

University of Windsor

Scholarship at UWindsor

Great Lakes Institute for Environmental
Research Publications

Great Lakes Institute for Environmental
Research

11-29-2017

Environmental DNA detection of rare and invasive fish species in two Great Lakes tributaries

Katherine D. Balasingham
University of Windsor

Ryan P. Walter
University of Windsor

Nicholas E. Mandrak

Daniel D. Heath
University of Windsor

Follow this and additional works at: <https://scholar.uwindsor.ca/glierpub>



Part of the [Biochemistry, Biophysics, and Structural Biology Commons](#), [Biodiversity Commons](#), [Biology Commons](#), and the [Marine Biology Commons](#)

Recommended Citation

Balasingham, Katherine D.; Walter, Ryan P.; Mandrak, Nicholas E.; and Heath, Daniel D.. (2017). Environmental DNA detection of rare and invasive fish species in two Great Lakes tributaries. *Molecular Ecology*, 27, 112-127.
<https://scholar.uwindsor.ca/glierpub/129>

This Article is brought to you for free and open access by the Great Lakes Institute for Environmental Research at Scholarship at UWindsor. It has been accepted for inclusion in Great Lakes Institute for Environmental Research Publications by an authorized administrator of Scholarship at UWindsor. For more information, please contact scholarship@uwindsor.ca.

MS. KATHERINE D BALASINGHAM (Orcid ID : 0000-0002-8355-3051)

Article type : Original Article

Environmental DNA detection of rare and invasive fish species in two Great Lakes tributaries

Katherine D. Balasingham^a

Ryan P. Walter^{a,c,d}

Nicholas E. Mandrak^{b,d}

Daniel D. Heath^a

^a Department of Biology, Great Lakes Institute for Environmental Research, University of Windsor,
2990
Riverside Drive West, ON, N9C 1A2, Canada

^b Department of Biological Sciences, University of Toronto Scarborough, 1265 Military Trail, Toronto,
ON M1C 1A4, Canada

^c *Present Address:* Department of Biological Sciences, California State University, 800 N State College
Blvd, Fullerton, CA 92831, United States

^d Great Lakes Laboratory for Fisheries and Aquatic Sciences, Fisheries and Oceans Canada,
Burlington, ON L7R 4A5, Canada

Running Title: eDNA analysis of riverine fish community

Keywords:

environmental DNA, next-generation sequencing, endangered species, invasive, fish community

Correspondence:

Katherine D. Balasingham

Department of Biology

Current/Permanent Address:

123 Irenemount Crescent, Markham

ON, L3S 3C3, Canada

Contact: 1-647-588-5799

Email: katherine_balasingham@hotmail.com

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mec.14395

This article is protected by copyright. All rights reserved.

Abstract

The extraction and characterization of DNA from aquatic environmental samples offers an alternative, non-invasive approach for the detection of rare species. Environmental DNA, coupled with PCR and next-generation sequencing (“metabarcoding”), has proven to be very sensitive for the detection of rare aquatic species. Our study used a custom designed group-specific primer set and next-generation sequencing for the detection of three species at risk; (Eastern Sand Darter, *Ammocrypta pellucida*; Northern Madtom, *Noturus stigmosus*; and Silver Shiner, *Notropis photogenis*), one invasive species (Round Goby, *Neogobius melanostomus*) and an additional 78 native species from two large Great Lakes tributary rivers in southern Ontario, Canada; the Grand River and the Sydenham River. Out of 82 fish species detected in both rivers using capture-based and eDNA methods, our eDNA method detected 86.2% and 72.0% of the fish species in the Grand River and the Sydenham River, respectively, which included our four target species. Our analyses also identified significant positive and negative species co-occurrence patterns between our target species and other identified species. Our results demonstrate that eDNA metabarcoding that targets the fish community as well as individual species of interest provides a better understanding of factors affecting the target species spatial distribution in an ecosystem than possible with only target species data. Additionally, eDNA is easily implemented as an initial survey tool, or alongside capture-based methods, for improved mapping of species distribution patterns.

Introduction

The need to monitor aquatic species distributions without the using invasive capture methods has led to the development of alternative sampling approaches such as the analysis of environmental DNA (eDNA). In aquatic environments, the eDNA of species is derived from biological sources such as urine, shed skin cells, eggs and sperm, blood, and feces and may remain detectable in the system for extended periods of time (Ficetola et al. 2008; Taberlet et al. 2012; Thomsen et al. 2012; Balasingham et al. 2016). The DNA can then be extracted from water samples using non-

invasive sampling approaches established for eDNA analyses (Bohmann et al. 2014; Barnes & Turner 2016). This is especially useful for species at risk (SAR) that are endangered, threatened, or extirpated in Canada according to the *Species at Risk Act* (SARA; SARA 2016) and are, thus, present at low abundances (i.e. rare). Several studies have employed eDNA methods for SAR detection successfully; for example, Chinook Salmon (*Oncorhynchus tshawytscha*) in the Upper Columbia River (Laramie et al. 2014), Japanese Giant Salamander (*Andrias japonicus*) in the Katsura River basin in Japan (Fukumoto et al. 2015), European Weather Loach (*Misgurnus fossilis*) in Jutland, Denmark (Sigsgaard et al. 2015), and Eastern Hellbender (*Cryptobranchus alleganiensis*) in North Carolina, USA (Spear et al. 2015) and in Pennsylvania (Pitt et al. 2017).

With successful applications of eDNA analyses to species with low abundance, eDNA approaches have expanded to encompass invasive species early in their invasion (i.e. still few in number) to facilitate the implementation of early eradication or control programs (Anderson 2005; Hulme 2006). Recent work on invasive species detection using eDNA methods include the detection of American Bullfrog (*Lithobates catesbeianus*) in ponds throughout southern France (Dejean et al. 2012), Bluegill (*Lepomis macrochirus*) in Japanese ponds (Takahara et al. 2013; Doi et al. 2015), Asian carps (*Hypophthalmichthys* spp.) in northern USA (Jerde et al. 2011; Mahon et al. 2013), New Zealand Mudsnail (*Potamopyrgus antipodarum*) in streams in Idaho, USA (Goldberg et al. 2013), and the spread of Ruffe (*Gymnocephalus cernua*) in the Great Lakes (Tucker et al. 2016). With increased detection sensitivity, less impact on the sampled ecosystem and ease of applicability, eDNA analysis provides an alternative, valuable sampling method for aquatic species.

While eDNA has been used extensively for determining the distribution of target species, simple presence data does not provide an ecological context for conservation, management, and invasion risk assessments. Sampled water should contain DNA from all inhabitants in the system and could thus be utilized to characterize whole communities (Yamanaka & Minamoto 2016). Moreover,

Accepted Article

gathering presence data for the whole community indirectly benefits ecosystem management as it monitors for unknown species (e.g. recently introduced species) as a “passive surveillance” of the system (Simmons et al. 2016). By targeting highly conserved regions that flank polymorphic regions (Thomsen & Willerslev 2015) for PCR primer design, NGS of the resulting amplicons can be used for metabarcoding of the whole community, with a high likelihood of detecting cryptic, elusive, or rare species that are often missed using capture or observation-based methods (Pompanon et al. 2011; Port et al. 2016; Evans et al. 2016; Valentini et al. 2016). This method of metabarcoding has also been shown to be semi-quantitative when using spiked samples and spiked communities (Diaz-Real 2015; Pereira et al. 2017).

While eDNA analyses of target species distribution and community composition are powerful, the results must be interpreted with caution. eDNA quantity and quality is affected by biotic factors such as life stage (Maruyama et al. 2014; Fukumoto et al. 2015; Spear et al. 2015) and biomass (Pilliod et al. 2014; Klymus et al. 2015), and abiotic factors such as temperature (Strickler et al. 2015), turbidity (Diffey et al. 2002), flow rate and water chemistry (Takahara et al. 2012; Deiner & Altermatt 2014; Turner et al. 2015; Balasingham et al. 2016). eDNA detection sensitivity is also affected by sampling protocol (Bohmann et al. 2014; Renshaw et al. 2015; Turner et al. 2015) and DNA extraction methods (Matheson et al. 2010; McKee et al. 2015; Eichmiller et al. 2016); however, eDNA protocols continue to improve to produce reliable and sensitive eDNA analyses (Jerde et al. 2011; Jerde et al. 2015; Dejean et al. 2012; Rees et al. 2014; McKee et al. 2015).

Effective management and conservation of aquatic ecosystems require data on the distribution of rare and invasive species. The inclusion of community composition data thus provides an ecological perspective on SAR spatial patterns in relation to the presence of other native or invasive species. In this study, we characterized the spatial distribution of three SAR, one invasive species, and common native species in two large Great Lakes tributary rivers, the Grand River and

the Sydenham River, using eDNA metabarcoding methods. We used a custom-designed group-specific PCR primer for the CO1 mtDNA gene designed to target all known fishes in the Great Lakes basin (identified using historical capture-based survey data). We used NGS and intensive water sampling to maximize data gathered on the spatial relationships among SAR, invasive species, and the native fish community composition. Our use of combined eDNA and NGS methods (metabarcoding) is a powerful approach for analyzing community composition and co-occurrence to investigate important species interactions, including rare, invasive, and common native species. Future studies designed to detect the distribution of target species should use this metabarcoding with group-specific primers and custom reference species sequence databases for optimal detection of species of interest. Our approach to metabarcoding simultaneously generates critical information of shared habitats and potential ecological interactions of target species, especially in regard to SAR and invasive species.

Materials & Methods

Study Area

The Sydenham River is a tributary of Lake St. Clair (Fig. 1), which historically had a high level of aquatic biodiversity; approximately 80 fish species (Staton et al. 2003; Marson & Mandrak 2009). Currently, 51 fish species have been reported (SCRCA 2013a; SCRCA 2013b) in the Sydenham River. The Sydenham River has two main branches, the North Sydenham River, which is approximately 70 km long with a mean annual discharge of $16.9 \text{ m}^3 \cdot \text{s}^{-1}$, and the East Sydenham River which is approximately 100 km long with a mean annual discharge of $34.4 \text{ m}^3 \cdot \text{s}^{-1}$. The total catchment drainage area of both branches is $2\,725 \text{ km}^2$ (Metcalf-Smith et al. 2003; Staton et al. 2003).

The Grand River drains into Lake Erie (Fig. 1) and is in the largest watershed in southern Ontario, approximately 300 km long with catchment drainage area of 6800 km^2 (Li et al. 2016). It has an annual mean discharge rate of approximately $64 \text{ m}^3 \cdot \text{s}^{-1}$ (Singer & So 1980). The watershed has

more than half of all freshwater fish species in Ontario (158 species; Mandrak & Crossman 1992), with 83 species present in Grand River in 1999 (Wright & Imhof 2001).

Study Species

The four target species in this study include three SAR; Eastern Sand Darter (*Ammocrypta pellucida*), Northern Madtom (*Noturus stigmosus*), and Silver Shiner (*Notropis photogenis*), and one invasive species, the Round Goby (*Neogobius melanostomus*). The target SAR have overlapping historical or current distributions in at least one of the two study rivers. The invasive Round Goby is known to be present at high densities at many locations in both study rivers, with known detrimental impacts on many native and at-risk species (Poos et al. 2010; DFO 2011; DFO 2012).

The Eastern Sand Darter (ESD) is a small (71 mm), benthic darter and the only member of the genus *Ammocrypta* found in Canada (COSEWIC 2009). It is listed as Threatened by SARA in Canada (SARA 2016). Its distribution in Canada is limited to Big Creek, Big Otter Creek (likely extirpated), Grand River, Lake Erie and its tributaries, Sydenham River, and Thames River (DFO 2011). ESD populations in Grand River were deemed “good” and more than 735 individuals have been captured since 1987 when it was first reported (DFO 2011). Populations in Sydenham River were deemed “poor” with only 43 individuals recorded in the last two decades (DFO 2011) and only 12 caught in a more recent study (Ginson et al. 2015). ESD inhabits lakes and shallow streams or rivers, with sandy bottoms for burrowing (COSEWIC 2009). Main threats to ESD populations are habitat degradation caused by eutrophication and pollution from agricultural practices and urban development, invasive species (including Round Goby), and disease (DFO 2011; Finch et al. 2013).

