaw cisco *Coregonus alpenae* (Koelz 1924) have been reported, with *C. alpenae*, *C. nigripinnis*, *C. reighardi* and *C. johannae* currently considered extinct (Todd & Smith, 1992; Davis & Todd, 1998; Roth *et al.*, 2012). *Coregonus kiyi* and *C. nigripinnis* are reported as extirpated, and lavaret *Coregonus lavaretus* (L. 1758) and maraene *Coregonus maraena* (Bloch 1779) were introduced into Lake Huron but failed to establish (Roth *et al.*, 2012). Among the commercially harvested species, larvae of *C. artedi* and *C. clupeaformis* have previously been reported as very difficult to differentiate on the basis of morphological characteristics (Todd & Stedman, 1989). Both *C. clupeaformis* and *P. cylindraceum* have been specifically classified as 'Valued Ecosystem Components' (VEC) in Lake Huron because of their ecological significance, economic value and presence of important spawning locations in Lake Huron (J. A. Holmes & D. I. G. Noakes, unpubl. data; T. Brown, unpubl. data). For the purpose of this proof-of-concept study, *C. clupeaformis* is the target species because it is the most intensively harvested species by the Lake Huron indigenous and non-Indigenous fisheries (Morh & Nalepa, 2005; LaRiviere & Crawford, 2013).

Visual identification of larval *C. clupeaformis* relies on morphological keys, notably those prepared by Auer (1982) and Cucin & Faber (1985). Such visual identification, however, can be costly, labour intensive and unreliable for samples that contain multiple individuals and species (Pfrender *et al.*, 2010). In addition, morphological species identification of larval samples can be particularly difficult because individuals may be badly damaged during sampling, and often only fragmentary remains are recovered (D. D. Ager, I. Cord & P. H. Patrick, unpubl. data). Genetic techniques, including DNA barcoding and restricted fragment length polymorphisms (RFLPs), have the potential to greatly increase accuracy and reliability of larval fish identification (Hebert, 2003; Kochzius *et al.*, 2010; Ko *et al.*, 2013) as well as the identification of fragmentary remains (Quinteiro *et al.*, 1998; Mackie *et al.*, 1999). Several genetic methods have previously been applied to identify *C. clupeaformis*, including: allozymes to identify *C. clupeaformis* populations (Casselman *et al.*, 1981); microsatellites to infer glacial lineages (Lu *et al.*, 2001); RFLPs to evaluate genetic diversity and geographic structure (Bernatchez & Dodson, 1991); and DNA barcoding for species identification (Hubert *et al.*, 2008). Although these genetic techniques are effective in identifying *C. clupeaformis*, and provide some benefits over visual identification, they can still be very costly and time consuming for large-scale implementation. For these reasons, it is necessary to develop novel genetic methods for reliably identifying *C. clupeaformis* in large, mixed samples of wild ichthyoplankton.

Real-time PCR (qPCR) has been proposed as an alternative method for species identification to address the issues of cost, time and handling of multispecies batch

samples like the mashes (collection of multiple species and individuals) resulting from ichthyoplankton surveys (Bustin, 2005). Real-time PCR has proven to be a reliable method for species identification in many fields, including pest identification (Huang *et al.*, 2010; Naaum *et al.*, 2012) and seafood market fraud (Taylor *et al.*, 2002; Rasmussen Hellberg & Morrissey, 2011). This method has also been applied successfully in several aspects of fish ecology including the identification of fish eggs (Taylor *et al.*, 2002; Bayha *et al.*, 2008; Gleason & Burton, 2011), the identification of marine fish parasites (McBeath *et al.*, 2006) and the differentiation of commercially important salmonid species (Rasmussen Hellberg *et al.*, 2010). Real-time PCR can also allow for relative quantification of individual larvae in ichthyoplankton samples (Pan *et al.*, 2008). In terms of technical advantages, real-time PCR identification can be performed rapidly, producing results in as little as 1 h, and requires no post-PCR processing steps including gel electrophoresis. In addition, real-time PCR can also be performed onsite with portable platforms, providing enhanced flexibility and responsiveness for field assessments of species identification (Naaum *et al.*, 2012). This makes real-time PCR an ideal alternative to current methods for species identification of *C. clupeaformis*.

In this study, genus-specific PCR primers and species-specific probe for real-time PCR were designed for *C. clupeaformis* using DNA barcode sequences. DNA barcoding, a method of species identification that utilizes sequence variation in the mitochondrial cytochrome *c* oxidase subunit I (COI) gene to discriminate species (Hebert *et al.*, 2003), has proven to be particularly useful for fishes (Ward *et al.*, 2009) including those from North America (Hubert *et al.*, 2008; April *et al.*, 2012). Larval fishes have been successfully identified using DNA barcoding in Australia (Pegg *et al.*, 2006), among coral reefs in the Indo-Pacific (Hubert *et al.*, 2010), along the Caribbean coast of Panama (Victor *et al.*, 2009), and along the Yucatan Peninsula, Mexico (Valdez-Moreno *et al.*, 2010). The DNA barcode region was used in the development of this assay because of the low intra-specific and high inter-specific variability of this region, and due to the high quality and availability of publicly available sequences in the Barcode of Life Data Systems (BOLD) (Ratnasingham & Hebert, 2007). Moreover, this facilitated the use of barcoding (*i.e.* Sanger sequencing) to verify species identifications.

MATERIALS AND METHODS

Larval samples were provided by third parties as preserved tissues in 95% ethanol. Larval specimens were first examined by visual identifiers. Caudal-fin clips were taken from each larval specimen. Two samples, Atlantic salmon *Salmo salar* L. 1758 and *P. cylindraceum* were collected from adult muscle samples. Each individual was additionally DNA barcoded to confirm the species identification. DNA was extracted using the Qiagen DNeasy Blood and Tissue Kits (Qiagen; www.qiagen.com/ca) following the manufacturer's instructions. To prevent cross contamination, DNA extractions were completed in a separate, sterile room from PCR, and PCR work was carried out in a PCR workstation that was UV sterilized after each run. Gel electrophoresis was also carried out in an environment kept separate from extraction and PCR.

