

**ORIGINAL RESEARCH**

# Invasion genetics from eDNA and thousands of larvae: A targeted metabarcoding assay that distinguishes species and population variation of zebra and quagga mussels

Nathaniel T. Marshall<sup>1,2</sup> | Carol A. Stepien<sup>2</sup> 

<sup>1</sup>Genetics and Genomics Group, Department of Environmental Sciences, The University of Toledo, Toledo, Ohio

<sup>2</sup>Genetics and Genomics Group, NOAA Pacific Marine Environmental Laboratory, Seattle, Washington

**Correspondence**

Carol A. Stepien, Genetics and Genomics Group, NOAA Pacific Marine Environmental Laboratory, Seattle, WA.

Emails: carol.stepien@noaa.gov; stepien.carol@gmail.com; cstepien@uw.edu

**Funding information**

U.S. Environmental Protection Agency, Grant/Award Number: GL-00E01289 and GL-00E01898; Hudson River Foundation Bain Fellowship; International Association for Great Lakes Research scholarship

**Abstract**

Identifying species and population genetic compositions of biological invasions at early life stages and/or from environmental (e)DNA using targeted high-throughput sequencing (HTS) metabarcode assays offers powerful and cost-effective means for early detection, analysis of spread patterns, and evaluating population changes. The present study develops, tests, and applies this method with a targeted sequence assay designed to simultaneously identify and distinguish between the closely related invasive Eurasian zebra and quagga mussels (*Dreissena polymorpha* and *D. rostriformis*) and their relatives and discern their respective population genetic patterns. Invasions of these dreissenid mussel species have markedly changed freshwater ecosystems throughout North America and Europe, exerting severe ecological and economic damage. Their planktonic early life stages (eggs and larvae) are morphologically indistinguishable, yet each species exerts differential ecological effects, with the quagga often outcompeting the zebra mussel as adults. Our targeted assay analyzes genetic variation from a diagnostic sequence region of the mitochondrial (mt)DNA cytochrome oxidase I (COI) gene, to assess temporal and spatial inter- and intra-specific genetic variability. The assay facilitates analysis of environmental (e)DNA from water, early life stages from thousands of individuals, and simultaneous analysis of 50–100 tagged field-collected samples. Experiments evaluated its accuracy and performance using: (a) mock laboratory communities containing known DNA quantities per taxon, (b) aquaria with mixed-species/haplotype compositions of adults, and (c) field-collected water and plankton versus traditional sampling of adult communities. Results delineated species compositions, relative abundances, and population-level diversity differences among ecosystems, habitats, time series, and life stages from two allopatric concurrent invasions in the Great Lakes (Lake Erie) and the Hudson River, which had separate founding histories. Findings demonstrate application of this targeted assay and our approach to accurately and simultaneously discern species- and population-level differences across spatial and temporal scales, facilitating early detection and ecological understanding of biological invasions.

Laboratory group relocated from affiliation 1 to affiliation 2 in October 2016.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

## KEYWORDS

*Dreissena*, environmental DNA, Great Lakes, high-throughput sequencing, invasive species