The Northern Madtom (NMT) is a small (120 mm) benthic, cavity-nesting, catfish listed as Endangered under SARA in Canada (SARA 2016). It prefers big rivers with slightly turbid water and substrates comprised of sand, gravel, and silt for nesting (DFO 2012). The Canadian distribution is

restricted to the Detroit River, Lake St. Clair and its tributaries, St. Clair River, Thames River, and Sydenham River although it has not been collected in the Sydenham River since 1975 (COSEWIC 2012; DFO 2012). NMT is not known to be present in the Grand River either currently or historically. The main threats to NMT include habitat degradation such as nutrient loading and habitat fragmentation, and competition with invasive species such as the Round Goby for food and habitat space (Edwards et al. 2012).

The Silver Shiner (SS) is a small (143 mm) minnow with a Canadian distribution consisting of four known populations, in Bronte Creek, Grand River, Sixteen Mile Creek, and Thames River. It is assessed as Threatened by COSEWIC (COSEWIC 2011) but listed as Special Concern by SARA in Canada (SARA 2016). There are no current or historic populations of SS in the Sydenham River, although a single individual was reportedly caught in the east Sydenham River in 2003 (DFO 2013). A population in Grand River is considered “fair” with only a few individuals collected (DFO 2013). SS prefers fast-flowing systems with deep riffles and sand, clay and gravel substrates (McKee & Parker 1982). The main threat to SS populations is habitat degradation via contamination from poor water management (i.e. nutrient and/or sediment loading), habitat fragmentation, and introduced species (DFO 2013).

The Round Goby (RG) originally arrived to the Great Lakes via ballast water from Eastern Europe, with initial reports in Lake St. Clair in 1990 (Jude et al. 1992). It is a small (average body length is 100 mm (N’Guyen et al. 2016)), highly invasive, benthic species that has spread throughout the Great Lakes and is expanding rapidly upstream in tributaries (Poos et al. 2010; Bronnenhuber et al. 2011). RG burrow in rocky substrate or open crevices to lay eggs during the spawning season from April to September, directly competing with NMT and ESD for benthic habitat space (Kornis 2011; Kornis et al. 2012; Edwards et al. 2012). It feeds on mussels (e.g. dreissenids), insect larvae, and young-of-the-year of native species such as NMT (DFO 2012; Burkett & Jude 2015).

Furthermore, RG feed nocturnally which also directly competes with NMT for foraging space (COSEWIC 2002; DFO 2012). The direct impacts of RG on SS is not currently known (DFO 2013).

Water Sampling

Water sampling was conducted by DFO via boat in September to early October 2013 in the Grand River and in mid to late October 2013 in the Sydenham River using a sampling protocol modified from Jerde et al. (2011). All field equipment was sterilized by soaking in 10% bleach for 10 minutes and rinsed using double-distilled water (ddH₂O). Water samples from 44 sites in the Sydenham River and 43 sites in the Grand River were collected in 500 mL Nalgene bottles and stored in the field in coolers containing ice. Surface samples were collected just under the surface, while samples collected at greater depths were characterized as bottom samples. Blank river site control samples (river controls) were included at a small subset of sample sites (3 from Sydenham River and 4 from Grand River) by opening a 500 mL Nalgene bottle filled with ddH₂O to expose the contents to the air, sealed, and stored in the cooler alongside the river water samples. The majority of samples collected at each site were replicated (2-3 bottom samples and 10-15 surface samples) and we used a subset of 2-5 samples per site for eDNA analysis. In total, we included 184 Sydenham River samples (109 surface, 72 bottom, and 3 river controls) and 170 Grand River samples (108 surface, 58 bottom, and 4 river controls).

eDNA Extraction

Within 24 hours of collection, all water samples were filtered using Whatman® glass microfiber filter papers (47 mm diameter, 1.2 µm pore size; Whatman, Maidstone, UK). Prior to filtering each river sample, 500 mL of ddH₂O was filtered on a separate filter to act as lab controls, followed by the filtration of the river sample on new filters (up to 4 filters for samples with high sediment load) using the same filtration apparatus. This allowed each sample to have its own lab control – we also included lab controls for the 7 river control samples. If any lab control tested

positive for target DNA during PCR amplification, the corresponding sample would be excluded due to potential contamination (i.e. false positives). Each filter was placed in a 15 mL Falcon tube and stored at -20 °C until DNA extraction.

For DNA extraction, filters were cut into halves using sterile forceps and razor blades cleaned between each use using 95% ethanol. One half of the filter was stored and the second half was used immediately for DNA extraction. The half filters were cut into strips to help with digestion and placed into 2 mL screwcap tubes containing 400 µL of 1.0 mm glass beads packed dry (BioSpec Cat. No. 11079110), 400 µL of ddH₂O, 400 µL of phenol-chloroform: isoamyl alcohol (25:24:1), and 400 µL of cetyl trimethylammonium bromide (CTAB) digestion buffer. The tubes were homogenized for two minutes at 3000 strokes per minute using Mini-Beadbeater-24 (Fisher Scientific LTD, BioSpec.) to allow complete cellular breakdown and protein digestion. The samples were centrifuged at 13 000 rpm for 20 minutes, then the supernatant was transferred to a new 1.5 mL Eppendorf tube, and then vortexed with equal volumes of chloroform-isoamyl for a second phase separation. The mixed solution was centrifuged again, supernatant transferred to a new 1.5 mL Eppendorf tube, then mixed with an equal volume of isopropanol and one tenth volume of 3M sodium acetate (pH 5.2). The mixtures were left overnight at

-20 °C. DNA was pelletized by centrifugation, washed once with 70% ice-cold ethanol, and re-suspended in 30 µL of 10 mM TE Buffer and 1.0 µL of 20 mg·µL⁻¹ RNase A to eliminate RNA present in the sample. All extracted eDNA was stored at -20 °C until further analysis.

Primer Design

We designed a group-specific fish CO1 PCR primer set (PS1) developed to amplify our four target species and 114 additional species that potentially inhabit the two Great Lakes tributaries based on DFO catch data and FishMaP (McLaughlin et al. 2010; Meixler et al. 2011; see Supp. Table

S1). The PS1 primers were designed using COI sequences collected from NCBI Genbank and Barcode of Life Database (BOLD; Ratnasingham & Hebert 2007) for 118 fish species and aligned using Geneious v. 6.1 (Biomatters, www.geneious.com). The primers were synthesized with 5' tails: Uni-A (forward) and Uni-B (reverse) for NGS library preparation (Table 1). NGS allowed us to identify both the target and non-target species sequences unambiguously as well as provide a semi-quantitative measure of signal strength (number of reads). We determined the primer-template match for each species from both rivers by DFO to estimate expected PCR amplification efficiency. Highly conserved regions surrounding polymorphic regions were selected for primer design, targeting a fragment size of ~250 bp. Degenerate base sites were avoided within 5 bp of the 3' end to minimize mispriming (Kwok et al. 1990; Epp et al. 2012). Primer-BLAST (Ye et al. 2012) and Net Primer (Premier Biosoft, www.premierbiosoft.com/netprimer) were used to verify primer specificity, low self-complementarity, uniform annealing temperatures, no secondary structure, and appropriate fragment length. PS1 efficiency was sensitivity tested for all four target species using a 10-fold dilution series and SYBR™ Green qRT-PCR with initial DNA template concentrations of 38.4 ng·μL⁻¹ for RG, 35.4 ng·μL⁻¹ for SS, 46.4 ng·μL⁻¹ for ESD, and 30.0 ng·μL⁻¹ for NMT. The qRT-PCR conditions were set to an initial denaturation at 95 °C for 10 minutes, followed by 45 cycles of 95 °C for 15 seconds and 60 °C annealing temperature for 60 seconds. A 10-fold serial dilution for each target species was used to estimate the lowest target DNA concentration that our PS1 primers could detect (Armbruster & Pry 2008; Kim et al. 2014).

eDNA PCR

eDNA PCR reactions consisted of 2.5 μL of 10X *Taq* reaction buffer, 25 mM MgCl₂, 0.5 μM of each forward and reverse PS1 primer, 0.2 mM of each dNTP, 0.2 mg·mL⁻¹ BSA, 0.1 units *Taq* polymerase, 1.0 μL of eDNA sample, and ddH₂O for a total reaction volume of 25 μL. PCR conditions were set to an initial denaturation at 95 °C for 2 minutes, then 40 cycles of 30 seconds at 94 °C, 30 seconds annealing temperature at 52 °C, 30 seconds at 72 °C for extension, final extension at 72 °C

for 10 minutes, and a final hold at 4 °C. Five PCR controls (4 target species benchmark DNA and a no-template blank) were also included for each prepared mastermix. The river control samples (n = 7) were treated as river samples and included in the NGS library preparation described below; however, the lab control samples were DNA extracted and underwent the first round of PCR, as described above, and assessed for amplification by inspection on an agarose gel; they were not included in the NGS library.

Next-Generation Sequencing Preparation

After eDNA PCR amplification, the PS1 PCR products were cleaned using Agencourt AMPure XP (Beckman Coulter, Mississauga, ON, Canada) to remove primer dimer and fragments less than 100 bp. Unique barcodes were ligated to the PCR amplicons using a second, short-cycle PCR for NGS library preparation; library preparation thus required two rounds of PCR. Second round PCRs consisted of 2.5 µL of 10X *Taq* reaction buffer, 25 mM MgCl₂, 0.2 mM of each dNTP, 0.5 µM P1+Uni-B adaptor (5'- CCTCTCTATGGGCAGTCGGTGATacgccaccgagc-3'), 0.5 µM A+barcode+key+Uni-A adaptor (5'- CCATCTCATCCCTGCGTGTCTCCGACTCAGxxxxxxGATacctgcctgccg), 0.1 units *Taq* polymerase, 10 µL of cleaned PCR product, and ddH₂O for a total reaction volume of 25.5 µL. Short-cycle PCR conditions started with a 2-minute denaturation at 95 °C followed by 6 cycles of 95 °C denaturation for 30 seconds, 60 °C annealing temperature for 30 seconds, 72 °C extension for 30 seconds, and a final extension at 72 °C for 5 minutes. Second round PCR products with attached adaptors were purified again using Agencourt AMPure XP, 10 µL of each sample combined in a 1.5 mL tube (72 samples per tube), and precipitated at -20 °C overnight using equal volumes of isopropanol and one tenth volume of 3M sodium acetate (pH 5.2). The combined PCR products (i.e. library) was centrifuged, isopropanol discarded, and the pellet was washed once with 70% ice-cold ethanol. The tubes were centrifuged again, ethanol discarded, and the library was eluted in 50 µL of ddH₂O. Next, 20 µL of the library was gel extracted and bioanalyzed to determine final DNA concentration using Agilent High Sensitivity DNA chip on an Agilent 2100 Bioanalyzer (Agilent

Technologies, Germany). Finally, the library was diluted to a final concentration of 55 pmol·L⁻¹ and sequenced on a 318-chip using the Ion Chef™ System (Life Technologies, USA).

Next-Generation Sequencing Filtering

NGS data was processed using Quantitative Insights into Microbial Ecology (QIIME) software (Caporaso et al. 2010) to remove sequences that were smaller than 200 bp, de-replicate sequences, remove sequences with more than three primer-template mismatches, or that do not meet the default minimum average quality score of 25 (corresponds to an average error rate of 1% in accepted reads). We used the fish species COI sequences previously used for PS1 primer design to create a custom reference database for each river. The filtered NGS sequences were compared against the custom database using BLASTn with default parameters changed from an expected (E) value of 10 to 10⁻⁶⁰ and percent identity from 0% to 96% to ensure high confidence in the returned species identification. We classified a species as “present” at a site only if it returned at least three eDNA sequence matches in a sample. Disregarding one hit (singletons) or two hits (doubletons) avoids diversity overestimation produced by PCR or sequence artefacts, but accepting eDNA matches of three allows rare sequences to be counted; an important consideration for rare species (Zhan et al. 2014).