PCR amplification for the DNA barcode region was carried out in 12·5 µl reaction volumes with 6·25 µl of 10% trehalose, 2 µl of ddH_2O, 1·25 µl of 10X buffer, 0·625 µl of $MgCl_2$ (50 mM), 0·1 µl of 10 µM each of forward and reverse primers (universal fish primers VF1i_t1 and VR1i_t1; Ivanova *et al.*, 2007); 0·0625 µl of 10 mM dNTPs, 0·06 µl of platinum Taq polymerase (5 U μl^{-1}) and 2 µl of DNA. The reaction profile was: an initial hot start at 94° C for

TABLE I. Sequences of species-specific primer and probe set for *Coregonus clupeaformis*. The amplicon length for the primers is 122 bp of the DNA barcode region. All primers and probes target the cytochrome *c* oxidase subunit I (COI) DNA barcode region

Primer name	Primer sequence (5′–3′)
Coregonus clupeaformis forward	GCC CTA AGC CTT TTA ATC
Coregonus clupeaformis reverse	GGC ATA ACT ATA AAG AAA ATC ATA
Coregonus clupeaformis probe	6FAM-CC GTG ACG ATC ACA TTA TAA ATC TGA T-BHQ1

6FAM, fluorescent reporter 6-carboxyfluorescein; BHQ1, black hole quencher 1.

120 s, followed by 40 cycles with a denaturation at 94° C for 30 s, annealing temperature of 52° C for 40 s and extension at 72° C for 60 s, followed by a final extension at 72° C for 10 min and hold at 4° C indefinitely. Amplification success of PCR products was ascertained visually using pre-cast 2% Agarose E-gels (Invitrogen; www.thermofisher.com), where a single, distinct, unambiguous band on the gel indicated successful amplification.

DNA sequencing reactions were carried out in 14 µl reaction volumes with 1 µl of BigDye v3.1, 1 µl of 5X SeqBuffer, 1 µl of 10 µM primer (VF1i_t1 for forward and VR1i_t1 for reverse), 10 µl of ddH_2O and 1·5 µl of PCR product. PCR products were bidirectionally sequenced using an ABI 3730 DNA Analyser (Applied Biosystems; www.thermofisher.com). Bi-directional sequence contig assemblies were created and edited using Sequencher v. 4.9 (Gene Codes Corporation; www.genecodes.com) and multiple sequence alignments were generated manually using MEGA 5.2 (Tamura *et al.*, 2007). Sequences were uploaded to the Barcode of Life Data System (BOLD; Ratnasingham & Hebert, 2007) public project 'Stokes Bay, Ontario, Lake Whitefish' (project code: SBOLW) and subsequently submitted to GenBank (accession numbers KP978018–KP978067) *via* BOLD. All barcodes were queried against the sequences of known provenance using the BOLD ID engine (www.boldsystems.org; Ward *et al.*, 2009) using the 'Species ID' option. A DNA barcode species match was defined as a sequence similarity greater than 99% to a unique species in the reference database.

The primers and probe for real-time PCR were designed using all North American Coregoninae DNA barcodes available publicly in BOLD. European species were excluded from probe development (Table II). A dataset entitled 'Dataset for Coregonids' (DS-CORG), given a unique digital object identification number (10.5883/DS-CORG/), was created in BOLD with all related species sequences and other Great Lakes fish sequences from the container project 'Lake Whitefish (*Coregonus clupeaformis*)' (project code: LOLW), the public projects 'Sharbot Lake Hatchery Larvae' (project code: SLHL; accession numbers KP978226–KP978312), 'Stokes Bay, Ontario, Lake Whitefish' (project code: SBOLW; accession numbers KP978018–KP978067) and 'Lake Wide Lake Huron Lake Whitefish' (project code: LWLHW; accession numbers KP978068–KP978225). Available DNA barcodes were collapsed into unique haplotypes using DNA Barcoding Tools (www.ibarcode.org; Singer & Hajibabaei, 2009). Primers and probes were designed using Allele ID 7.75 (Premier Biosoft International; www.premierbiosoft.com) using default settings. Due to the sequence similarities among coregonid species and within species complexes (*i.e. C. artedi* complex and *C. clupeaformis* complex), primers were designed to be genus specific instead of species specific (Table III). The *C. clupeaformis*-specific probe was designed and tagged with the fluorescent reporter 6-carboxyfluorescein (FAM) at the 5′ end and BQH-1 Quencher at the 3′ end (Integrated DNA Technologies; www.ididna.com). The primer and probe were screened *in silico* using the primer BLAST search by blasting the probe in the position of one of the primers (either forward or reverse) to determine if it matched any of the species (www.ncbi.nlm.nih.gov/tools/primer-blast; Ye *et al.*, 2012), with matches found to *C. clupeaformis* and humpback whitefish *Coregonus pidschian* (Gmelin 1789), for the primers and probe. Forward and reverse primers matched with the genus *Coregonus*, but did not match with the genus *Prosopium* or other non-target species tested with this assay (Table III). The

TABLE II. Species data for sequences used in the design of the primer and probe set. Barcode of Life Data Systems (BOLD; http://boldsystems.org/) project code indicates the projects that contained sequences used with the number of sequences taken from each project in parentheses; the number of sequences for each species and haplotypes per species are also indicated. Full specimen details for sequences are available from the BOLD Dataset 'Dataset of Coregonids', project code DS-CORG

Species	BOLD project code	Number of sequences	Number of haplotypes
Coregonus artedi	BCF (7), GBGCA (2), SBOLW (14)	23	5
Coregonus autumnalis	ANGBF (1), BCF (6), GBGCA (2)	9	3
Coregonus clupeaformis	BCF (8), CYTC (6), GBGCA (2), LWLHW (66), SBOLW (27), SLHL (26)	135	15
Coregonus hoyi	BCF (5), GBGCA (2), SLHL (31)	38	6
Coregonus huntsman	BCF (1), GBGCA (1)	2	2
Coregonus kiyi	BCF (1), GBGCA (1)	2	2
Coregonus laurettae	ANGBF (1), BCF (7), GBGCA (1)	9	5
Coregonus maraena	IFCZE	6	2
Coregonus nasus	ANGBF (1), BCF (7), GBGCA (4)	12	4
Coregonus nigripinnis	BCF (2), GBGCA (1)	3	3
Coregonus peled	GBGCA (1), IFCZE (6)	7	7
Coregonus pidschian	ANGBF (1), GBGCA (1)	2	1
Coregonus sardinella	ANGBF (1), BCF (6), GBGCA (2)	9	4
Coregonus zenithicus	BCF (3), GBGCA (1)	4	2
Prosopium coulterii	BCF (2), GBGCA (2)	4	4
Prosopium cylindraceum	BCF (9), GBGCA (4), LWLHW (8)	21	9
Prosopium williamsoni	BCF (11), GBGCA (5)	16	7
Stenodus leucichthys	ANGBF (1), BCF (8), GBGCA (3)	12	3

species-specific probe did not match non-target species in the genera *Prosopium* or *Coregonus*, with the exception of *C. pidschian*. It is important to note that *C. pidschian* may cause a false positive with this assay; however, because this species has an Arctic distribution, it should not affect identification of specimens from the Laurentian Great Lakes (Freyhof & Kottelat, 2008).