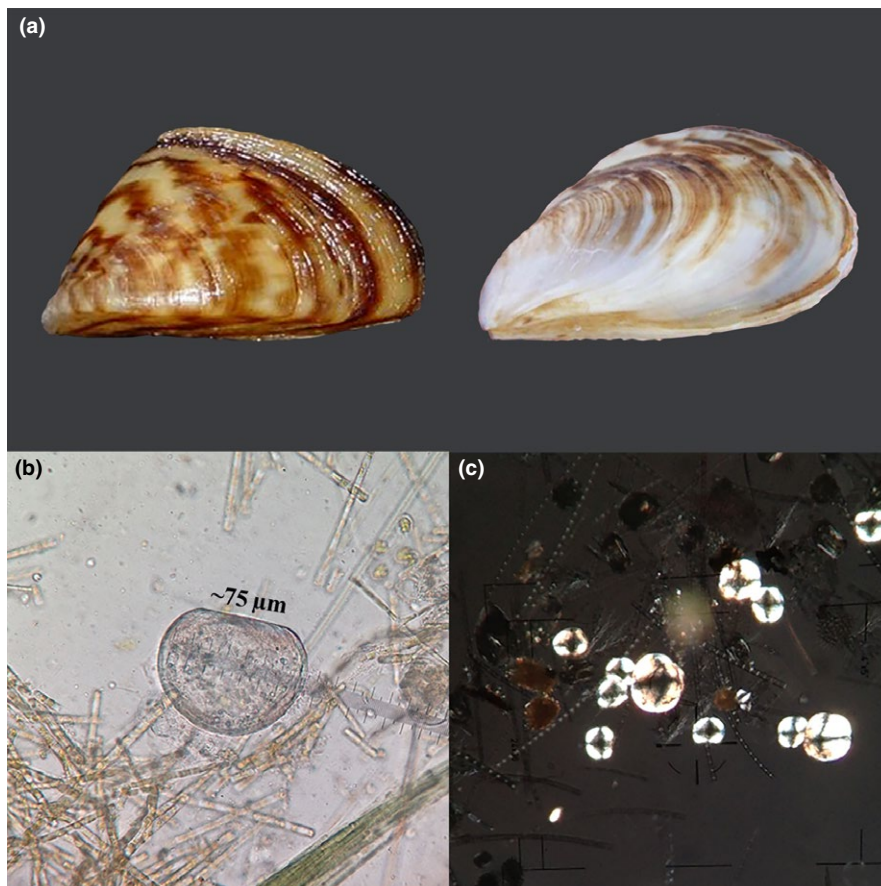
## 1 | INTRODUCTION

The growing global prevalence of invasive species profoundly shapes contemporary ecosystems (Pimentel, 2014; Simberloff, 2013) and negatively affects commerce (Keller, Lodge, Lewis, & Shogren, 2009; Pimentel, Zuniga, & Morrison, 2005). Ecological effects of invasive species range from competition with native species for food and space, alterations in habitat structure, and spread of pathogens and parasites (Gallardo, Clavero, Sánchez, & Vilà, 2016; Kvach & Stepien, 2008; Simberloff, 2013). The North American Laurentian Great Lakes is one of the most serious examples of a heavily invaded ecosystem, where >184 aquatic invasive species (AIS) have established (NOAA, 2018). Of these, the invasions of two closely related bivalve dreissenid mussels from Eurasia (Figure 1)—the zebra mussel *Dreissena polymorpha* (Pallas, 1771) and the quagga mussel *D. rostriformis* (Deshayes, 1838)—are regarded to have exerted some of the greatest ecological and economic impacts. Dreissenids are extensive biofoulers, prolific filterers, and ecosystem “engineers” that convert soft to hard benthos, often hosting entire communities of accompanying organisms on their shell surfaces and interstices (Stepien et al., 2013; Ward & Ricciardi, 2007).

Over the past 300 years, the zebra mussel has spread throughout much of Europe from its native Ponto Caspian region, and the quagga mussel for the last 130 years (van der Velde, Rajagopal, & Bij de Vaate, 2010; Zhulidov et al., 2010). The first North American appearance of dreissenids was reported in 1986, with the zebra mussel's discovery in Lake St. Clair of the Great Lakes, attributed to ballast water released from one or more trans-oceanic vessels (Hebert, Muncaster, & Mackie, 1989). This was followed by the 1989 detection of the quagga mussel in Lake Erie (May & Marsden, 1992).

Zebra mussel populations in the Great Lakes initially increased very rapidly, whereas the quagga mussel remained rare for several years, being primarily confined to deeper water habitats in eastern Lake Erie (Karatayev et al., 2014; Stepien, Hubers, & Skidmore, 1999). Near-complete replacement of the zebra mussel by the quagga mussel since has occurred throughout the central and eastern Lake Erie basins, with the western basin now harboring a mixed-species community, where shallower waters provide refugia for the zebra mussel (Karatayev et al., 2014). Lake Erie has the longest history of harboring established populations of both species (Benson, 2013), constituting a focal area for this study.

In comparison, the zebra mussel first invaded the Hudson River in 1991 (Strayer et al., 1996), followed by the quagga mussel in 2008 (Strayer



**FIGURE 1** (a) Adult zebra mussel *Dreissena polymorpha* (left) and quagga mussel *D. rostriformis* (right) (each specimen about 2.0 cm in length, (photographs not copyrighted and used courtesy of USGS NAS; <https://nas.er.usgs.gov/>)). *Dreissena* veliger larvae from our plankton samples (b) under compound light microscope (single larva) and (c) using cross-polarized light (multiple larvae) taken by NTM in this study

& Malcom, 2013). This zebra mussel population initially grew very rapidly, accounting for >50% of the benthic biomass by 1992 (Strayer et al., 1996). Genetic analyses using microsatellites showed that zebra mussel populations from the Hudson River and Lake Erie significantly differed, implicating different founding origins (Brown & Stepien, 2010; Stepien et al., 2013). By 2010, the quagga mussel had spread throughout the Hudson River's entire freshwater reach, yet has remained only a small fraction (<10%) of the dreissenid community (Strayer & Malcom, 2013; Strayer, pers. comm.). The respective histories and relative compositions of the two species in the Hudson River and Lake Erie thus comprise an ecological contrast for the present investigation.

Early detection and continued monitoring of dreissenid introductions and spread patterns are high management agency priorities in the United States (NOAA, 2018; USGS, 2018), Canada (DFO, 2018; MNRF, 2018), and Europe (van der Velde et al., 2010; Zhulidov et al., 2010). Ecological effects of dreissenid invasions largely depend on their respective species composition, time elapsed since their establishment, population density, and hydrological characteristics of the waterbody (Benson, 2013). Predicting their ecological impacts thus relies on up-to-date information about their relative species composition and abundances. However, these sister taxa are difficult to distinguish to species as adults, sometimes even by trained malacologists, and cannot be differentiated at early life stages (as eggs or larvae; Nichols & Black, 1994; Claxton & Boulding, 1998).

Dreissenids are prolific, reproducing multiple times each season, with >1,000,000 eggs per female zebra mussel (Borcherding, 1991), and have free-swimming veliger larvae that widely disperse (Ackerman, Sim, Nichols, & Claudi, 1994; Figure 1b–c). Adult mussels attach in tremendous numbers to vessels, tackle, and drifting objects and can survive out of water for a week or longer (Collas et al., 2014; Stepien et al., 2013); they thus are readily transported among bodies of water. New introductions often pose significant biosecurity risks (Robinson, Burgman, & Cannon, 2011), since dreissenid populations often are not identified until they are already established and widespread, and by then are difficult (or impossible) to eradicate (Zaiko, Minchin, & Olenin, 2014). Genetic tools that effectively identify and diagnose dreissenids at early life stages (from plankton) and at low densities (e.g., environmental (e)DNA from water samples) would significantly improve timely detection, identification, and monitoring.

Over the past two decades, identifying individual species from their DNA via “barcoding” (see Hebert, Cywinska, Ball, & deWaard, 2003) has increased diagnosis of taxa (Radulovici, Archambault, & Dufresne, 2010). Analysis of eDNA shed into water (via urine, waste, mucus, sloughed cells) has facilitated early detection of some AIS, but mostly has been limited to the detection of single species (Darling & Mahon, 2011; Ficetola, Miaud, Pompanon, & Taberlet, 2008; Jerde, Mahon, Chadderton, & Lodge, 2011). For example, eDNA was used to discover the quagga mussel in the Rhine River harbor of Basel, Switzerland, which had been missed by traditional monitoring (De Ventura, Kopp, Seppälä, & Jokela, 2017). Similarly, the zebra mussel was discerned from eDNA in the Red River off Lake Winnipeg, Canada, in 2014, a year before the first adult was located (Gingera, Bajno, Docker, & Reist, 2017).