Community and Species Occurrence Analysis

The eDNA presence data for the target SAR and invasive species were mapped to display their spatial distribution. Specific Aquatic Ecosystem Classification (AEC) values (which define habitats along water systems based on environmental characteristics such as soil type, climate, and geology (Melles et al. 2013) were included to show how species occurrence patterns may be associated with habitat type along the two rivers.

We also compared the eDNA presence data with the data from capture-based methods used by DFO in 2010 to 2013 during June, July and August for the Grand River, and from 2010 to 2012

during June, August, and September for the Sydenham River. The species occurrence results from the two sampling methods were combined to determine the percent of species detected using eDNA methods.

Whole-river community analysis of species co-occurrences was completed using R Studio v. 0.99.892 (RStudio Team 2015) with the package “cooccur” (Veech 2013; Griffith et al. 2016) to determine positive, negative, and random (i.e. no significance) spatial co-occurrence distribution relationships between species, using the default settings of the package. We also kept the default threshold (thresh = TRUE), which removed any species pairs that do not share any sites; this removes species pairs that have insufficient co-occurrence information (Veech 2013). We performed species co-occurrence analyses using the DFO capture-based data and our eDNA-based data separately to determine if the same patterns of species co-occurrences were identified. While we report all detected species co-occurrences, we focus on our target species. We performed the co-occurrence analyses for both sampled rivers at sites that were sampled using capture-based methods and e-DNA methods.

Semi-Quantitative eDNA Analysis

Additionally, eDNA detections for our four target species were visualized semi-quantitatively as the number of eDNA sequences for a target species at a specific site divided by the total number of eDNA sequences returned at the same site. This method illustrates the abundance of target species signal at a given site based on its concentration of eDNA sequences out of all eDNA sequences obtained at the site as a proportion. If a species occurs in high abundance in a particular area, samples collected in those areas should reflect a higher proportion of eDNA sequence returns for the species as a measure of relative signal strength.

Results

Primer Design

Our group-specific primer PS1 was designed to have minimum mismatch for all species found in the Grand River and the Sydenham River. Figure 1 shows the level of forward and reverse primer mismatch, with 0 total mismatch for majority (68.3%) of the species, 24.4% had a mismatch in only one of the primers, and 8.5% mismatched in both primers. Only 9.8% of fishes had a mismatch within 5 bp of the 3' end for either the forward or reverse primer, but there were no mismatches within 5 bp of the 3' end for both primers for our four target species. PS1 amplified DNA for all four of our target species across the dilution series qRT-PCR test. Detection sensitivity plateaued at approximately 10^{-5} template dilution, which reflects the PS1 primer approximate detection limit (Figure 2).

Next-Generation Sequence Data

Our next-generation sequencing produced 6.5 million raw reads. A total of 2 430 736 reads were returned for 354 (170 Grand River and 184 Sydenham River) samples with an average of 7107 sequences per sample (range = 1 to 50 607 reads/sample). After BLASTing against the custom fish COI sequence database, one out of seven river controls (site S20) was contaminated and returned a total of 5 sequence reads, all were RG; however, the corresponding lab control was negative for fish DNA (no PCR amplification). We thus concluded that field contamination was low, although we did remove the one site that returned 5 RG sequences from further analyses. None of the 354 lab control samples produced visible PCR amplification using our group-specific primers (PS1), indicative of low lab-based contamination level.

After removing species or samples that returned < 3 eDNA sequences from further analysis (and removing sequences that did not match any of the fish sequences in our custom database), 954 811 fish sequences were returned: 75.5% matched RG (721 267 sequences), 0.28% matched SS (2 720 sequences), 0.22% matched ESD (2 079 sequences) and 0.18% matched NMT (1 683 sequences)

(Table 2). The remaining 23.8% sequence reads (227 062 sequences) matched 61 co-occurring native species (Supp. Table S2A and Supp. Table S2B). Grand River had a total of 344 725 returned eDNA sequence reads from 43 sites comprising 50 fish species from 50 bottom and 90 surface samples (8 bottom and 18 surface samples removed). Sydenham River resulted in a total of 610 087 eDNA sequence reads from 43 sites comprising 54 fish species from 65 bottom and 101 surface samples (7 bottom and 8 surface samples removed).

RG was detected in 277 samples in total from both rivers (113 Grand River and 164 Sydenham River samples); this corresponds to the majority of sampled sites. The number of RG eDNA sequence reads per site varied widely, with a mean of 1 842 eDNA sequences per sample in the Grand River (range = 3 to 13 729 reads/sample) and a mean of 3 128 sequences per sample in the Sydenham River (range = 4 to 25 734 reads/sample). RG was not found at one Sydenham River site (S28) and five Grand River sites (G15, G17, G18, G19, and G24; Fig 3). ESD were detected in four Grand River sites (5 samples total) with a mean of 192 sequences per sample (range = 59 to 485 reads/sample), and four Sydenham River sites (4 samples total) with a mean of 280 sequences per sample (range = 184 to 489 reads/sample). SS was only detected in Grand River at three sites (G7, G10, and G40) with a mean of 701 eDNA sequences (range = 10 to 1 044 reads/site). NMT was detected in one Grand River site (G6) in a surface sample with 48 eDNA sequence matches. NMT was also detected at seven sites in Sydenham River (S4, S8, S14, S15, S21, S30, and S41) with a mean of 281 sequences (range = 3 to 1 160 reads/site) (Fig. 3). Finally, ESD had higher sequence reads in surface samples than bottom samples, NMT had greater eDNA read count in surface samples, and SS had greater eDNA read count in bottom samples (Table 3).

Community and Species Co-Occurrence

All Grand River sample sites were dominated by cool, turbid water with slow flow (AEC = MTS (moderate, turbid, slow)). Sydenham River sample sites were also dominated by cool, turbid, slow water (AEC = MTS) in lower and upper East Sydenham River segments, and warm, turbid, slow-flowing water in the North Sydenham River segment (AEC = LTS (low, turbid, slow); Fig. 3).

Overall, the eDNA method detected 82.6% and 62.5% of the species that capture-based methods detected in Grand River and the Sydenham River, respectively (Table 4). For all species detected using both methods, eDNA detected 86.2% and 72.0% of the total in the Grand River and the Sydenham River, respectively (Table 4).

Species co-occurrence was analyzed using presence data from capture-based and eDNA methods, producing results for positively co-occurring species pairs (occur together more often than expected, $P < 0.05$), negatively co-occurring species pairs (occur together less often than expected, $P < 0.05$), and randomly co-occurring species pairs (occurrences not different from that expected by chance, $P > 0.05$). Using presence/absence data from capture-based methods in the Grand River, our species co-occurrence analysis included 47 species across 29 sites (that were present in the eDNA sampling), producing 1 081 species pair combinations of which 710 (65.7%) were removed. Species pairs that were removed due to insufficient co-occurrence data. Co-occurrence analysis for the remaining 371 species pairs resulted in 13 positive, 11 negative, and 347 random species co-occurrences (Fig.4A, Supp. Table S3A) in the Grand River, based on DFO capture-based data. In the Sydenham River, 56 species across 17 sites were analyzed producing 1 540 species pairs of which 959 (62.3%) pairs were removed and 581 pairs were analyzed for species co-occurrence. There were 49 positive, 7 negative, and 492 random species pairs (Fig.4A, Supp. Table S3B).

For eDNA-based detection in the Grand River, a total of 1 225 species pairs were analyzed for 50 detected species across 43 sites. A total of 947 (77.3%) species pairs were removed and 278 species pairs were analyzed further. There were 42 positive, 0 negative, and 236 random co-occurrences (Fig. 4B, Supp. Table S3C). In the Sydenham River, 1 431 species pairs were analyzed for 54 species across 43 sites, in which 1 137 (79.5%) were removed, and 294 species pairs remained for further analysis. There were 65 positive, 0 negative, and 229 random co-occurrences (Fig.4B, Supp. Table S3D). For NMT, ESD, and SS there were no significant negative or positive co-occurrences. RG had one positive co-occurrence with Common Carp (*Cyprinus carpio*) in the Grand River.

Semi-Quantitative eDNA Data

Figure 5 illustrates the proportion of eDNA NGS reads for each target species relative to the total number of fish-assigned NGS reads arranged from upstream to downstream sites. The three target SAR species had very low eDNA detection proportions (< 10%; Fig. 5A, Supp. Table S2B) across the majority of the sites where the species was detected, except for ESD in site G01 (30.9%) and NMT in site S21 (30.3%). In site G01, a total of 375 eDNA sequences were collected, 116 belonging to ESD and the remaining identified as RG and Brindled Madtom (*Noturus miurus*) (162 and 87 sequences respectively). In Site S21, 3 712 eDNA sequences were obtained and 1 124 belonged to NMT and the remaining identified as Golden Redhorse (*Moxostoma erythrurum*) and RG (11 and 2 577 sequences respectively). Downstream eDNA percent detection of RG in the Grand River reflects an oscillating pattern, whereas detection proportions in the Sydenham River increase when moving downstream (Fig. 5B, Supp. Table S2B).

Discussion

Detection of Species at Risk

By using a PCR primer set designed to amplify a broad range of fish species, followed by metabarcoding using a custom sequence-database, we not only detected our species of interest, but we also gathered distributional data for the fish community that allowed species co-occurrence analysis. Across both watersheds, we detected 65 of 82 fishes known to inhabit the two rivers using eDNA and NGS analysis; these detections included all four of our target SAR and invasive species. We show that our combined eDNA/NGS approach with community-specific primers was able to provide information on the distribution of invasive, native, and rare species in flowing freshwater systems, and allowed us to make comparisons of habitat use and potential species interactions without the need to physically capture individuals.

ESD was detected mainly in the lower East Sydenham River using capture-based methods, but we detected high eDNA sequence read numbers (> 100) farther upstream in the upper East

Sydenham River and in North Sydenham River. The eDNA-based distribution may be indicative of movement of individuals upstream or perhaps the discovery of new populations; however, it is not consistent with the hydrological flow of the eDNA signal and is thus unlikely an artifact. Capture-based methods detected ESD in several sites throughout Grand River whereas eDNA detections were isolated to the upper Grand River. Lack of ESD eDNA detections downstream could also indicate that there are fewer individuals located downstream, resulting in weaker eDNA concentrations in downstream samples. Fewer individuals in a larger river system can also lead to rapid dilution of target eDNA. Since eDNA has the potential for downstream flow in lotic systems, eDNA signals can reflect DNA sources upstream from the site of detection (Thomsen & Willerslev 2015); however, there is no mechanism for downstream eDNA sources to be detected upstream. SS was detected only in Grand River using eDNA, and this distribution agrees with DFO capture data (DFO 2013).