Optimization and initial tests were carried out on a Smart Cycler II (Cepheid; www.cepheid.com/us) platform following manufacturer's guidelines. Reactions were carried out in 25 µl volumes containing 17 µl of ddH_2O, 5 µl of template DNA (7·1 ng µl^{-1}), 1·25 µl of forward primer (0·5 µM), 1 µl of reverse primer (0·4 µM) and 0·75 µl of probe (0·3 µM). OmniMix HS lyophilized mastermix (Cepheid) was used according to the manufacturer's guidelines. PCR cycling conditions were an initial hot start of 95° C for 120 s, followed by 35 cycles of 95° C for 11 s, annealing at 62° C for 30s and extension at 72° C for 10 s for improved specificity. Standard curves were created from 10-fold serial dilutions of *C. clupeaformis* DNA using DNA from 7·1 ng µl^{-1} to 0·71 pg µl^{-1} to determine efficiency of the assay (Fig. 1). Cycle threshold (C_t) values obtained from the serial dilutions were then plotted against the logarithm of the template DNA (ng), and slope (m) was calculated by linear regression. Reaction efficiency (E) for real-time PCR was calculated, where $E = 10^{(-1m^{-1})} - 1$ (Bustin, 2005).

The probe and primer pair for *C. clupeaformis* was surveyed against larvae of the following non-target species to test specificity: (1) five individuals for two congeneric coregines *C. artedi* (6·3–14·23 ng µl^{-1}) and *C. hoyi* (7·1 ng µl^{-1}), (2) five individuals for a non-congeneric but spatially overlapping coregonine *P. cylindraceum* (78·1 ng µl^{-1}), (3) single individual of a non-coregonine salmonid *S. salar* (7·1 ng µl^{-1}), (4) single individuals of non-salmonids with spatio-temporally overlapping larvae yellow perch *Perca flavescens* (Mitchell 1814) (2·7 ng µl^{-1}), white sucker *Catostomus commersonii* (Lacépède 1803) (5·77 ng µl^{-1}) and

TABLE III. Location of forward and reverse primers on the 650 base pair amplicon of the DNA barcode region of cytochrome *c* oxidase subunit I (COI) and corresponding probe location and alignment among species of *Coregonus* (*C.*), *Prosopium* (*Pr.*), *Stenodus* (*S.*), *Perca* (*Pe.*) and *Etheostoma* (*E.*). Created by using the NCBI-Primer BLAST tool

Species	Forward primer	Probe (5′–3′) 6FAM-CC GTG ACG ATC ACA TTA TAA ATC TGA T-BHQ1	Probe location	Reverse
C. clupeaformis	40–57	.. … … … … … … … ….	98–124	138–161
C. albula	43–60	.. … ….T … … … … ….	101–127	141–164
C. artedi	40–57	.. … ….T … … … … ….	98–124	138–161
C. autumnalis	40–57	.. … ….T … … … … ….	98–124	138–161
C. hoyi	40–57	.. … ….T … … … … ….	98–124	138–161
C. huntsmani	40–57	.. … ….T … … … … ….	98–124	138–161
C. kiyi	40–57	.. … ….T … … … … ….	98–124	138–161
C. laurettae	40–57	.. … ….T … … … … ….	98–124	138–161
C. lavaretus	51–68	.. … ….T … … … … ….	109–135	149–172
C. maraena	37–54	.. … ….T … … … … ….	95–121	135–158
C. migratorius	33–50	.. … ….T … … … … ….	91–117	131–154
C. muksun	40–57	.. … ….T … … … … ….	98–124	138–161
C. nasus	40–57	.. … ….T … … … … ….	98–124	138–161
C. nigripinnis	41–61	.. … ….T … … … … ….	102–128	142–165
C. pidschian	40–57	.. … … … … … … … ….	98–124	138–161
C. peled	37–54	.. … ….T … … … … ….	95–121	135–158
C. pollan	40–57	.. … ….T … … … … ….	98–124	138–161
C. sardinella	40–57	.. … ….T … … … … ….	98–124	138–161
C. ussuriensis	25–42	.. … ….T … … … … ….	83–109	123–146
C. zenithicus	40–57	.. … ….T … … … … ….	98–124	138–161
E. exile	*	.. ..A ….C ..C ..A ..T ..T ..A .	98–124	*
Pe. flavescens	*	.C … … ….C ..A ..T ..T ..A .	98–124	*
Pr. abyssicola	*	.. ..A ….T … … ….T ….	98–124	*
Pr. coulterii	*	.T ..A … … … … … … ….	98–124	*
Pr. cylindraceum	*	.. … ….T … … ….T ….	98–124	*
Pr. gemmifer	*	.. ..A ….T … … ….T ….	98–124	*
Pr. spilonotus	*	.. ..A ….T … … ….T ….	98–124	*
Pr. williamsoni	*	.. ..A ….T … … ….T ….	98–124	*
S. leucichthys	*	.. … ….T … … … … ….	138–161	142–165

*Primers do not bind.

longnose sucker *Catostomus catostomus* (Forster 1773) (12·03 ng μl^{-1}) and (5) single individuals for other non-salmonid Great Lakes fishes found in Lake Huron, which are commonly encountered walleye *Sander vitreus* (Mitchell 1818) (7·1 ng μl^{-1}), rainbow smelt *Osmerus mordax* (Mitchell 1814) (3·7 ng μl^{-1}), ghost shiner *Notropis buchanani* Meek 1896 (8·0 ng μl^{-1}) and Iowa darter *Etheostoma exile* (Girard 1859) (4·7 ng μl^{-1}). The DNA from each of the non-target species was acquired following the aforementioned methods, using DNA extracted from the caudal-fin clips of larval specimens previously caught in Lake Huron (positively identified using DNA barcoding, with the exception of *S. salar* DNA, which was acquired from a muscle subsample of a store-bought fillet (also identified by the DNA barcode).