Although those studies detected dreissenids (De Ventura et al., 2017; Egan et al., 2015; Gingera et al., 2017; Ram et al., 2011), their methodology was unable to differentiate between zebra and quagga mussel species using a single assay and could not address population-level variation (e.g., haplotypes). Targeted metabarcoding assays with high-throughput sequencing (HTS) and bioinformatic processing can be designed to simultaneously resolve entire communities to species level, as well as distinguish population genetic variation (Elbrecht, Vamos, Steinke, & Leese, 2018; Parsons, Everett, Dahlheim, & Park, 2018; Sigsgaard et al., 2016; Stepien, Snyder, & Elz, 2019).

## 1.1 | Research objective and applications

The research objective was to develop, test, and apply a new targeted HTS metabarcoding assay that simultaneously identifies and differentiates among all six *Dreissena* species (see Stepien et al., 2013; Marshall & Stepien, 2019), including zebra and quagga mussels, and discerns their respective population-level variability (i.e., haplotype representation and relative abundances) at all life stages. Experiments tested the assay's accuracy and performance on mixed communities of zebra and quagga mussel species and their haplotypes in: (a) a series of mock laboratory communities containing known quantities of DNA, (b) aquarium experiments, and (c) field-collected eDNA water and plankton (larvae) versus results from traditional sampling. Dreissenid mussel community compositions were compared between two allopatric ecosystems in Lake Erie and the Hudson River, which have separate invasion histories. Field experiments evaluated community differences across a temporal collection series in Lake Erie encompassing the reproductive season, and along a spatial gradient along the Hudson River. We also compared and contrasted species- and population-genetic variability from two sites in Lake Erie, one in the western basin where the zebra mussel is more prevalent and the other in the central basin, which is dominated by the quagga mussel. These are statistically compared with results from the Hudson River sites, which also are dominated by the zebra mussel.

Our assay aimed to simultaneously assess inter- and intra-specific genetic variation from up to thousands of individuals in environmental samples, allowing for detection from eDNA in water and plankton of early life stages. It was designed to alleviate morphological mis-identifications. At the community level, accurate identification of dreissenids to species and assessment of their population genetic variability are projected to facilitate understanding their relative reproductive outputs, spread patterns, ecological interactions, successes, and adaptations over temporal and spatial scales.

## 2 | MATERIALS AND METHODS

### 2.1 | HTS assay design

First, we screened and aligned sequences for the “barcode” region of the mitochondrial (mt) cytochrome oxidase I (COI) gene from NIH

NCBI GenBank (<https://www.ncbi.nlm.nih.gov/genbank>) and our own laboratory databases for all six species of *Dreissena* along with the two species of its sister genus *Mytilopsis* (Marshall & Stepien, 2019; Stepien et al., 2013). Primer pairs were designed by evaluating the aligned sequences with ClustalX in MEGA7 (Kumar, Stecher, & Tamura, 2016) and selecting regions of diagnostic variation that are flanked by conserved regions. The aligned database identified 20 unique zebra mussel haplotypes (labeled ZM-A through ZM-T) and 17 quagga mussel haplotypes (labeled QM-A through QM-Q) in a 570-nucleotide (nt) region of COI (Supporting information Tables A1–A4). Two markers were developed and tested to target different nonoverlapping variable regions of ~200 nucleotides (nt), which encompassed intraspecific variation for zebra and quagga mussels and species-specific diagnosis for all six species of *Dreissena*, along with the two species of the closely related genus *Mytilopsis* (see Stepien et al., 2013 for taxonomy). We then focused on testing the two markers on zebra and quagga mussels, which were designed to each distinguish different single nucleotide polymorphisms (SNPs).

Primer set 1 (COIA-F: 5'AGTGTYYTKATTCTGTTTRGAGCTWAGKGC3', -R: 5'GAYAGGTARAACCCAAAACWCTWAC3') flanks a 169-nt region that contains 24 fixed differences between zebra and quagga mussels. It delineates 10 zebra mussel and eight quagga mussel haplotypes and differentiates among all six *Dreissena* species (Tables A1–A2). Primer set 2 (COIB-F: 5'GGATTYGGKGGKGGTTGAAC3', -R: 5'GRAAWCTKGTMACCAATAGAWGT3') flanks a 175-nt region with 26-nts that differentiate zebra and quagga mussels. It delineates seven zebra mussel and nine quagga mussel haplotypes and also distinguishes among all six *Dreissena* species (Tables A3–A4). Our two primer sets distinguish seven of the nine quagga mussel COI haplotypes described by Marescaux et al. (2016), found throughout its native and invasive ranges.

## 2.2 | Mock communities and aquaria experiments

Performance of the assays was tested on eight mock communities containing various known concentrations of DNA zebra and quagga mussel haplotypes (termed amplicon sequence variants (ASVs; see Callahan, McMurdie, & Holmes, 2017)). We thus tested whether the

mock communities statistically maintained the relative ASV proportions during HTS (following our method previously published in Klymus, Marshall, & Stepien, 2017). Mixtures of purified DNA containing various proportions of 47–6,000 copies each of five zebra mussel haplotypes (ZM-A, ZM-C, ZM-G, ZM-M, and ZM-O) and three quagga mussel haplotypes (QM-A, QM-F, and QM-G) for the COI gene were analyzed (Table 1). We quantified the target copy number per DNA extraction for each ASV in a series of competitive polymerase chain reactions (PCR) that co-amplified the native template (NT) and an internal standard (IS; see Klymus et al., 2017). Our COIA assay (Tables A1 and A2) was designed to differentiate among three zebra mussel haplotypes (ZM-C, ZM-G, and the ZM-A,M,O group) and three quagga mussel haplotypes (QM-A, QM-F, and QM-G; Tables A1 and A2); thus, six ASVs were identifiable in the mock community with this marker. The COIB assay was designed to differentiate four zebra mussel haplotypes (ZM-G, ZM-M, ZM-O, and the ZM-A,C group) and one quagga mussel haplotype group (QM-A,F,G group); thus, five ASVs would be diagnosable in the mock community (Tables A3 and A4).