NMT was detected in Sydenham River and returned 48 eDNA sequence matches at one Grand River site (G6). In both cases, the eDNA detections do not agree with DFO capture records and historical reports of NMT presence. Although NMT have not been reported in the Sydenham River since 1975 (Edwards et al. 2012), the high number of eDNA sequence matches coupled with detection in at least four independent sites (Fig. 5A) suggests that there is a source of NMT eDNA in Sydenham River, perhaps missed by capture-based methods due to low abundance and cryptic behaviour (COSEWIC 2002). Furthermore, our eDNA sequences had a 96%-99% identity match to NMT COI sequences in our custom database, and 90-92% sequence match for Stonecat (*Noturus flavus*) or Brindled Madtom (*Noturus miurus*), the most closely related species that co-occurs with our target species. An alternative possibility is that we detected residual eDNA dating to their known presence in the 1970's; however, such a long residence time for residual eDNA is highly unlikely (Thomsen et al. 2012; Balasingham et al. 2016), especially as we detected NMT eDNA in surface water samples. While the detection of NMT at one Grand River site was based on 48 recovered sequence reads, the fact that it was at a single upstream site indicates it is, at best, a signal of a very

few, isolated individuals. Although eDNA sample contamination is a possibility, lab controls for this region of the Grand River did not produce amplification for any fish species (although there was no river control taken at this specific site). Previous work has shown that eDNA has a higher detection sensitivity, especially for rare or cryptic species, than capture-based methods, and has extended the known range of a variety of species, including invertebrates (Tréguier et al. 2014), fishes (Janosik & Johnston 2015; Sigsgaard et al. 2015; Boothroyd et al. 2016) and amphibians (Pierson et al. 2016). For example, Sigsgaard et al. (2015) detected the European Weather Loach (*Misgurnus fossilis*) using eDNA at a site where it had not been detected by capture-based methods for 17 years. Pierson et al. (2016) found that eDNA had 20X the detection probability for Patch-Nosed Salamander (*Urspeleperes brucei*) than traditional leaf litterbag surveys. Janosik and Johnston (2015) showed that traditional seining and netting identified the rare Slackwater Darter (*Etheostoma boschungii*) in one out of 49 sampled sites whereas eDNA identified it in 23 of the sites. Boothroyd et al. (2016) detected Spotted Gar (*Lepisosteus oculatus*) eDNA in Thames River where it had previously been thought to be extirpated for over 50 years. Hence, while our eDNA detection of SAR did not directly correspond to capture detections, the differences are likely attributable to the higher detection sensitivity of eDNA methods and the potential for target species movement between the time of the capture surveys and the water sampling. Overall, our results highlight the potential for eDNA to be used alongside capture-based methods, or as a stand-alone detection methodology for monitoring and mapping target SAR.

Detection of Invasive Species

The spread of aggressive invasive species such as the RG can cause detrimental changes to local habitat and negatively impact native biodiversity via competition and direct predation (Thompson & Simon 2014; Burkett & Jude 2015). Not surprisingly, a large portion (76.1%) of our eDNA sequence reads matched RG DNA, with detections in 93% of all sampled sites across both rivers. The RG is a known aggressive invader, with high-density capture-based detections in both the

Sydenham River and Grand River since its secondary invasion of Great Lakes tributaries (Poos et al. 2010). eDNA analysis detected RG in nine Grand River and five Sydenham River sites where no individuals were captured using capture-based methods. The new Grand River RG sites were upstream of past RG captures sites, indicating that it had likely moved upstream from the time that capture-based methods were used to when eDNA sampling occurred, or possibly missed when in low abundance at the invasion front. The new Sydenham River sites were also upstream from past capture records, however two extra sites with RG eDNA detections were downstream of past capture data, hence downstream migration of RG eDNA may have contributed to those detections.

Interestingly, sites in upper Grand and East Sydenham rivers had an overall higher number of detected fish species (≥ 10 species) including the three target SAR (Fig. 3), indicating that those areas may be localized hotspots for fish biodiversity. The use of appropriate, non-invasive management of RG to limit its numbers while maintaining the relatively higher biodiversity in the upper reaches of these two large Great Lakes tributaries should be considered. For example, RG pheromone traps designed to attract only RG would be ideal (Kornis et al. 2012; Smith 2014). Curiously, the few sites that did not produce positive RG eDNA detections had been sites where RG were captured previously. Thus, RG was present at those sites and the failure of our eDNA analysis to detect them is likely due to lower eDNA quality/quantity (or PCR inhibition) in those samples. This is supported by the overall low number of eDNA sequence reads for all other species at those sites (range = 1 to 136 reads/species; Supp. Table S2A). The impact of false negatives using eDNA methods can be high when monitoring the spread of invasive species. However, in this study, eDNA had higher detection rates for the invasive RG relative to capture-based methods, and eDNA extended the RG range farther upstream in both rivers, as would be expected given that our sampling ranged from 1-3 years after the capture-based surveys, indicating potential movement of the invasion front upstream with few individuals contributing eDNA. At the sites where capture-based methods did not detect RG, RG were captured at neighbouring sites in low numbers. Our eDNA detection showed RG in all of those

sites, demonstrating that false negative detections occur using both capture-based and eDNA detection methods. Thus, the ideal approach to fish community characterization is a combination of both eDNA and capture-based methods, especially in critical early invasion monitoring.

Species Co-Occurrence and Community Composition

Co-occurrence data help to explain patterns in target species spatial distribution and potentially allows the identification of critical fish community interactions for SAR and invasive species. Capture-based methods identified statistically significant positive and negative co-occurrences in both rivers, albeit many fewer species pairs were detected than with the eDNA methods (Fig. 4A). For example, using the capture presence data, Spotfin Shiner (*Cyprinella spiloptera*) negatively co-occurred with Channel Catfish (*Ictalurus punctatus*) in the Grand River, but showed no evidence of negative co-occurrence in the Sydenham River (Fig. 5A, Supp. Table S2C). Hence, while both species were found in these rivers, they were not detected together at individual sites. In the Sydenham River, both Channel Catfish and Spotfin Shiner were caught in sites in the lower and East Sydenham River where the water was warm, turbid, and slow moving (AEC = LTS), but the habitat utilized by these two species in the Grand River was of type MTS (cool, turbid, and slow flowing). The cooler water in the Grand River may be driving habitat differentiation between the Channel Catfish and Spotfin Shiner, as the Channel Catfish was found mostly in downstream sites where some tributaries have warmer water. Hence, despite habitat homogeneity being fairly high in both the Grand River and the Sydenham River, the few negative co-occurrences may represent highly divergent resource or habitat use between species pairs, potential resource competition, or differences in dispersal patterns.

The majority of species co-occurrences detected using eDNA were as expected based on our knowledge of fish species interactions. Overall, our eDNA data identified statistically significant positive species co-occurrences in both rivers (Fig. 4B). Only the RG eDNA data showed one positive

co-occurrence with Common Carp in the Grand River, a species that is also fairly common throughout the river and was also observed to feed on dreissenids, the diet of RG (Ruetz III et al. 2012).

In the Grand River, all samples were collected at sites with high groundwater flow and fast channel flow with cool, turbid water (AEC = MTS), highlighting the relatively homogenous riverine habitat throughout our study area. This likely explains the lack of statistically significant negative co-occurrences, since the habitat type was similar throughout the study area, increasing the likelihood that most significant co-occurrences are based on species preferring similar habitat resources. In the North Sydenham River segment, the study area was dominated by low groundwater flow and low channel flow with warm, turbid waters, whereas the lower and upper East Sydenham River segments were dominated by MTS stream types. The general lack of significant co-occurrences with our target SAR is likely due to their low sample size which limits the ability to detect significant spatial patterns with other co-occurring species. Similarly, no significant co-occurrences were found with our target invasive species as the RG was detected in virtually every sample site in both rivers and, thus, does not display any spatial limitations, or strong ecological interactions relative to other co-occurring species. This also highlights the broad range of habitats that are exploitable by the highly invasive RG. The greater number of significant negative species co-occurrences obtained using capture-based presence data is likely due to higher detection thresholds for capture-based techniques such that low-density species are missed (Shaw et al. 2016). Conversely, eDNA can settle, degrade, or migrate downstream which may result in biased species co-occurrence patterns, and the variation in eDNA detection rates among species due to varying rates of DNA release into the habitat may also affect fine-scale eDNA species distribution patterns (Davy et al. 2015). Additionally, the use of a group-specific primer may result in primer bias and poor amplification of low abundance species in certain sites. However, NGS read count in meta-barcoding studies is generally found to correspond linearly with target species DNA concentration. For example, Pereira et al. (2017) tested a universal primer targeting *Legionella pneumophila* V4 and V5 16S rRNA regions in spiked drinking

and nuclease-free water in 10-fold dilutions for comparative NGS analysis. The authors found NGS reads of eDNA mock communities (a mixture of six different *Legionella* sp. Including the spiked *L. pneumophila*) was quantitative of *L. pneumophila* abundance for each dilution with a strong correlation ($R^2 = 0.88$ for nuclease-free water and $R^2 = 0.89$ for drinking water). Diaz-Real et al. (2015) also used NGS and COI degenerate primers to measure abundance of two bird feather mites (*Proctophyllodes clavatus* and *P. sylviae*) in different abundance ratios and found that NGS read count increased linearly with individual species abundance in mock communities. Nevertheless, we suggest that future eDNA studies should spike field-collected eDNA samples with all target species' DNA and compare the PCR efficiency against replicates spiked with individual target species' DNA to better visualize primer bias and amplification efficiency. We also suggest future studies incorporate multiple marker loci, as this approach has been shown to increase species detection sensitivity through reducing primer bias (Wilcox et al. 2013; Miya et al. 2015; Evans et al. 2016; Shaw et al. 2016).

eDNA Signal Strength

NGS data provides a semi-quantitative metric for species abundance, thus we visualized our eDNA presence data for each species as semi-quantitative proportions of the total number of eDNA reads produced per site for each species (Fig. 5, Supp. Table S2B). This eDNA sequence read proportion metric provides a semi-quantitative measure of signal strength and can, thus, be used as a proxy for detection confidence (stronger signal indicates greater confidence) or, more indirectly, species abundance (Amend et al. 2010; Carew et al. 2013; Elbrecht & Leese 2015). For example, RG, which generally produced higher eDNA reads per site than other species, had similar eDNA proportions in neighbouring sites in both rivers, indicating that the eDNA semi-quantitative abundance data are consistent with the expected relationship between species abundance and favourable habitats. For the three SAR, eDNA proportions were generally less than 10%, which is expected as they are at very low densities and would contribute to little the eDNA in the water

collected samples. In the cases where SAR eDNA proportions were high (approximately 30%), those samples only included amplified DNA from three species, thus providing a greater chance for a high proportion of rare eDNA.

Conclusion

Overall, our eDNA-based species-detection approach identified at least 70% of the known species present in both the Grand and Sydenham rivers (86.4% and 72.4% respectively). Conservation and management programs require detailed information on the spatial distribution of SAR to facilitate the preservation of critical habitat or to develop spatially explicit management plans. In the past, the distribution of fish species was determined using the physical capture and identification of individuals, but those methods are logistically difficult, expensive, unintentionally harmful to target and co-occurring species, harmful to the ecosystem, and may exhibit low detection sensitivity for species that are rare, cryptic, or inhabit difficult to sample areas. Our study demonstrates that a one-time sampling effort for water from two very large Great Lakes tributaries can successfully detect the majority of the contemporary fish community, including SAR, which were detected in several sites not identified using capture-based methods. Moreover, our eDNA analysis also detected the invasive RG at sites not identified in the DFO capture-based fish inventory program. eDNA analysis coupled with NGS also provides the opportunity to passively assess community structure and, thus, determine important species co-occurrence or interactions, since eDNA samples contain template DNA from all species inhabiting the same environment. Community assessment identifies key species interactions and overall species distribution across sites. While eDNA community characterization has limitations, such data can be used to screen for critical habitat (e.g. biodiversity hotspots) and can, thus, focus management assessment in ways not possible based on presence data for individual target species. As eDNA methods continue to be refined, future studies should employ eDNA analysis and NGS together with capture-based methods, for large-scale monitoring of whole community structure and important interspecific relationships with critical target species.