The *C. clupeaformis* primers and probe pair were also tested on mixtures of DNA from known species obtained from larval specimens of *C. clupeaformis*, *C. hoyi* and *S. vitreus* sampled from the Ontario Ministry of Natural Resources White Lake Fish Culture Station. The *C. clupeaformis* primers and probe were tested on the following mixtures: 50% each of *C. clupeaformis* and *C. hoyi* DNA; 50% each of *C. clupeaformis* and *S. vitreus*; 50% each of *C. hoyi* and *S. vitreus*; 33%

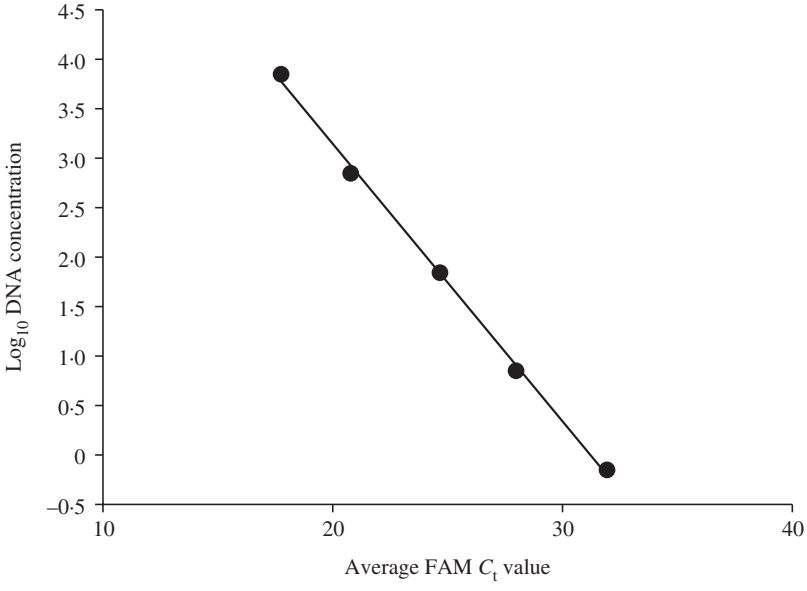

FIG. 1. Standard curve generated from 10-fold serial dilutions of three replicates of *Coregonus clupeaformis* DNA from 7·1 ng μl^{-1} to 0·71 pg μl^{-1}. FAM, fluorescent reporter 6-carboxyfluorescein. The curve was fitted by $y = -0{\cdot}2805x + 8{\cdot}7548$ ($r^2 = 0{\cdot}998$).

each of *C. clupeaformis*, *C. hoyi* and *S. vitreus*; 17% *C. clupeaformis* and 10% *C. clupeaformis* (Table IV). Each mixture was made in a volume of 30 µl, with 5 µl aliquoted into each tube for testing. DNA of non-target species was standardized to a concentration of 7·1 ng μl^{-1}, to be consistent with that of *C. clupeaformis* being used. C_t values were recorded for all samples (Table IV).

To assess the utility of this assay on a high-throughput instrument, it was also evaluated using a StepOnePlus Real-Time PCR system (Applied Biosystems). All reactions were carried out using MicroAmp fast 96-well reaction plates (100 µl) (Applied Biosystems), containing the same volumes and concentrations as used on the Cepheid SmartCycler II. To set an appropriate comparable threshold on the new instrument and based on recommendations from Applied Biosystems User Guide, the fluorescence threshold was manually set to 8000 to ensure it was in the amplification phase based on reference target samples of *C. clupeaformis* previously run on the Cepheid SmartCycler II. One replicate of each of the same non-target species and mixtures (as above), as well as DNA from six individuals of *C. clupeaformis* were tested on the StepOnePlus.

Assay specificity was validated in two ways using melt-curve analysis and gel electrophoresis. First, melt-curve analysis using PerfeCTa SYBR Green FastMix (Quanta BioScience Inc.; www.quantabio.com) was carried out on the Cepheid Smart Cycler II. Reactions were carried out in 25 µl volumes containing 5·25 µl of ddH_2O, 5 µl of template DNA, 1·25 µl of forward primer (0·5 µM), 1 µl of reverse primer (0·4 µM), and 12·5 µl of SYBR Green. Four samples were tested including a no-template control (5 µl of ddH_2O), two non-target species (5 µl of *C. commersonii* and *C. catostomus*, respectively) and one target species (5 µl of *C. clupeaformis*) to determine the presence of a single melt peak with duplicates of each sample. Specificity was also validated by gel electrophoresis by placing 8 µl of qPCR product in a 2% agarose gel in Tris/Borate/EDTA (TBE) buffer. Amplification of a product, and product size were used to verify the specificity of the assay for the target species.

To address issues of non-target amplification observed in initial testing, three independent samples from larval *P. flavescens* were re-tested in duplicate on both the Cepheid SmartCycler II and the StepOnePlus Real-Time PCR system. These larval *P. flavescens* were collected for a

TABLE IV. DNA mixtures containing different proportions of *Coregonus clupeaformis* (CC) (7·1 ng μl^{-1}) mixed with non-target species *Coregonus hoyi* (CH) (7·1 ng μl^{-1}) and *Sander vitreus* (SV) (7·1 ng μl^{-1}). Absolute concentration of CC DNA in each mixture provided. Cycle threshold (C_t) values for each replicate shown along with average C_t value

Proportion of CC (%)	CC (μl)	SV (μl)	CH (μl)	C_t	C_t average	(CC)
50	15	0	15	18·09	17·97	5·13
				18·2		
				17·61		
50	15	15	0	18·84	18·27	4·3
				18·12		
				17·84		
0	0	15	15	NA	NA	0
				NA		
				NA		
33	10	10	10	18·51	18·44	3·4
				18·52		
				18·72		
17	5	10	15	20·09	20·09	1·16
				19·98		
				20·2		
10	3	12	15	20·42	20·27	1·16
				20·3		
				20·09		

previous study from hatchery reared fish and were collected and stored independently in ethanol, which differs from the other test samples that were collected from the wild in larval tows; this has the potential to increase risk for surface cross contamination of individuals as they were collected and stored together.