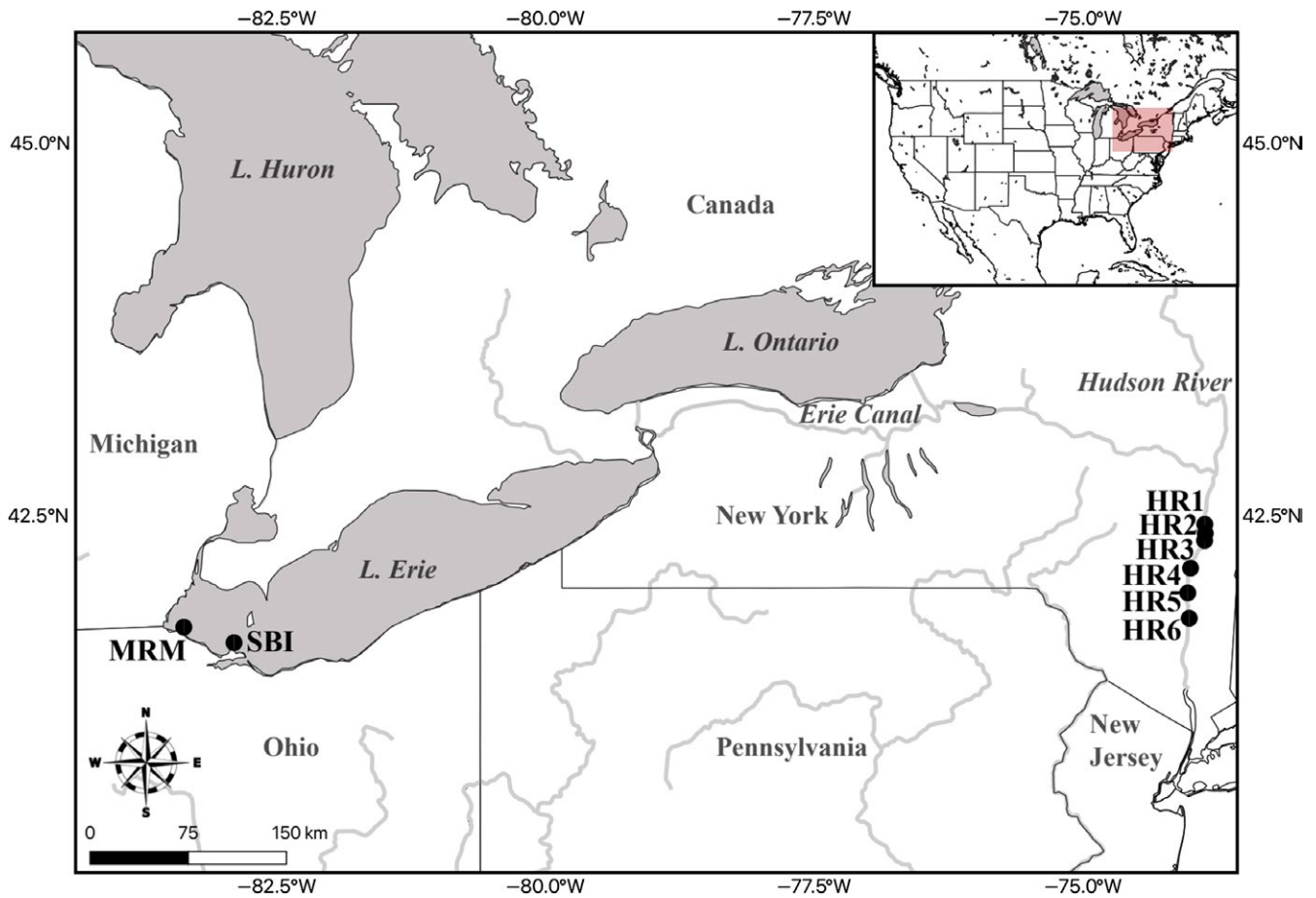
Adult quagga mussels were collected from Belle Isle in the Detroit River, Detroit, MI (42.34, -82.99), and zebra mussels from the western Lake Erie basin in Oregon, OH (41.67, -83.29), both in April 2016 (Figure 2). Three 38-L aquaria, air stones, and tubing were cleaned with 10% bleach and filled with 15 L of dechlorinated tap water. Aquaria were covered in clean plastic wrap and oxygenated throughout the experiments. Prior to introduction of the mussels, the water in each aquarium was thoroughly mixed, and a 500 ml aliquot was taken using a sterile container for eDNA analysis; this served as a negative control. The removed water was replaced with dechlorinated water from a sterile container. Mixed proportions of adults of the two species, ranging from 26 to 36 individuals total, were placed into each aquarium (Table 2). At days 2, 7, and 14, the water in each aquarium was sampled as above for eDNA analysis.

Water samples (including the negative controls) each were filtered through a 0.2- $\mu$ m polyethersulfone (PES) filter, from which eDNA was extracted using the Qiagen DNeasy<sup>®</sup> Blood and Tissue Kit (Qiagen Inc., Valencia, CA, USA), with the amounts of proteinase K and ATL doubled in order to fully submerge the filter. Extracted

Haplotype	Mock Communities							
	1	2	3	4	5	6	7	8
QM-A	6,000	3,000	1,500	750	375	188	94	47
ZM-A	3,000	1,500	750	375	188	94	47	6,000
ZM-M	1,500	750	375	188	94	47	6,000	3,000
ZM-C	750	375	188	94	47	6,000	3,000	1,500
QM-G	375	188	94	47	6,000	3,000	1,500	750
QM-F	188	94	47	6,000	3,000	1,500	750	375
ZM-G	94	47	6,000	3,000	1,500	750	375	188
ZM-O	47	6,000	3,000	1,500	750	375	188	94

**TABLE 1** Sequence copy numbers of each species and haplotype that were used per mock community (MC 1–8), of the target mitochondrial DNA cytochrome oxidase I (COI) gene amplicon

Note. ZM = *Dreissena polymorpha* (zebra mussel), QM = *D. rostriformis* (quagga mussel). Haplotypes are lettered.