Acknowledgments

THIS PROJECT WAS FUNDED BY THE NATURAL SCIENCES AND ENGINEERING RESEARCH COUNCIL (NSERC) OF CANADA, CANADIAN AQUATIC INVASIVE SPECIES NETWORK (CAISN), THE DEPARTMENT OF FISHERIES AND OCEANS (DFO) SPECIES AT RISK PROGRAMME, A NATURAL SCIENCES AND ENGINEERING RESEARCH COUNCIL (NSERC) DISCOVERY GRANT TO DDH, AND A SPECIES AT RISK RESEARCH FUND (SARRFO) GRANT TO RPW, DDH, AND NEM. WE THANK ROBIN GÁSPÁRDY, KELLY A. MCNICHOLS-O'ROURKE, JASON BARNUCZ, AND ALEX PRICE FROM DFO FOR COLLECTION OF WATER SAMPLES, EQUIPMENT, AND CATCH DATA. LABORATORY SUPPORT WAS PROVIDED BY SARA JAMIESON, STACEY MCINTYRE, DILRAJ SANGHERA, SAMEER JAFAR, KETA PATEL, MARIAM SAMEEM, JUSTIN MYCHEK-LONDER, AND CALVIN KELLEDOCK. THANK YOU TO DR. SUBBA RAO CHAGANTI, DR. XIAOPING HE, DR. KYLE WELLBAND, RYAN SCOTT, AND RUSSELL HEPBURN FROM THE UNIVERSITY OF WINDSOR, FOR AID IN PROTOCOL DESIGN AND ON-GOING DISCUSSION OF THIS PROJECT.

REFERENCES

- Amend AS, Seifert KA, Bruns TD (2010) Quantifying microbial communities with 454 pyrosequencing: does sequence abundance count? *Molecular Ecology*, **19**, 5555-5565.
- Anderson LWJ (2005) California's reaction to *Caulerpa taxifolia*: a model for invasive species rapid response. *Biological Invasions*, **7**, 1003-1016.
- Armbruster DA, Pry T (2008) Limit of blank, limit of detection and limit of quantification. *The Clinical Biochemist Reviews*, **29**, S49-S52.
- Balasingham KD, Walter RP, Heath DD (2016) Residual eDNA detection sensitivity assessed by quantitative real-time PCR in a river ecosystem. *Molecular Ecology Resources*, doi: 10.1111/1755-0998.12598.
- Barnes MA, Turner CR (2016) The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, **17**, 1-17.
- Bohmann K, Evans A, Gilbert MTP *et al.* (2014) Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution*, **29**, 358-367.
- Boothroyd M, Mandrak NE, Fox M, Wilson CC (2016) Environmental DNA (eDNA) detection and habitat occupancy of threatened spotted gar (*Lepisosteus oculatus*). *Aquatic Conservation: Marine and Freshwater Ecosystems*, **26**, 1107-1119.
- Bronnenhuber JE, Dufour BA, Higgs DM, Heath DD. (2011). Dispersal strategies, secondary range expansion and invasion genetics of the nonindigenous Round Goby, *Neogobius melanostomus*, in Great Lakes tributaries. *Molecular Ecology*, **20**, 1845-1859.
- Burkett EM, Jude DJ (2015) Long-term impacts of invasive Round Goby *Neogobius melanostomus* on fish community diversity and diets in the St. Clair River, Michigan. *Journal of Great Lakes Research*, **41**, 862-872.
- Carew ME, Pettigrove VJ, Metzeling L, Hoffmann AA (2013) Environmental monitoring using next generation sequencing: rapid identification of macroinvertebrate bioindicator species. *Frontiers in Zoology*, **10**, doi: 10.1186/1742-9994-10-45.
- Caporaso JG, Kuczynski J, Stombaugh J *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**(5): 335-336.
- COSEWIC (2002) COSEWIC assessment and update status report on the Northern Madtom *Noturus stigmosus* in Canada.

- Committee on the Status of Endangered Wildlife in Canada. Ottawa. vii + 15 pp. (www.sararegistry.gc.ca/status/status_e.cfm)
- COSEWIC (2009) COSEWIC assessment and status report on the Eastern Sand Darter *Ammocrypta pellucida*, Ontario populations and Quebec populations, in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa. vii + 49 pp. (www.sararegistry.gc.ca/status/status_e.cfm)
- COSEWIC (2011) COSEWIC assessment and status report on the Silver Shiner *Notropis photogenis* in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa. Xi + 45 pp. (www.sararegistry.gc.ca/status/status_e.cfm)
- COSEWIC (2012) COSEWIC Status Report on Northern Madtom (*Noturus stigmosus*) in Canada. Prepared for Committee on the Status of Endangered Wildlife in Canada. 2-month Interim Report. 48 pp.
- Griffith DM, Veech JA, Marsh CJ (2016) cooccur: Probabilistic Species Co-Occurrence Analysis in R. *Journal of Statistical Software*, **69**, 1-17.
- Davy CM, Kidd AG, Wilson CC (2015) Development and validation of environmental DNA (eDNA) markers for detection of freshwater turtles. *PLoS ONE*, **10**: e0130965
- Deiner K, Altermatt F (2014) Transport Distance of Invertebrate Environmental DNA in a Natural River. *PLoS ONE*, **9**, e88786.
- Dejean T, Valentini A, Miquel C *et al.* (2012) Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog *Lithobates catesbeianus*. *Journal of Applied Ecology*, **49**, 953-959.
- Diaz-Real J, Serrano D, Piriz A, Jovani R (2015) NGS metabarcoding proves successful for quantitative assessment of symbiont abundance: the case of feather mites on birds. *Experimental and Applied Acarology*, **67**, 209-218.
- DFO (2011) Recovery Potential Assessment of Eastern Sand Darter (*Ammocrypta pellucida*) in Canada. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2011/020.
- DFO (2012) Recovery Potential Assessment of Northern Madtom (*Noturus stigmosus*) in Canada. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2012/051.
- DFO (2013) Recovery Potential Assessment of Silver Shiner (*Notropis photogenis*) in Canada. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2012/068.
- Doi H, Takahara T, Minamoto T *et al.* (2015) Droplet Digital Polymerase Chain Reaction (PCR) Outperforms Real-Time PCR in the Detection of Environmental DNA from an Invasive Fish Species. *Environmental Science Technology*, **49**, 5601-5608.
- Edwards AL, Laurin AY, Staton SK (2012) Recovery Strategy for the Northern Madtom (*Noturus stigmosus*) in Canada. Species at Risk Act Recovery Strategy Series. Fisheries and Oceans Canada, Ottawa. viii +42 pp.
- Eichmiller JJ, Miller LM, Sorensen PW (2016) Optimizing techniques to capture and extract environmental DNA for detection and quantification of fish. *Molecular Ecology Resources*, **16**, 56-68.
- Elbrecht V, Leese F (2015) Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass – Sequence Relationships with an Innovative Metabarcoding Protocol. *PLoS ONE*, **10**, e0130324.
- Epp LS, Boessenkool S, Bellemain EP *et al.* (2012) New environmental metabarcodes for analyzing soil DNA: potential for studying past and present ecosystems. *Molecular Ecology*, **21**, 1821-1833.
- Evans NT, Olds BP, Renshaw MA *et al.* (2016) Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Molecular Ecology Resources*, **16**, 29-41.
- Ficetola GF, Miaud C, Pompanon F, Taberlet P (2008) Species detection using environmental DNA from water samples. *Biology Letters*, **4**, 423-425.
- Finch M, Faber JE, Koops MA (2013) Biological traits of eastern sand darter (*Ammocrypta pellucida*) in the lower Thames River, Canada, with comparisons to a more southern population. *Ecology of Freshwater Fish*, **22**, 234-245.
- Fukumoto S, Ushimaru A, Minamoto T (2015) A basin-scale application of environmental DNA assessment for rare endemic species and closely related exotic species in rivers: a case study of giant salamanders in Japan. *Journal of Applied Ecology*, **52**, 358-365.
- Ginson R, Walter RP, Mandrak NE, Beneteau CL, Heath DD (2015) Hierarchical analysis of genetic structure in the habitat-specialist Eastern Sand Darter (*Ammocrypta pellucida*). *Ecology & Evolution*, **5**, 695-708.
- Goldberg CS, Sepulveda A, Ray A, Baumgardt J, Waits LP (2013) Environmental DNA as a new method for early detection of New Zealand mudsnails (*Potamopyrgus antipodarum*). *Freshwater Science*, **32**, 792-800.
- Hulme PE (2006) Beyond control: wider implications for the management of biological invasions. *Journal of Applied Ecology*, **43**, 835-847.
- Janosik AM, Johnston CE (2015) Environmental DNA as an effective tool for detection of imperiled fishes. *Environmental Biology of Fishes*, **98**, 1889-1893.
- Jerde CL, Mahon AR, Chadderton WL, Lodge DM (2011) "Sight-unseen" detection of rare aquatic species using environmental DNA. *Conservation Letters*, **4**, 150-157.
- Jerde CL, Mahon AR (2015) Improving confidence in environmental DNA species detection. *Molecular Ecology Resources*, **15**, 461-463.
- Jude DJ, Reider RH, Smith GR (1992) Establishment of Gobiidae in the Great Lakes basin. *Canadian Journal of Fisheries and Aquatic Sciences*, **49**, 416-421.
- Kim TG, Jeong S-Y, Cho K-S (2014) Comparison of droplet digital PCR and quantitative real-time PCR for examining population dynamics of bacteria in soil. *Applied Microbiology and Biotechnology*, **98**, 6105-6113.
- Klymus KE, Richter CA, Chapman DC, Paukert C (2015) Quantification of eDNA shedding rates from invasive bighead carp (*Hypophthalmichthys nobilis*) and silver carp (*Hypophthalmichthys molitrix*). *Biological Conservation*, **183**, 77-84.