RESULTS

Using the Cepheid SmartCycler II, the primer and probe set was tested against nine non-target species, including both *C. hoyi* and *C. artedi*, which are two closely related species to *C. clupeaformis* and also found in the Great Lakes. No amplification was observed with *C. hoyi*, *C. artedi*, *S. salar*, *C. commersonii*, *C. catostomus* or *S. vitreus*. Some amplification was observed with non-target species: *P. flavescens* ($C_t = 29·66$) and *E. exile* ($C_t = 31·99$). Three independent samples from larval *P. flavescens* were re-tested with duplication on both the Cepheid SmartCycler II and the StepOnePlus Real-Time PCR system and showed no non-target amplification. No false negatives were observed, and no signal was observed for no-template blank controls when testing the optimal protocol on *C. clupeaformis* DNA or non-target species. The efficiency of the reaction, as calculated from the standard curve ($r^2 = 0·998$) was 90·77% (Fig. 1). The assay successfully identified target DNA in as low as 0·1% proportion of a DNA mixture (Table IV).

TABLE V. Validation of protocol and primer and probe set on StepOnePlus Real-Time PCR machine and comparison of cycle threshold (C_t) values. All samples extracted from larval caudal-fin clips, except for *Salmo salar* and *Prosopium cylindraceum*, which were collected from adult muscle samples. See Table IV for (CC) DNA (ng μl^{-1}) in mixtures 1–5

Species	C_t (Cepheid)	C_t (StepOnePlus)
Coregonus clupeaformis	17·49	17·81
C. clupeaformis	20·56	21·58
C. clupeaformis	21·89	22·06
C. clupeaformis	20·49	20·98
C. clupeaformis	20·68	21·23
Salmo salar	NA	NA
Etheostoma exile	31·99	31·67
Perca flavescens	29·66	29·89
Osmerus mordax	NA	NA
Notropis buchanani	NA	NA
Coregonus artedi	NA	NA
Coregonus hoyi	NA	30·31
Prosopium cylindraceum	NA	NA
Catostomus commersonii	NA	NA
Catostomus catostomus	NA	NA
Mixture 1 (CC:CH)	17·97	19·22
Mixture 2 (CC:SV)	18·27	18·39
Mixture 3 (SV:CH)	NA	NA
Mixture 4 (CC:CH:SV)	18·44	19·45
Mixture 5 (CC:CH:SV)	20·09	20·29

CC, *Coregonus clupeaformis*; CH, *Coregonus hoyi*; SV, *Sander vitreus*; NA, no C_t value (fluorescence unobserved).

The protocol, primers and probe set were also evaluated on a StepOnePlus Real-Time PCR system, a platform capable of high-throughput analysis. This assay was successfully used to identify the target species, *C. clupeaformis*, with comparable C_t values to those obtained using the Cepheid SmartCycler II for the same samples (Table V). No false negatives were observed, and no signal was observed for no-template controls. No amplification was observed for non-target species *S. salar*, *C. artedi*, *O. mordax*, *C. commersonii*, *C. catostomus* or *P. cylindraceum*. Some fluorescence output was observed with *C. hoyi* ($C_t = 30{\cdot}31$), *P. flavescens* ($C_t = 29{\cdot}89$) and *E. exile* ($C_t = 31{\cdot}67$) (Table V).

Specificity of melt-curve validation of qPCR primers revealed a single melt peak for *C. clupeaformis* at 81·8° C, indicating a single PCR product. This melt peak is comparable to the estimated melting temperature of assay product calculated by Allele ID ($T_m = 79{\cdot}2$° C). No melt peaks were observed in no-template controls or in replicates of non-target species, therefore validating the specificity of the *C. coregonus* forward and reverse primer (Fig. 2). Specificity was further tested by agarose gel electrophoresis. No-template controls, non-target species and *C. coregonus* were amplified using universal fish cocktail primers and then compared against qPCR product of no-template control, non-target species and *C. coregonus* using the *C. coregonus* primers (Fig. 3). Bands were present for only *C. coregonus* when

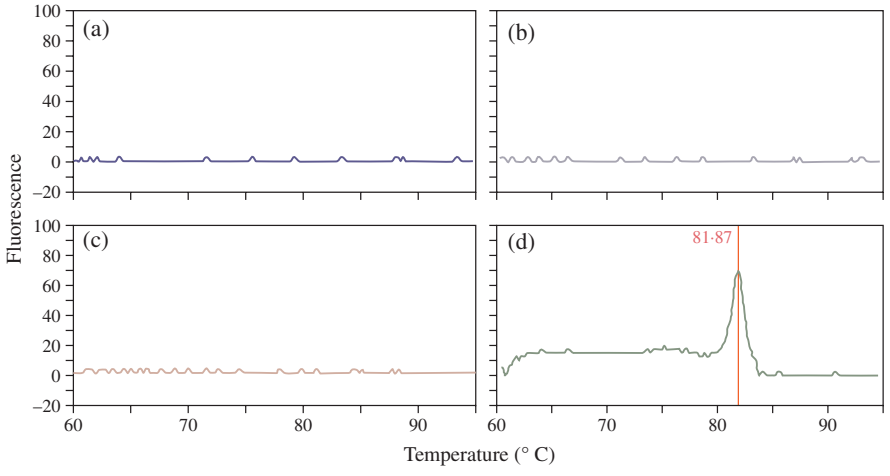

FIG. 2. Melt-curve analysis peaks for (a) no-template control, (b) non-target species *Catostomus catostomus*, (c) non-target species *Catostomus commersonii* and (d) target species *Coregonus clupeaformis*. Only one distinct melt peak is present in (d) for *Coregonus clupeaformis*. No secondary peaks indicate no primer dimer formation.

visualizing products of qPCR at the expected amplicon length of 122 bp for the product of qPCR.