**FIGURE 2** Site map for dreissenid mussel plankton and eDNA water samples, from the Hudson River (HR, locations 1–6) and Lake Erie (Maumee River's mouth (MRM) and South Bass Island (SBI))

DNA was run through a Zymo Research One Step PCR Inhibitor Removal kit (<http://www.zymoresearch.com/rna/rna-clean-up/rt-pcr-inhibitor-removal/onestep-pcr-inhibitor-removal-kit>).

At the end of the experiment (day 14) and after the final water collection, mussels were removed, air dried for one hour, and individually weighed (g). They then were slowly cooled in a refrigerator for 24 hr and then individually labeled and frozen. DNA from each was extracted, amplified, and Sanger sequenced for the COI gene using primers and methodology from Folmer, Black, Hoeh, Lutz, and Vrijenhoek (1994). PCRs contained 1X Radiant<sup>®</sup> (Alkali Scientific) PCR buffer, 0.75 mM dNTPs, 3 mM MgCl<sub>2</sub>, 0.5 μM of each primer, 1.5U Radiant<sup>®</sup> Taq, 100 ng of template DNA, and the necessary volume of ddH<sub>2</sub>O to comprise 25 μl. The haplotypes were compared to those in GenBank and our laboratory database (Stepien et al., 2013).

### 2.3 | Field eDNA water and plankton samples

Paired 1L water samples were taken for eDNA analysis from three Hudson River locations (Figure 2), including: Stuyvesant, NY (abbreviated HR1; 42.39, -73.78; depth = ~3.7 m), Coxsackie, NY (HR2; 42.35, -73.80; located four km downriver from Stuyvesant; depth =

~3.4 m), and Stockport, NY (HR3; 42.31, -73.77; located five river km downriver from Coxsackie; depth = ~4.3 m), on August 6, 2016. They were collected in sterilized bleach-washed containers, one from 15 cm below the surface and the other 15 cm above the benthos (the latter were taken and capped by a team of SCUBA divers from the Cary Institute of Ecosystem Studies, who were careful not to touch or disturb the benthos). Water samples were labeled, placed on ice in the field, and then frozen at -20°C in the laboratory. Simultaneously, the SCUBA diving team surveyed the adult dreissenid mussel communities by collecting 10–12 rocks (15–40 cm in size), from which the mussels were counted in the Cary Institute laboratory (totaling 2,984 mussels from Stuyvesant, 2,174 from Coxsackie, and 3,345 from Stockport). Length–mass regression analysis was performed on a representative range of sizes for ~300 mussels to estimate respective biomass per species. In the laboratory, half of each water sample (500 ml) was pelleted in 10 sterile 50-mL falcon tubes at 7,500 g centrifugation for 30 min. DNA was extracted with a cetyl-trimethyl ammonium bromide (CTAB) protocol (Turner et al., 2014), and purified with a Zymo Research One Step PCR Inhibitor Removal kit, with the resultant DNA used for HTS.

**TABLE 2** Numbers (N), weight (g), and percentage of species and haplotypes (lettered) used in aquarium environmental DNA experiments (Aquaria 1–3) with assay COIA. Mussels were weighed at the end of the experiment (at 14 days)

Species and Haplotype	Aquarium 1				Aquarium 2				Aquarium 3			
	N	%	Mass (g)	%	N	%	Mass (g)	%	N	%	Mass (g)	%
ZM-A	24	66.67	3.03	60.21	12	38.71	1.14	35.88	15	57.69	2.68	29.43
ZM-M	6	16.67	0.75	14.58	5	16.13	0.65	18.97	1	3.85	0.06	0.63
Total Zebra Mussels	30	83.33	3.78	74.79	17	54.84	1.79	51.80	16	61.54	2.74	30.06
QM-A	6	16.67	1.28	25.21	13	41.94	1.56	45.30	10	38.46	6.37	69.94
QM-G	0	0.00	0	0.00	1	3.23	0.10	2.90	0	0.00	0	0.00
Total Quagga Mussels	6	16.67	1.28	25.21	14	45.16	1.66	48.20	10	38.46	6.37	69.94
Total Mussels	36		5.06		31		3.45		26		9.11	

Triplicate plankton samples were collected from four Hudson River sites (Figure 2)—Stuyvesant, NY (HR1), North Germantown, NY (HR4; 42.16, -73.89), Rhinecliff, NY (HR5; 41.92, -73.95), and Poughkeepsie, NY (HR6; 41.72, -73.94)—on May 31, 2016, with a 63- $\mu$ m Wisconsin Cole-Parmer 40-AD50 plankton net (<https://www.coleparmer.com>) attached to a 15.2 m line. The Poughkeepsie site was resampled on June 2, 2016, and August 6, 2016. For comparative analyses, plankton also was taken from two Lake Erie locations, at the Maumee River's mouth (MRM; 41.69, -83.43; ~1.5 m deep, sandy bottom, zebra mussel refugium) and South Bass Island (SBI; 41.63, -82.84; ~5 m deep, rocky bottom, quagga mussel dominated), in the afternoon every two weeks from June through September 2016. The net was thrown from the seawall/shore and, upon sinking, was maintained at the maximum depth the line would allow for 60 s, then manually retrieved using a hand-over-hand technique ~0.5 m/s, resulting in ~0.19 m<sup>3</sup> (~190 L) of water sampled. Upon retrieval, each sample was placed into labeled 15-mL falcon tubes containing 95% EtOH.