- Kornis MS (2011) Distribution, impact, and life histories of round gobies in the Laurentian Great Lakes and their tributaries: lessons for invasion biology. (Doctoral dissertation). Retrieved from The University of Wisconsin-Madison Limnology Library.
- Kornis MS, Mercado-Silva N, Vander Zanden MJ (2012). Twenty years of invasion: a review of Round Goby (*Neogobius melanostomus*) biology, spread and ecological implications. *Journal of Fish Biology*, **80**, 235-285.
- Kwok S, Kellogg DE, McKinney N *et al.* (1990) Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus type 1 model studies. *Nucleic Acids Research*, **18**, 999-1005.
- Laramie MB, Pilliod DS, Goldberg CS (2015) Characterizing the distribution of an endangered salmonid using environmental DNA analysis. *Biological Conservation*, **183**, 29–37.
- Li Z, Huang G, Wang X, Han J, Fan Y (2016) Impacts of future climate change on river discharge based on hydrological inference: A case study of the Grand River Watershed in Ontario, Canada. *Science of the Total Environment*, **548-549**, 198-210.
- Mahon AR, Jerde CL, Galaska M *et al.* (2013) Validation of eDNA surveillance sensitivity for the detection of Asian carps in controlled and field experiments. *PLoS ONE*, **8**, e58316.
- Mandrak NE, Crossman EJ (1992) A checklist of Ontario freshwater fishes. Royal Ontario Museum, Toronto, 176 pp.
- Marson D, Mandrak NE (2009) Survey of the fish assemblages in the non-wadeable waters of the Sydenham River in 2003. *Canadian Manuscript Reports of Fisheries and Aquatic Sciences*, **2916**: v + 21 p.
- Maruyama A, Nakamura K, Yamanaka H, Kondoh M, Minamoto T (2014) The release rate of environmental DNA from juvenile and adult fish. *PLoS ONE*, **9**, e114639.
- Matheson CD, Gurney C, Esau N, Lehto R (2010) Assessing PCR inhibition from humic substances. *The Open Enzyme Inhibition Journal*, **3**, 38-45.
- McKee PM, Parker BJ (1982) The distribution, biology, and status of the fishes *Campostoma anomalum*, *Clinostomus elongates*, *Notropis photogenis* (Cyprinidae), and *Fundulus notatus* (Cyprinodontidae) in Canada. *Canadian Journal of Zoology*, **60**, 1347-1358.
- McKee AM, Spear SF, Pierson TW (2015) The effect of dilution and the use of a post-extraction nucleic acid purification column on the accuracy, precision, and inhibition of environmental DNA samples. *Biological Conservation* **183**: 70-76.
- McLaughlin RL, Jones ML, Mandrak NE, Stacey D (2010) FishMaP On-Line: A Web Application Supporting Science-Based Decisions Concerning Fish Movement and Passage. Great Lakes Fishery Commission.
- Melles S, Jones N, Schmidt B. 2013. Aquatic Research Series 2013-05: Aquatic Ecosystem Classification for Ontario: a technical proposal. Ontario Ministry of Natural Resources. 52 pp.
- Metcalfe-Smith JL, Di Maio J, Staton SK, De Solla SR (2003) Status of the freshwater mussel communities of the Sydenham River, Ontario, Canada. *The American Midland Naturalist*, **150**, 37-50.
- Meixler MS, Mandrak NE, McLaughlin RL (2011) Web based watershed atlas for fishes of the Great Lakes. Great Lakes Fishery Commission.
- Miya M, Sato Y, Fukunaga T *et al.* (2015) MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection of more than 230 subtropical marine species. *Royal Society Open Science*, **2**, 150088.
- N’Guyen A, Hirsch PE, Adrian-Kalchauer I, Burkhardt-Holm P (2016) Improving invasive species management by integrating priorities and contributions of scientists and decision makers. *Ambio*, **45**, 280-289.
- Pereira RPA, Peplies J, Brettar I, Höfle MG (2017) Development of a genus-specific next generation sequencing approach for sensitive and quantitative determination of the *Legionella* microbiome in freshwater systems. *BMC Microbiology*, **17**, 1-14.
- Pierson TW, McKee AM, Spear SF *et al.* (2016) Detection of an enigmatic plethodontid salamander using environmental DNA. *Copeia*, **104**, 78-82.
- Pilliod DS, Goldberg CS, Arkle RS, Waits LP (2014) Factors influencing detection of eDNA from a stream-dwelling amphibian. *Molecular Ecology Resources*, **14**, 109–116.
- Pitt AL, Shinskie JL, Tavano JJ, *et al.* (2017). Decline of a giant salamander assessed with historical records, environmental DNA and multi-scale habitat data. *Freshwater Biology*, **62**, 967-976.
- Pompanon F, Coissac E, Taberlet P (2011). Metabarcoding, une nouvelle façon d’analyser la biodiversité/Metabarcoding a new way to analyze biodiversity. *Biofutur*, **319**, 30-21.
- Poos M, Dextrase AJ, Schwalb AN, Ackerman JD (2010) Secondary invasion of the Round Goby into high diversity Great Lakes tributaries and species at risk hotspots: potential new concerns for endangered freshwater species. *Biological Invasions*, **12**, 1269-1284.
- Port JA, O’Donnell JL, Romero-Maraccini OC *et al.* (2016) Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Molecular Ecology*, **25**, 527-541.
- Ratnasingham S, Hebert PDL (2001) BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Molecular Ecology Notes*, **7**, 355-364.
- Rees HC, Maddison BC, Middleditch DJ *et al.* (2014) The detection of aquatic animal species using environmental DNA – a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology*, **51**, 1450-1459.
- Renshaw MA, Olds BP, Jerde CL, McVeigh MM, Lodge DM (2015) The room temperature preservation of filtered environmental DNA samples and assimilation into a Phenol-Chloroform-Isoamyl alcohol DNA extraction. *Molecular Ecology Resources*, **21**, 2045-2050.
- RStudio Team (2015) RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. URL: <http://www.rstudio.com/>

- Ruetz III CR, Reneski MR, Uzarski DG (2012) Round goby predation on Dreissena in coastal areas of eastern Lake Michigan. *Journal of Freshwater Ecology*, **27**, 171-184.
- Shaw JLA, Clarke LJ, Wedderburn SD, Barnes TC, Weyrich LS, Cooper A (2016) Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biological Conservation*, **197**, 131-138.
- Sigsgaard EE, Carl H, Møller PR, Thomsen PF (2015) Monitoring the near-extinct European weather loach in Denmark based on environmental DNA from water samples. *Biological Conservation*, **183**, 46-52.
- Simmons M, Tucker A, Chadderton WL *et al.* (2016) Active and passive environmental DNA surveillance of aquatic invasive species. *Canadian Journal of Fisheries and Aquatic Sciences*, **73** 76-83.
- Singer SN, So SK (1980) Characterization, magnitude and impact of urban runoff in the Grand River basin. Water Resources Branch, the Ontario Ministry of the Environment.
- Smith JL (2014) Female Round Goby (*Neogobius melanostomus*) Movement Responses to Pheromones: An Investigation of Current Methods and Future Needs. (Master's Dissertation) Retrieved from The University of Windsor Leddy Library. Paper 5052.
- Species at Risk Act (SARA). 2016. Section 1-2, Interpretation. Published by the Minister of Justice. Accessed April 26, 2016 at <http://laws-lois-justice.gc.ca>
- Species at Risk Act (SARA). 2016. Schedule 1 (Subsections 2(1), 42(2), and 68(2)). Accessed on April 19, 2016 at http://www.registrelep-sararegistry.gc.ca/species/schedules_e.cfm?id=1
- Species at Risk Act (SARA). 2016. Schedule 3 (Section 130). Accessed on April 25, 2016 http://www.registrelep-sararegistry.gc.ca/species/schedules_e.cfm?id=3
- Staton SK, Dextrase A, Metcalfe-Smith JL *et al.* (2003) Status and trends of Ontario's Sydenham River ecosystem in relation to aquatic species at risk. *Environmental Monitoring and Assessment*, **88**, 283-310.
- St. Clair Region Conservation Authority (SCRCA) Watershed Report Card. (2013a). Upper Sydenham River: Watershed Report Card 2013. <http://www.scrca.on.ca/wp-content/uploads/2013/09/Report-Card-2013-Upper-Sydenham-River.pdf>
- St. Clair Region Conservation Authority (SCRCA) Watershed Report Card. (2013b). Lower East Sydenham: Watershed Report Card 2013. <http://www.scrca.on.ca/wp-content/uploads/2013/09/Report-Card-2013-Lower-East-Sydenham.pdf>
- Spear S, Groves JD, Williams LA, Waits LP (2015) Using environmental DNA methods to improve detectability in a hellbender (*Cryptobranchus alleganiensis*) monitoring program. *Biological Conservation*, **183**, 38-45.
- Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH (2012) Environmental DNA. *Molecular Ecology*, **21**, 1789-1793.
- Takahara T, Minamoto T, Yamanaka H, Doi H, Kawabata ZI (2012) Estimation of fish biomass using environmental DNA. *PLoS ONE*, **7**, e35868.
- Takahara T, Minamoto T, Doi H (2013) Using environmental DNA to estimate the distribution of an invasive fish species in ponds. *PLoS ONE*, **8**, e56584.
- Thompson HA, Simon TP (2014) Diet shift response in round goby, *Neogobius melanostomus*, based on size, sex, depth, and habitat in the western basin of Lake Erie. *Journal of Applied Ichthyology*, **30**, 955-961.
- Thomsen PF, Kielgast J, Iversen LL *et al.* (2012) Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology*, **21**, 2565-2573.
- Thomsen PF, Willerslev E (2015) Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, **183**, 4-18.
- Tréguier A, Paillisson J-M, Dejean T *et al.* (2014) Environmental DNA surveillance for invertebrate species: advantages and technical limitations to detect invasive crayfish *Procambarus clarkia* in freshwater ponds. *Journal of Applied Ecology*, **51**, 871-879.
- Tucker AJ, Chadderton WL, Jerde CL, *et al.* (2016) A sensitive environmental DNA (eDNA) assay leads to new insights on Ruffe (*Gymnocephalus cernua*) spread in North America. *Biological Invasions*, **18**, 3205-3222.
- Turner CR, Uy KL, Everhart RC (2015) Fish environmental DNA is more concentrated in aquatic sediments than surface water. *Biological Conservation*, **183**, 93-102.
- Valentini A, Taberlet P, Miaud C *et al.* (2016) Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, **25**, 929-942.
- Veech (2013) A probabilistic model for analyzing species co-occurrence. *Global Ecology and Biogeography*, **22**, 252-260.
- Wilcox TM, McKelvey KS, Young MK *et al.* (2013) Robust detection of rare species using environmental DNA: the importance of primer specificity. *PLoS ONE*, **8**, e59520.
- Wright J, Imhof J (2001) Technical background report for the Grand River fisheries management plan. Grand River Conservation Authority, the Ontario Ministry of Natural Resources.
- Yamanaka H, Minamoto T (2016) The use of environmental DNA of fishes as an efficient method of determining habitat connectivity. *Ecological Indicators*, **62**, 147-153.
- Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL (2012) Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, **13**, 134, doi:10.1186/1471-2105-13-134.
- Zhan A, Xiong W, He S, Maclsaac HJ (2014) Influence of artifact removal on rare species recovery in natural complex communities using high-throughput sequencing. *PLoS ONE*, **9**, e96928.

Contributions

RPW, NEM, and DDH contributed equipment and samples for eDNA analysis. KDB designed the custom river sequence database, and analyzed NGS data. KDB wrote the first draft of the manuscript, and RPW, NEM, and DDH edited and provided feedback for the manuscript, and made necessary changes.

Data Accessibility

NGS sequences will be available on NCBI SRA database (BioProject ID: PRJNA407636; BioSample Accession: SAMN07659297)

Supporting Information (Supplementary Tables S1, S2A, S2B, S2C, S3A, S3B, S3C, and S3D) as well as additional input files and parameters used in this study will be available on Dryad doi:10.5061/dryad.d52q0

Tables

Table 1.

Group-specific primer (PS1) set sequences developed in this study for the detection of rare, invasive, and common native fish species from environmental DNA. Associated average melting temperature (T_M), and amplicon size (bp). Annealing temperature used for the primer set was 52°C.

Primer Name	Sequence 5'-3'	Fragment Size (bp)	T_M (°C)
PS1-F	ACCTGCCTGCCGTATTTGGYGYCYTGRGCCGGRATAGT	247	85.0
PS1-R	ACGCCACCGAGCCARAARCTYATRTTRTTYATTCG		62.3

Note: Uni-A tail attached to 5' end of forward primers. Uni-B tail attached to 5' end of reverse primers.

Table 2.

Summary of eDNA NGS data for primer PS1 for samples with ≥ 3 eDNA sequence matches to target fish species. In total, 170 Grand River samples (108 surface, 58 bottom, 4 river controls) from 43 sites and 184 Sydenham River samples (109 surface, 72 bottom, 3 river controls) from 43^a were analyzed. ^a total sites reduced to 43 from 44 due to exclusion of site S20 because of possible DNA contamination.

	Eastern Sand Darter (<i>Ammocrypta pellucida</i>)	Silver Shiner (<i>Notropis photogenis</i>)	Northern Madtom (<i>Noturus stigmosus</i>)	Round Goby (<i>Neogobius melanostomus</i>)
Grand River				
No. of eDNA sequences	961	2 720	48	208 167
No. of Bottom Samples	1	3	0	43
No. of Surface Samples	4	1	1	70
No. Sites	4	3	1	38
Sydenham River				
No. of eDNA sequences	1 118	0	1 635	513 100
No. of Bottom Samples	2	0	4	65
No. of Surface Samples	2	0	3	99
No. Sites	4	0	7	42

Table 3.

Number of eDNA sequences detected in bottom and surface samples for each target fish species at risk in detected sites.

Northern Madtom			Eastern Sand Darter			Silver Shiner		
Sites	Bottom	Surface	Sites	Bottom	Surface	Sites	Bottom	Surface
G6		48	G1		116	G7	1 700	
			G9		113	G10	1 011	
			G10		673	G40		9
			G23	59				
Total	0	48	Total	59	902	Total	2 711	9
S4	107		S4		206			
S8	14		S29		489			
S14		9	S39	184				
S15		3	S40	239				
S21		1 124						
S27	363							
S41	15							
Total	499	1 136	Total	423	695	Total	0	0

Table 4.