DISCUSSION

A real-time assay was successfully developed that can reliably differentiate between *C. clupeaformis* and other local coregonines found in Lake Huron (Table I). Of

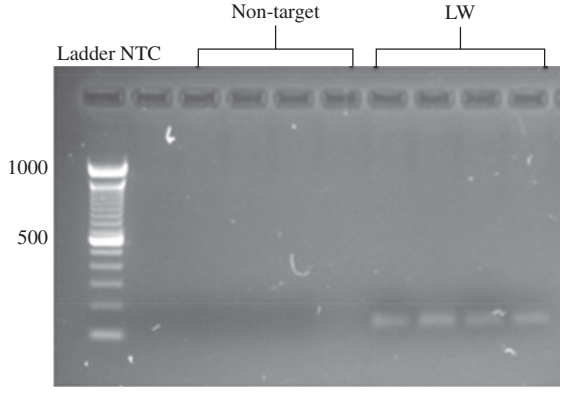

FIG. 3. Results of agarose gel electrophoresis of real-time PCR product using *Coregonus*-specific primers designed in this study. NTC, no-template control, non-target in lanes 3–6 include two replicates of *Catostomus catostomus* (lanes 3 and 4) and *Catostomus commersonii* (lanes 5 and 6). Bands observed for *Coregonus clupeaformis* (LW) in lanes 7–10 at *c.* 120 bp, which is expected length of product amplicon from real-time PCR. No bands present for the no-template control and non-target species.

particular importance is the ability to differentiate larval *C. clupeaformis* from two other congeneric species, *C. hoyi* and *C. artedi*, which can be easily misidentified as *C. clupeaformis* during the larval period (Scott & Crossman, 1973; Auer, 1982; Todd & Stedman, 1989). The assay correctly identified all target samples of *C. clupeaformis*, with no false negatives observed. Although some amplification was observed for non-target species, the C_t values were high ($C_t = 29{\cdot}66–31{\cdot}99$) compared to the target species ($C_t = 17{\cdot}81–22{\cdot}06$) and this was probably due to contamination of DNA samples when larval fish were initially bulk stored together in ethanol at the time of field collection. It is also important to note that the non-target genomic DNA was tested at an equivalent or lower concentration (2·7, 4·7 or 7·1 ng μl^{-1}) than the target genomic DNA (7·1 ng μl^{-1}). With repeated DNA extractions from larval *P. flavescens* stored separately in ethanol, nonspecific amplification was eliminated, suggesting that the original amplification observed was due to contaminated DNA samples resulting from combined storage of multiple species prior to the DNA extraction. The resampling of *C. hoyi* and *E. exile* from independent samples was not feasible for the purposes of this investigation. The differences in primer and probe binding regions (Table III), and the absence of signal observed in the retested *P. flavescens* samples, however, suggest that this was also the case for these other two non-target species, which were collected and stored in the same manner as the larval *P. flavescens* that were initially tested. Potential contamination will always be an issue when dealing with wild-caught ichthyoplankton using tows or other forms of larval sampling where samples are comingled. Organisms are caught in the net and held in the cod end with each other until the time of processing, which has the potential to cause surface contamination of individual larvae and the preservation fluid they are stored in.

The assay developed in this study represents a rapid identification method that can be implemented in the field, or in high-throughput laboratory facilities, as required. The assay is a good complement and alternative to visual identification methods, more efficient than other time- and cost-intensive genetic methods of identification and provides an effective solution to identifying individuals in wild larval fish assemblages. For example, the time and cost associated with visually identifying a single ichthyoplankton specimen is estimated to cost \$50–100 per h per fish (Golder Associates, unpubl. data), while DNA barcoding costs at least \$40 per sample (E. Zakharov, pers. comm.; Canadian Centre for DNA Barcoding www.ccdb.ca) and requires a minimum of 48 h to complete; while qPCR can be accomplished for as little as \$5 in 2 h for 96 specimens using a high-throughput machine like the StepOnePlus Real-Time PCR system, once an assay has been developed and optimized. Real-time PCR utilizing this assay has a wide breadth of application in larval ecology including fragmentary analysis (Jackson *et al.*, 2012), gut content analysis (Jarman *et al.*, 2004) and plankton enumeration and identification (Coyne *et al.*, 2005).

In addition to the above applications, this real-time assay can be effectively and efficiently applied to major environmental assessments of industrial effects on ichthyoplankton, including the entrainment of fish larvae associated with cooling water intake structures at power-generating and manufacturing facilities. For example, a major hindrance to accurately estimating entrainment rates at power plants is the identification of species composition (The Committee on Entrainment, 1978). Visual identification of entrainment samples remains a challenge because of the degradation of samples (Azila & Chong, 2010; Rabin, 2010), yet this information is key to assessing ecological impact. The success of the assay described in this laboratory study, and the validation of

the primer and probe set on both a portable and on a high-throughput platform, suggests that this method would be a suitable means to address the presences of a target species in entrainment ichthyoplankton samples. Future research involving the expansion of this method to include other species of commercial and ecological importance in the Great Lakes that may be at risk of entrainment is highly feasible. The use of multiple primer and probe sets in a multiplexed real-time PCR assay would simultaneously identify several species of interest, further reducing the costs of monitoring the presence of these species.

This research was supported by the Saugeen Ojibway Nation (Chippewas of Nawash Unceded First Nation, Saugeen First Nation) in collaboration with Bruce Power Limited (http://www.brucepower.com). We would also like to thank N. Freeman, K. Mulligan, G. Burke, A. Wincikaby, R. Eberts (University of Regina), C. Somers (University of Regina), W. L. Stott (USGS), T. Drew (White Lake Fish Culture Station, Sharbot Lake), C. Cumpseth and C. Ho (BOLD for bioinformatics support).

References

April, J., Mayden, R. L., Hanner, R. H. & Bernatchez, L. (2012). 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.

Auer, N. A. (Ed.) (1982). Identification of larval fishes of the Great Lakes basin with emphasis on the Lake Michigan drainage. *Great Lakes Fishery Commission Special Publication* **82–83**, 1–744.

Azila, A. & Chong, V. C. (2010). Multispecies impingement in a tropical power plant, Straits of Malacca. *Marine Environmental Research* **70,** 13–25.

Bayha, K. M., Graham, W. M. & Hernandez, F. J. Jr. (2008). Multiplex assay to identify eggs of there fish species from the northern Gulf of Mexico, using locked nucleic acid Taqman real-time PCR probes. *Aquatic Biology* **4,** 65–73.