Numbers of *Dreissena veliger* larvae per site from the plankton tows were estimated from three separate 500  $\mu$ l aliquots using a Sedgewick-Rafter counting cell (Thermofisher<sup>®</sup>, <https://www.thermofisher.com>) under cross-polarized light microscopy, following Johnson (1995). Density (number of *Dreissena veligers* per liter; Hosler, 2011) was determined from the mean of the three aliquots per sample, multiplying by the sample volume (15 ml) divided by the filtered volume (~190 L), and reported as the mean of replicates  $\pm$  standard error (SE). Numbers of *Dreissena* larvae were statistically compared using paired Student's *t* tests (Sokal & Rohlf, 1994). One plankton sample from each site was centrifuged at 6,000 rpm for 5 min, and the ethanol decanted. DNA was extracted from the resultant pellet using the CTAB soil extraction protocol (Turner et al., 2014), with the resultant DNA amplified for HTS.

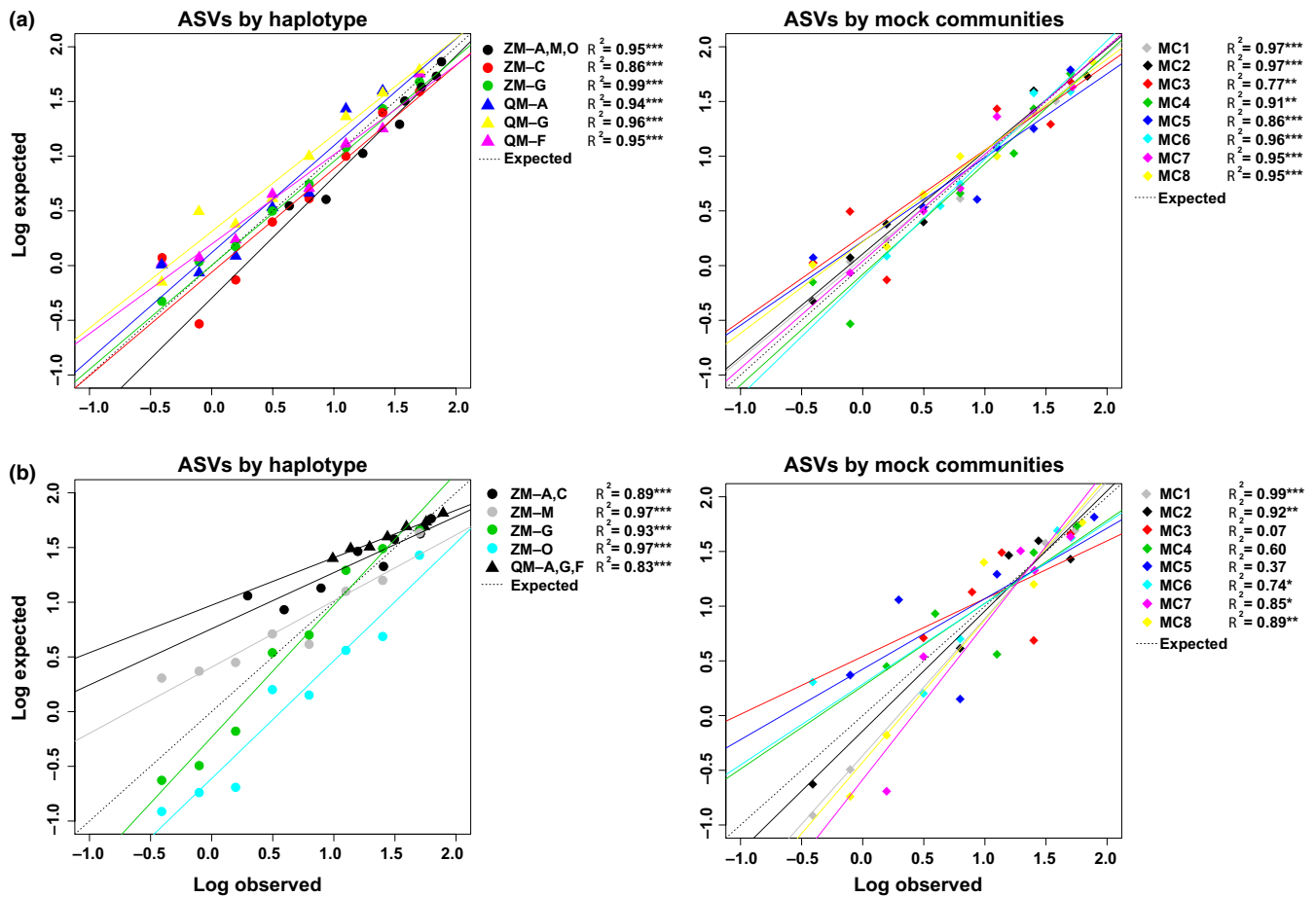
## 2.4 | MiSeq HTS library preparation

A four-step PCR library preparation was used prior to paired-end HTS on the Illumina<sup>®</sup> MiSeq platform (<https://www.illumina.com/systems/sequencing-platforms/miseq.html>). After each PCR, products were column cleaned with QIAquick<sup>®</sup> PCR purification kits (Qiagen; <https://www.qiagen.com>), yielding the template for the subsequent step. The

first step of four cycles used assay-specific primers (COIA or COIB) to append 21-nt APEX tails (F:5'GATCAGGCGTCTGTCGTGCTC3', R:5'CTCGACGACAGACGCCTGATC3') to the target regions, which aimed to reduce PCR bias by retaining the relative ASV proportions of the sample throughout library preparation (Krujtskov et al., 2008). PCRs contained 1X PCR buffer, 0.2 mM dNTPs, 0.5  $\mu$ M of each primer, 1U AmpliTaq<sup>®</sup>, 2.5  $\mu$ l template DNA, and the necessary volume of ddH<sub>2</sub>O to constitute 25  $\mu$ l. Conditions were as follows: 30-s initial denaturation at 95°C, followed by four cycles of 95°C for 30-s, 62°C for 30-s, and 72°C for 1-min, with 2-min final extension at 72°C. Next, the APEX primers were used to amplify targets adhered with the APEX tail, using the same PCR conditions for 35 cycles.