Comparison of the number of species detected (yes) and not detected (no) between eDNA and capture-based methods for the Grand River and the Sydenham River. The total number of species (82) is the sum of detected species using both methods across both rivers.

Grand River	eDNA			Sydenham River	eDNA				
	Yes	No	Total		Yes	No	Total		
Capture-Based Methods	Yes	38	8	46	Capture-Based Methods	Yes	35	21	56
	No	12	24	36		No	19	7	26
	Total	50	32	82		Total	54	28	82

Figures

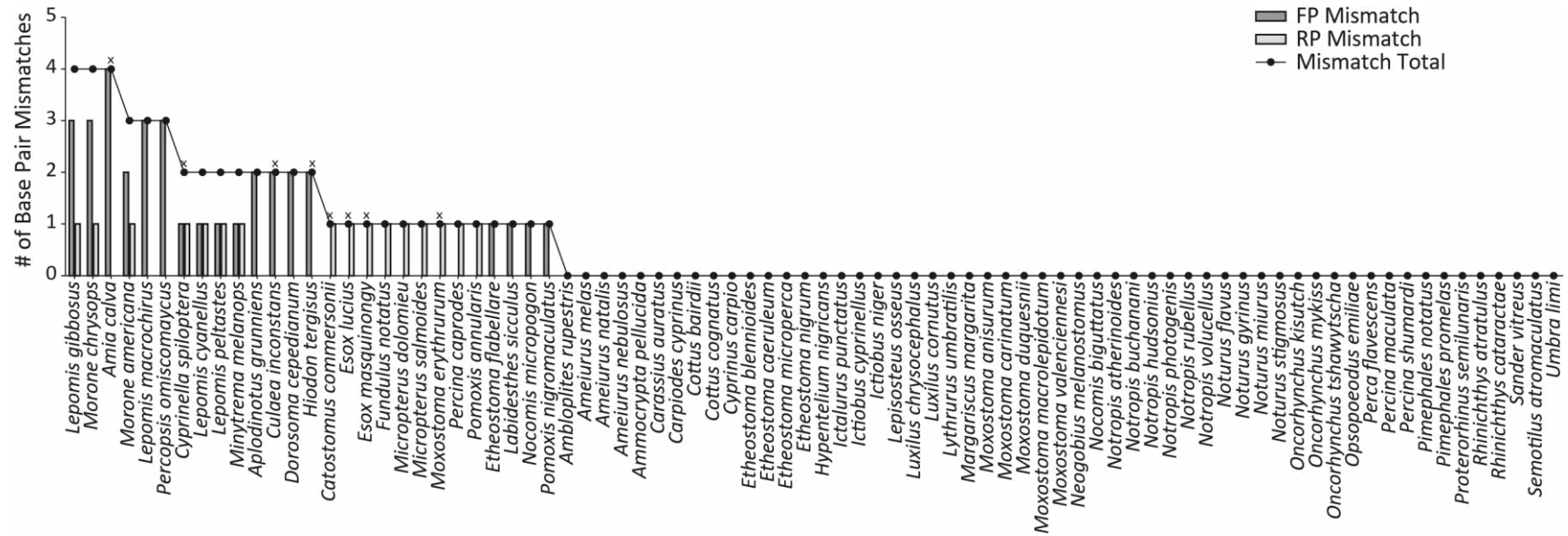


Figure 1. The number of PS1 primer mismatches (base pairs, bp) for forward primer (dark grey) and reverse primer (light grey) for the eDNA detected species. Black circles joined by the solid line indicate the total number of primer mismatches. The 'x' above certain species indicate a primer mismatch for species within 5 bp of the 3' end for either forward or reverse primer mismatch. N = 82 species detected from both eDNA and capture-based methods (out of 118 species in custom database).

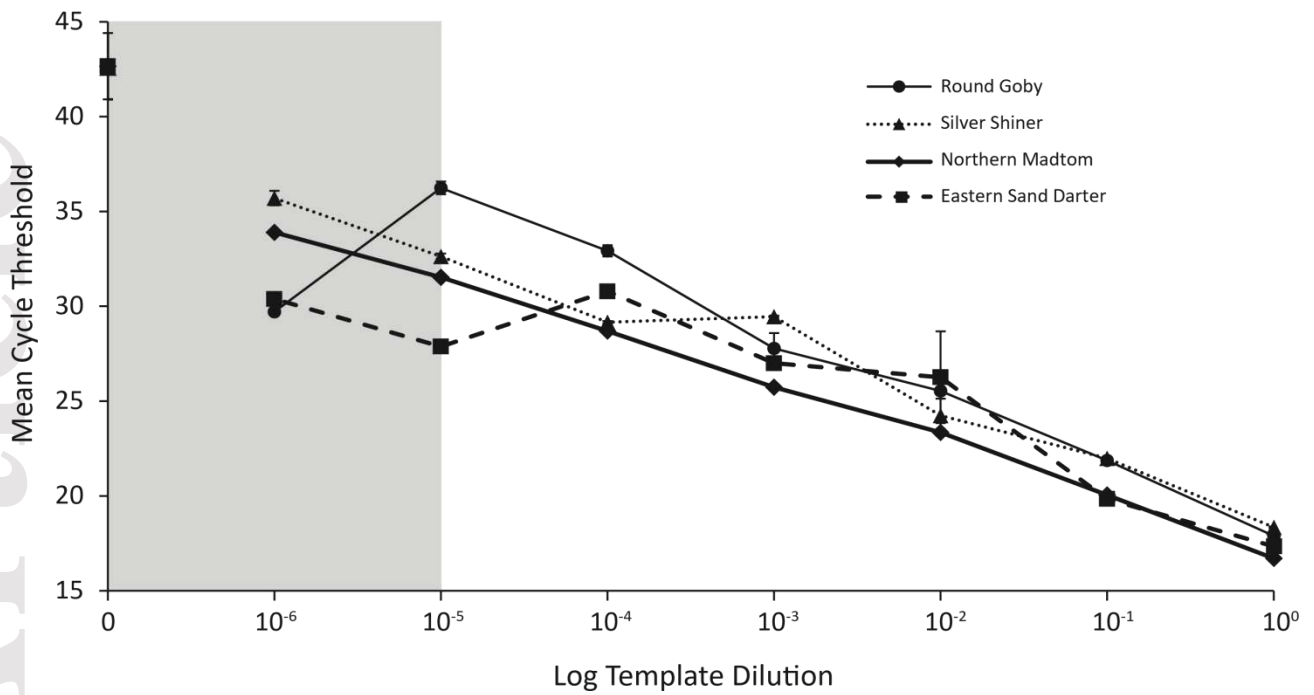


Figure 2. Sensitivity analysis of the group-specific COI PCR primer (PS1) designed to amplify Great Lakes fish species. The X-axis shows the 10-fold PCR template DNA dilution series used for qRT-PCR with PS1. Mean C_T (\pm SEM) is shown for each dilution (qRT-PCR done in triplicate), including the no-template-control shown as zero (box on Y-axis). Initial DNA template concentrations were $38.4 \text{ ng}\cdot\mu\text{L}^{-1}$ for Round Goby, $35.4 \text{ ng}\cdot\mu\text{L}^{-1}$ for Silver Shiner, $46.4 \text{ ng}\cdot\mu\text{L}^{-1}$ for Eastern Sand Darter, and $30.0 \text{ ng}\cdot\mu\text{L}^{-1}$ for Northern Madtom. Shaded region represents the plateau of the dilution curve and hence the limit of detection.

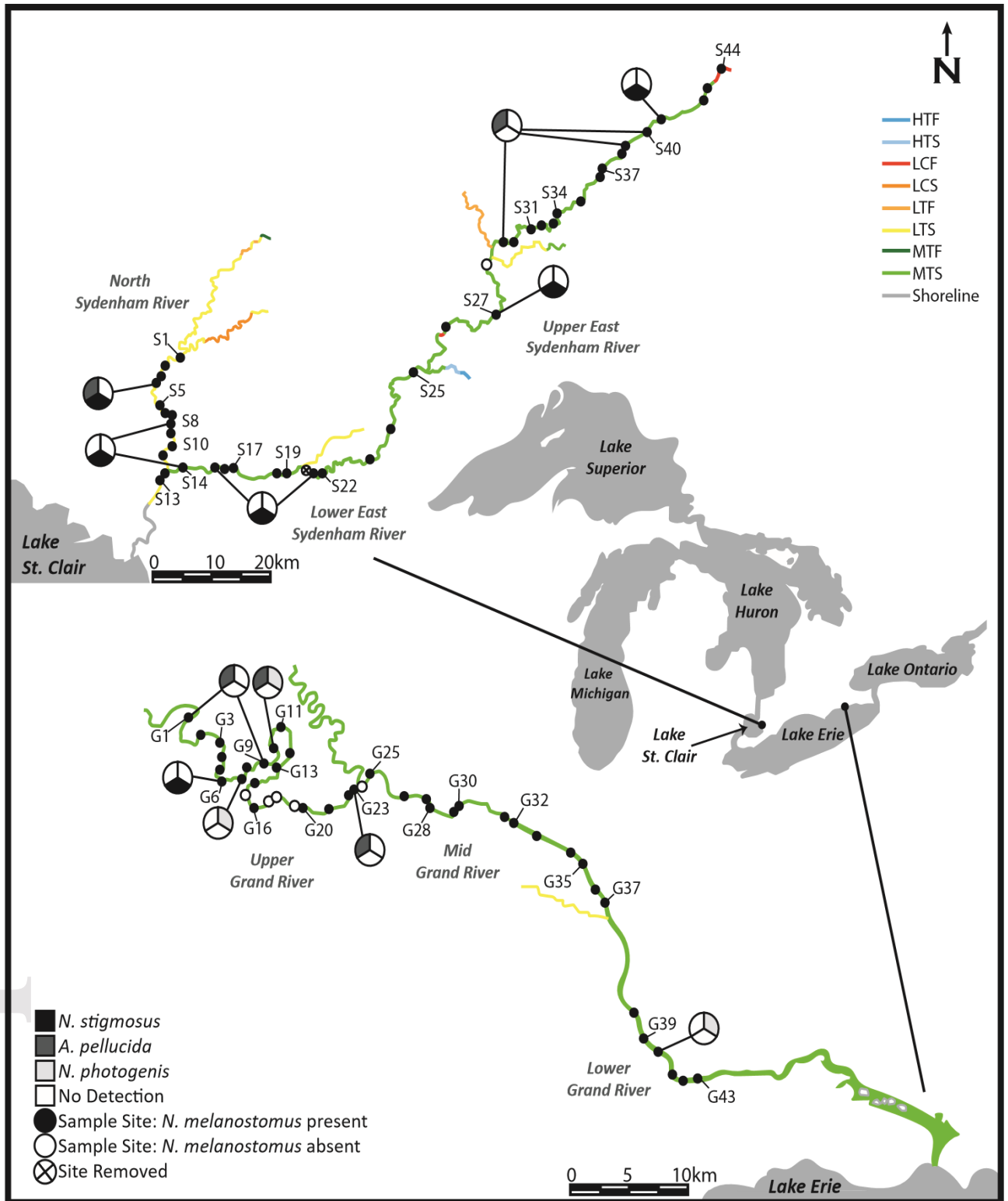


Figure 3. The spatial distribution based on the eDNA detection of three species at risk (Eastern Sand Darter (*Ammocrypta pellucida*), Northern Madtom (*Noturus stigmatosus*), Silver Shiner (*Notropis photogenis*)) and the invasive Round Goby (*Neogobius melanostomus*) in Sydenham River (44 sites) and Grand River (43 sites) in southern Ontario. Stream class and characteristics are shown in colour shading of the rivers and follow AEC protocol (Melles et al. 2013).