Bernatchez, L. & Dodson, J. J. (1991). Phylogeographic structure in mitochondrial DNA of the Lake Whitefish (*Coregonus clupeaformis*) in North America and its relationships to Pleistocene glaciations. *Evolution* **45,** 1016–1035.

Bustin, S. A. (2005). *Real-Time PCR. Encyclopedia of Diagnostic Genomics and Proteomics*. London: University of London.

Casselman, J. M., Collins, J. J., Crossman, E. J., et al. (1981). Lake whitefish (*Coregonus clupeaformis*) stocks of the Ontario waters of Lake Huron. *Canadian Journal of Fisheries and Aquatic Sciences* **38,** 1772–1789.

Coyne, K. J., Handy, S. M., Demir, E., Whereat, E. B., Hutchins, D. A., Portune, K. J., Doblin, M. A. & Cary, S. C. (2005). Improved quantitative real-time PCR assays for enumeration of harmful algal species in field samples using an exogenous DNA reference standard. *Limnology and Oceanography: Methods* **3,** 381–391.

Cucin, D. & Faber, D. J. (1985). Early life studies of lake whitefish (*Coregonus clupeaformis*), cisco (*Coregonus artedi*) and yellow perch (*Perca flavescens*) in Lake Opeongo, Ontario. *Ontario Fisheries Technical Report Series Number* **16.**

Davis, B. M. & Todd, T. N. (1998). Competition between larval lake herring (*Coregonus artedi*) and lake whitefish (*Coregonus clupeaformis*) for zooplankton. *Canadian Journal of Fisheries and Aquatic Sciences* **55,** 1140–1148.

Gleason, L. U. & Burton, R. S. (2011). High-throughput molecular identification of fish eggs using multiplex suspension bead arrays. *Molecular Ecology Resources* **12,** 57–66.

Hare, J. A., Cowen, R. K., Zehr, J. P., Juanes, F., & Day, K. H. (1994). Biological and oceanographic insights from larval labrid (Pisces: Labridae) identification using mtDNA sequences. *Marine Biology* **118,** 17–24.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B* **270,** 313–321.

Huang, K. S., Lee, S. E., Yeh, Y., Shen, G. S., Mei, E., Chang, C. M. (2010). Taqman real-time quantitative PCR for identification of western flower thrip (*Frankliniella occidentalis*) for plant quarantine. *Biology Letters* **6,** 555–557.

Hubert, N., Hanner, R., Holm, E., Mandrak, N. E., Taylor, E., Burridge, M., Watkinson, D., Dumont, P., Curry, A., Bentzen, P., Zhang, J., April, J. & Bernatchez, L. (2008). Identifying Canadian freshwater fishes through DNA barcoding. *PLoS One* **3,** e2490. doi: 10.1371/journal.pone.0002490

Hubert, N., Delrieu-Trottin, E., Irisson, J.-O., Meyer, C. & Planes, S. (2010). Identifying coral reef fish larvae through DNA barcoding: a test case with the families Acanthuridae and Holocentridae. *Molecular Phylogenetics and Evolution* **5S**, 1195–1203. doi: 10.1016/j.ympev.2010.02.023

Ivanova, N. V., Zemlak, T. S., Hanner, R. H. & Hebert, P. D. N. (2007). Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes* **7,** 544–548.

Jackson, C. B., Gallati, S. & Schaller, A. (2012). qPCR-based mitochondrial DNA quantification: influence of template DNA fragmentation on accuracy. *Biochemical and Biophysical Research Communications* **423,** 441–447.

Jarman, S. N., Deagle, B. E. & Gales, N. J. (2004). Group-specific polymerase chain reaction for DNA-based analysis of species diversity and identity in dietary samples. *Molecular Ecology* **13,** 1313–1322.

Ko, H. L., Wang, Y. T., Chiu, T. S., Lee, M. A., Leu, M. Y., Chang, K. Z., Chen, W. Y., & Shao, K. T. (2013). Evaluating the accuracy of morphological identification of larval fishes by applying DNA barcoding. *PLoS One* **8,** e53451. doi: 10.1371/journal.pone.0053451

Kochzius, M., Nölte, M., Weber, H., Silkenbeumer, N., Hjörleifsdottir, S., Hreggvidsson, G. O., Marteinsson, V., Kappel, K., Planes, S., Tinti, F., Magoulas, A., Garcia Vazquez, E., Turan, C., Hervet, C., Campo Falgueras, D., Antoniou, A., Landi, M. & Blohm, D. (2008). DNA microarrays for identifying fishes. *Mar Biotechnol* **10**, 207–217.

Kochzius, M., Seidel, C., Antoniou, A., Botla, S. K., Campo, D., Cariani, A., Garcia, E. G. Vazquez, Hauschild, J., Hervet, C., Hjörleifsdottir, S., Hreggvidsson, G., Kappel, K., Landi, M., Magoulas, A., Marteinsson, V., Nölte, M., Planes, S., Tinti, F., Turan, C., Venugopal, M. N., Weber, H. & Blohm, D. (2010). Identifying fishes through DNA barcodes and microarrays. *PLoS One* **5,** e12620. doi: 10.1371/journal.pone.0012620

LaRiviere, C. M. & Crawford, S. S. (2013). Indigenous principles of wild harvest and management: an Ojibway community as a case study. *Human Ecology* **41,** 947–960.

Lu, G., Basley, D. J. & Bernatchez, L. (2001). Contrasting patterns of mitochondrial DNA and microsatellite introgressive hybridization between lineages of lake whitefish (*Coregonus clupeaformis*); relevance for speciation. *Molecular Ecology* **10,** 965–985.

Mackie, I. M., Pryde, S. E., Gonzales-Sotelo, C., Medina, I., Peréz-Martín, R., Quinteiro, J., Rey-Mendez, M. & Rehbein, H. (1999). Challenges in the identification of species of canned fish. *Trends in Food Science and Technology* **10,** 9–14.

McBeath, A. J. A., Penston, M. J., Snow, M., Cook, P. F., Bricknell, I. R. & Cunningham, C. O. (2006). Development and application of real-time PCR for specific detection of *Lepeophtheirus salmonis* and *Caligus elongates* larvae in Scottish plankton samples. *Diseases of Aquatic Organisms* **73,** 141–150.