The third step amplified previous products using APEX primers appended with 7- to 17-nt (termed spacers) and a 33- to 34-nt sequencing primer on the 5' end. The spacer region increased nt diversity of sequence reads and enhanced cluster formation, improving sequencing quality (Fadrosh et al., 2014; Wu et al., 2015). Four primer sets were used, which differed in their spacer combinations (see Klymus et al., 2017). PCR conditions were identical, except that the cycle number was reduced to eight. The final (fourth) PCR incorporated Nextera paired-end indices (Illumina<sup>®</sup>, kit FC-121-1011), as well as the P5 and P7 adaptor sequences, allowing the sample to bind onto the Illumina<sup>®</sup> MiSeq flowcell. The 25  $\mu$ l PCRs contained a variable volume of ddH<sub>2</sub>O, 1x PCR buffer, 0.2 mM dNTPs, 2.5  $\mu$ l of each primer, 1.57U NEB Hotstart Taq polymerase (New England Biolabs<sup>®</sup>, Inc.), and 2.5  $\mu$ l of the previous PCR cleanup. Conditions were as follows: 30-s initial denaturation at 95°C, followed by eight cycles at 95°C for 30-s, 55°C for 30-s, and 68°C for 1-min, with a final 2-min 68°C extension. Four samples were run with the Mol16S primer set followed the library preparation, which was published in Klymus et al. (2017).

To avoid sequencing dimer product, targeted fragments first were size-selected with a 1.5% agarose gel cassette on Pippin Prep (Sage Science). PCR products were sized and quantified on a 2100 Bioanalyzer (Agilent Technologies). A negative control was run for each library preparation step and checked for contamination with gel electrophoresis, with the final step checked on the 2100 Bioanalyzer to verify the absence of product. No contamination was found in any instance. Pooled product concentrations



**FIGURE 3** Regression analyses of log transformed observed versus expected read percentages per haplotype (lettered and colored) of zebra mussel (ZM) and quagga mussel (QM) and mock community (MC) for the two assays (a) COIA and (b) COIB. Significant at  $* = 0.05$ ,  $** = 0.01$ ,  $*** = 0.001$

were measured on a Qubit fluorometer (Invitrogen), and the pooled samples then were run on an Illumina<sup>®</sup> MiSeq with 2X 300nt V3 chemistry at the Ohio State University's Molecular and Cellular Imaging Center in Wooster, OH (<http://mcic.osu.edu/>). An additional 40%–50% PhiX DNA spike-in control was added to improve data quality of low nt diversity samples (Klymus et al., 2017). HTS was performed on four MiSeq runs, which also contained tagged samples from other projects to increase diversity; all were pooled to yield >100,000 reads per sample.

## 2.5 | Bioinformatic and statistical analyses

MiSeq results were analyzed using the R package DADA2 (Callahan et al., 2016), except for three mock community samples that had low read depths (COIA mock community 5, and COIB mock communities 2 and 7). We instead merged and trimmed those with the Python package OBITools 1.2.11 (Boyer et al., 2016). For the other samples, the primer and spacer regions first were trimmed using a custom Perl script. Next, using the default settings in DADA2, error rates were estimated, sequences were merged and grouped into ASVs, and any erroneous sequences and chimeras were removed. Chimeras were not removed from the COIB dataset, since DADA2 mistakenly

regarded haplotype ZM-E as a chimeric sequence. Sequence ASVs were identified to dreissenid species and haplotypes in reference to COI sequences on GenBank and our own database. Possible erroneous ASVs were eliminated from our environmental datasets by removing any sequence present only in a single sample.