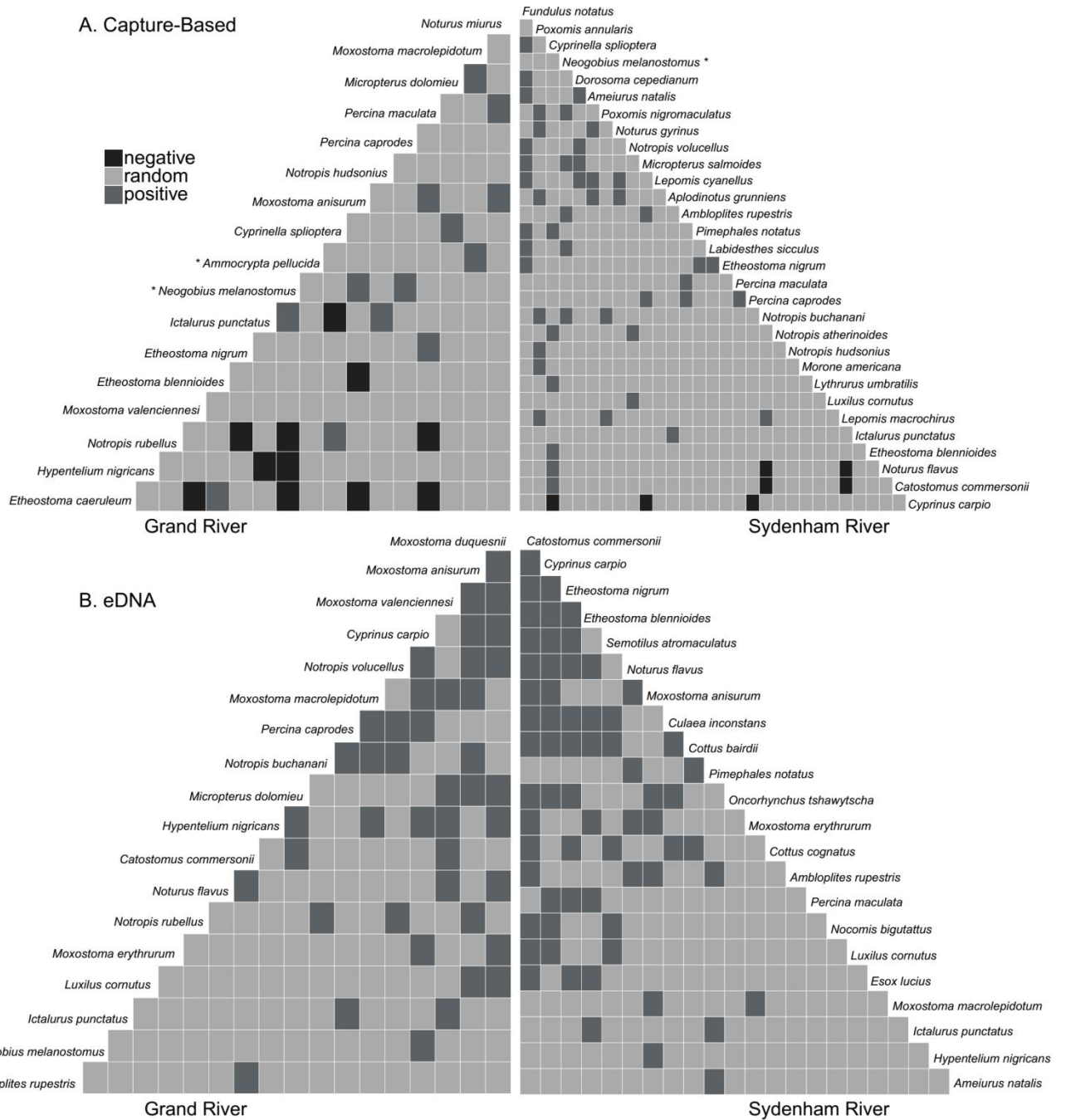


Figure 4. Species co-occurrence matrix for significant co-occurrences ($P < 0.05$) for (A) capture-based methods and (B) eDNA. Positive interactions (dark grey squares) are species that co-occur more often than expected, and negative interactions (black squares) are species that co-occur less often than expected. All other species interactions are random co-occurrences (light grey squares and do not differ from random expectations ($P > 0.05$; see Supp. Tables S3A, S3B, S3C, and S3D for species co-occurrences values). * Indicates target species.

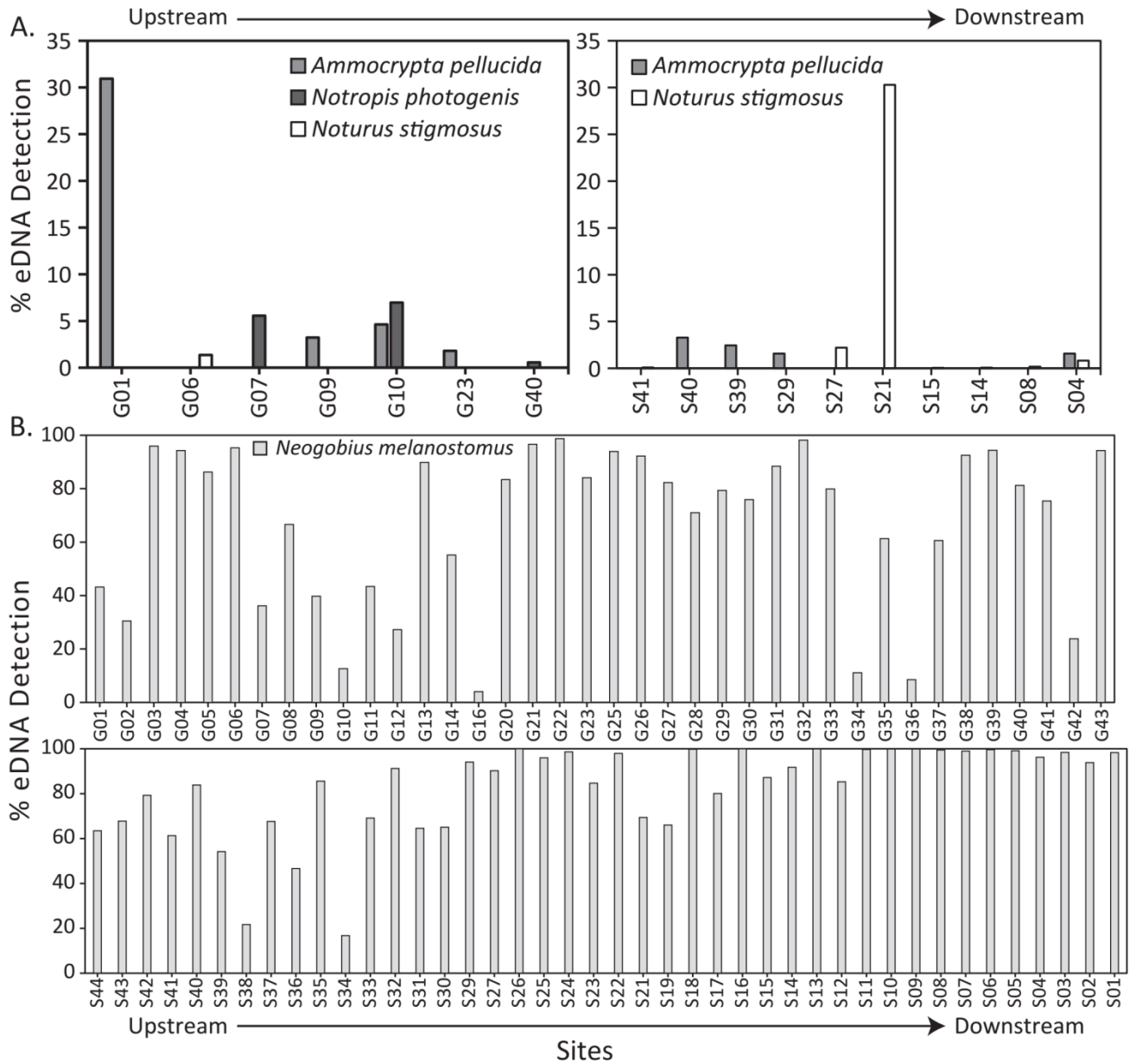


Figure 5. The percent total of eDNA next-generation sequences recovered from sites in the Grand River (G--) and the Sydenham River (S--) for target species. Panel (A) Species at Risk; *Ammocrypta pellucida*, *Notropis photogenis*, and *Noturus stigmosus*, and Panel (B) invasive *Neogobius melanostomus* in eDNA sample sites. Graphs display sites arranged from upstream to downstream.

Supplementary Information

Supplementary Table S1. List of fish species used in creating the custom COI database.

Supplementary Table S2A. Fish species currently or historically known to inhabit the Grand River and the Sydenham River (N = 82 species total) based on FishMaP. * Indicates specific target species in this study. Values given are the total number of eDNA sequences returned for the species at the sample site (G = Grand River, S = Sydenham River). Blank cells indicate no detection or number of eDNA sequences < 3. Site S20 removed due to eDNA contamination.

Supplementary Table S2B. Fish species currently or historically known to inhabit the Grand River and the Sydenham River (N = 82 species total) based on FishMaP. * Indicates specific target species in this study. Values given are the proportion of eDNA sequences returned for the species at the sample site (G = Grand River, S = Sydenham River), calculated as the number of eDNA sequences for a species divided by the total number of eDNA sequences obtained for the sample site. Blank cells indicate no detection or number of eDNA sequences < 3. Site S20 removed due to eDNA contamination.

Supplementary Table S2C. Fish species currently or historically known to inhabit the Grand River and the Sydenham River (N = 82 species total) based on FishMaP. * Indicates specific target species in this study. Values given are the number of live individuals caught at the sample site (G = Grand River, S = Sydenham River), using capture-based methods. Blank cells indicate no live individuals captured. Sites included overlap with eDNA sample sites.

Supplementary Table S3A. Statistically significant species co-occurrences (24 pairs; $P < 0.05$) using presence/absence data for 47 species detected using capture-based methods for the Grand River (R Studio v.0.99.892, package "cooccur"). ^a is the probability that the species pair co-occurs less often than observed (negative co-occurrence) and ^b is the probability that the species pair is likely to co-occur more often than observed. Number of species pairs (N) analyzed was 371. $N_{\text{random}} = 347$, $N_{\text{positive}} = 13$, and $N_{\text{negative}} = 11$.

Supplementary Table S3B. Statistically significant species co-occurrences (56 pairs; $P < 0.05$) using presence/absence data for 56 species detected using capture-based methods for the Sydenham River (R Studio v.0.99.892, package "cooccur"). ^a is the probability that the species pair co-occurs less often than observed (negative co-occurrence) and ^b is the probability that the species pair is likely to co-occur more often than observed. Number of species pairs (N) analyzed was 581. $N_{\text{random}} = 492$, $N_{\text{positive}} = 49$, and $N_{\text{negative}} = 7$.

Supplementary Table S3C. Statistically significant species co-occurrences (42 pairs; $P < 0.05$) using presence/absence data for 50 species detected using eDNA for the Grand River (R Studio v.0.99.892, package "cooccur"). ^a is the probability that the species pair co-occurs less often than observed (negative co-occurrence) and ^b is the probability that the species pair is likely to co-occur more often than observed. Number of species pairs (N) analyzed was 278. $N_{\text{random}} = 236$, $N_{\text{positive}} = 42$, and $N_{\text{negative}} = 0$.

Supplementary Table S3D. Statistically significant species co-occurrences (65 pairs; $P < 0.05$) using presence/absence data for 54 species detected using eDNA for the Sydenham River (R Studio v.0.99.892, package "cooccur"). ^a is the probability that the species pair co-occurs less often than observed (negative co-occurrence) and ^b is the probability that the species pair is likely to co-occur more often than observed. Number of species pairs (N) analyzed was 294. $N_{\text{random}} = 229$, $N_{\text{positive}} = 65$, and $N_{\text{negative}} = 0$.