Morh, L. C. & Nalepa, T. F. (Eds) (2005). Proceedings of a workshop on the dynamics of lake whitefish (*Coregonus clupeaformis*) and the amphipod *Diporeia* spp. in the Great Lakes. *Great Lakes Fishery Commission Technical Report* **66.**

Naaum, A. M., Foottit, R. G., Maw, H. E. L. & Hanner, R. (2012). Differentiation between *Aphis pomi* and *Aphis Spiraecola* using multiplex real-time PCR based on DNA barcode sequences. *Journal of Applied Entomology* **136,** 704–710.

Pan, M., McBeath, A. J. A., Hay, S. J., Pierce, G. J., Carey, O. & Cunningham, C. O. (2008). Real-time PCR assay for detection and relative quantification of *Liocarcinus depurator* larvae from plankton samples. *Marine Biology* **153,** 859–870.

Pegg, G. G., Sinclair, B., Briskley, L. & Aspden, W. J. (2006). MtDNA barcode identification of fish larvae in the southern Great Barrier Reef-Australia. *Scientia Marina* **70**(Suppl. 2), 7–12.

Pfrender, M. E., Hawkins, C. P., Bagley, M., Courtney, G. W., Creutzburg, B. R., Epler, J. H., Fend, S., Schindel, D., Ferrington, L. C. Jr., Hartzell, P. L., Jackson, S., Larsen, D. P.,

Lévesque, A., Morse, J. C., Petersen, M. J., Ruiter, D., & Whiting, M. (2010). Assessing macroinvertebrate biodiversity in freshwater ecosystems: advances and challenges in DNA-based approaches. *Quarterly Review of Biology* **85,** 319–340.

Quinteiro, J. C., Sotelo, G., Rehbein, H., Pryde, S. E., Medina, I., Pérez-Martín, R. I., Rey-Méndez, M. & Mackie, I. M. (1998). Use of mtDNA direct polymerase chain reaction (PCR) sequencing and PCR-restriction fragment length polymorphism methodologies in species identification of canned tuna. *Journal of Agricultural and Food Chemistry* **46,** 1662–1669.

Rabin, K. (2010). *Reeling in New York's Aging Power Plants: The Case for Fish-Friendlier Power*. New York, NY: Network for New Energy Choices.

Rasmussen Hellberg, R. S. & Morrissey, M. T. (2011). Advances in DNA-based techniques for the detection of seafood species substitution on the commercial market. *Journal of Laboratory Automation.* **16,** 308–322.

Rasmussen Hellberg, R. S., Morrissey, M. T. & Hanner, R. H. (2010). A multiplex PCR method for the identification of commercially important salmon and trout species (*Oncorhynchus* and *Salmo*) in North America. *Journal of Food Science* **75,** C595–C606.

Ratnasingham, S. & Hebert, P. D. N. (2007). BOLD: the barcode of life data system (http://www.barcodinglife.org). *Molecular Ecology Notes* **7,** 355–364.

Roth, B. M., Mandrak, N. E., Hrabik, T. H., Sass, G. G. & Peters, J. (2012). Fishes and decapod crustaceans of the Great Lakes Basin. In *Great Lakes Policy and Management: A Binational Perspective* (Taylor, W. W. & Lynch, A., eds), pp. 105–136. East Lansing, MI: Michigan State University Press.

Schlei, O. L., Crete-Lafreniere, A., Whiteley, A. R., Schlei, O. L., Crête-Lafrenière, A., Whiteley, A. R., Brown, R. J., Olsen, J. B., Bernatchez, L. & Wenburg, J. K. (2008). DNA barcoding of eight North American coregonine species. *Molecular Ecology Resources* **8,** 1212–1218.

Scott, W. B. & Crossman, E. J. (1973). Freshwater fishes of Canada. *Fisheries Research Board of Canada Bulletin* **184**.

Singer, A. C. & Hajibabaei, M. (2009). iBarcode.org: web-based molecular biodiversity analysis. *BMC Bioinformatics* **10**(Suppl. 6), S14.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24,** 1596–1599.

Taylor, M. I., Fox, C., Rico, I. & Rico, C. (2002). Species-specific TaqMan probes for simultaneous identification of (*Gadus morhua* L.), haddock (*Melanogrammus aeglefinus* L.) and whiting (*Merlangius merlangus* L.) *Molecular Ecology Notes* **2,** 599–601.

Teletchea, F. (2009). Molecular identification methods of fish species: reassessment and possible applications. *Reviews in Fish Biology and Fisheries* **19,** 265–293.

The Committee on Entrainment (1978). Introduction. In *Power Plant Entrainment: A Biological Assessment* (Schubel, J. R. & Marcy, B. C. Jr., eds), pp. 1–18. New York, NY: Academic Press.

Todd, T. N. & Smith, G. R. (1992). A review of differentiation in Great Lakes ciscoes. In *Biology and Management of Coregonid Fishes* (Todd, T. N. & Luczynski, M., eds), pp. 13–19. *Polski Arch Iwuh Hydrobiologii* **39.**

Todd, T. N. & Stedman, R. M. (1989). Hybridization of ciscoes (*Coregonus* spp.) in Lake Huron. *Canadian Journal of Zoology* **67,** 1679–1685.

Valdez-Moreno, M., Vasquez-Yeomans, L., Elias-Gutierrez, M., Ivanova, N. V. & Hebert, P. D. N. (2010). Using DNA barcodes to connect adults and early life stages of marine fishes from the Yucatan Peninsula, Mexico: potential in fisheries management. *Marine and Freshwater Research* **61,** 655–671.

Victor, B. C., Hanner, R., Shivji, M., Hyde, J. & Caldow, C. (2009). Identification of the larval and juvenile stages of the Cubera Snapper, *Lutjanus cyanopterus*, using DNA barcoding. *Zootaxa* **2215,** 24–36.

Ward, R. D., Hanner, R. & Hebert, P. D. N. (2009). The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology* **74,** 329–356.

Ye, L., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S. & Madden, T. L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13,** 134.

Electronic Reference

Freyhof, J. & Kottelat, M. (2008). *Coregonus pidschian*. In *IUCN 2011. IUCN Red List of Threatened Species. Version 2011.1* Available at http://www.iucnredlist.org/details/5375/0/