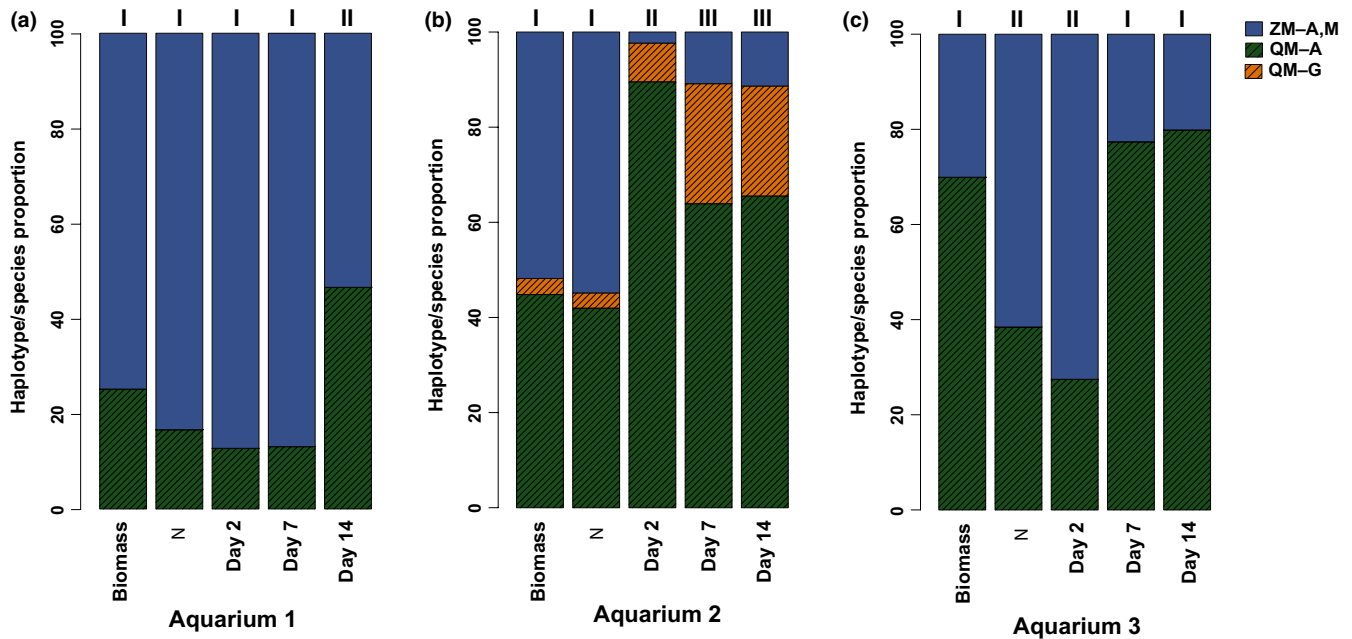
For mock communities, the percentage of reads per ASV was calculated and analyzed using R v3.4.3 (R Core Team, 2015) by regressing the logarithm of their observed versus expected percent read abundances to evaluate goodness of fit and whether the slopes differed from 1.00. Contingency tests (Sokal & Rohlf, 1994) were used to compare HTS reads from (a) aquarium eDNA water having known ASV compositions, (b) Hudson River eDNA water samples with traditional morphological census counts, and (c) plankton across temporal and spatial scales. To compare relatedness among multiple-ASV community assemblages, Nonmetric multidimensional scaling (NMDS; Kruskal, 1964) ordinations were performed using Bray-Curtis dissimilarity indices of the relative abundances of ASVs using the R package vegan (Oksanen et al., 2018) and the *metaMDS* and *anosim* functions to parse the data according to collection locations. Differences in the relative abundances of ASVs among the samples were evaluated using the pairwise Kulczynski (1927) Dissimilarity Index (KDI), which ranged from 0 (equal) to 1 (completely different).

**TABLE 3** Numbers of sequence reads and observed and expected percentages per taxon and mock community (MC 1–8) for the two assays (COIA and COIB)

Assay	Haplotype	MC 1			MC 2			MC 3			MC 4		
		Read number	% Obs.	% Exp.	Read number	% Obs.	% Exp.	Read number	% Obs.	% Exp.	Read number	% Obs.	% Exp.
COIA	ZM-A,M <sub>i</sub> O	44,985	31.4	38.0	99,521	53.7	69.0	42,468	20.6	34.5	21,288	10.9	17.3
	ZM-C	5,649	3.9	6.3	5,058	2.7	3.1	1,687	0.8	1.6	602	0.3	0.8
	ZM-G	1,695	1.2	0.8	1,544	0.8	0.4	98,326	47.8	50.2	52,667	26.9	25.1
	QM-A	82,114	57.3	50.2	71,510	38.6	25.1	54,804	26.6	12.6	8,828	4.5	6.3
	QM-G	6,321	4.4	3.1	4,521	2.4	1.6	6,454	3.1	0.8	1,540	0.8	0.4
	QM-F	2,577	1.8	1.6	3,237	1.7	0.8	2024	1.0	0.4	111,060	56.7	50.2
COIB	ZM-A,C	125,618	37.8	31.4	681	28.8	15.7	34,876	14.2	7.9	17,288	9.1	3.9
	ZM-M	43,243	13.0	12.6	66	2.8	6.3	13,122	5.3	3.1	5,523	2.9	1.6
	ZM-G	1,040	0.3	0.8	5	0.2	0.4	114,816	46.6	50.2	61,521	32.4	25.1
	ZM-O	508	0.2	0.4	641	27.1	50.2	12,560	5.1	25.1	7,076	3.7	12.6
	QM-A,FG	161,660	48.7	54.9	974	41.1	27.5	70,771	28.8	13.7	105,464	55.6	56.9
Assay	Haplotype	MC 5			MC 6			MC 7			MC 8		
		Read number	% Obs.	% Exp.	Read number	% Obs.	% Exp.	Read number	% Obs.	% Exp.	Read number	% Obs.	% Exp.
COIA	ZM-A,M <sub>i</sub> O	58	4.0	8.6	9,389	3.7	4.3	42,379	43.2	52.2	94,049	72.5	76.1
	ZM-C	17	1.2	0.4	97,801	38.8	50.2	24,253	24.7	25.1	13,038	10.1	12.6
	ZM-G	173	12.0	12.6	14,465	5.7	6.3	3,045	3.1	3.1	2,342	1.8	1.6
	QM-A	49	3.4	3.1	3,152	1.2	1.6	906	0.9	0.8	1,264	1.0	0.4
	QM-G	886	61.5	50.2	94,738	37.6	25.1	22,516	23.0	12.6	12,859	9.9	6.3
	QM-F	257	17.9	25.1	32,723	13.0	12.6	4,911	5.0	6.3	6,110	4.7	3.1
COIB	ZM-A,C	30,616	12.7	2.0	143,328	43.3	51.0	539	21.8	25.5	180,808	59.3	62.7
	ZM-M	6,448	2.7	0.8	7,192	2.2	0.4	1,066	43.1	50.2	53,123	17.4	25.1
	ZM-G	49,714	20.6	12.6	18,299	5.5	6.3	78	3.2	3.1	2082	0.7	1.6
	ZM-O	4,003	1.7	6.3	5,842	1.8	3.1	5	0.4	1.6	747	0.2	0.8
	QM-A,FG	154,937	64.1	78.4	162,485	49	39.2	784	31.7	19.6	68,943	22.6	9.8

Note. ZM = *Dreissena polymorpha* (zebra mussel), QM = *D. rostriformis* (quagga mussel). Haplotypes are lettered.





**FIGURE 4** Species and haplotype proportions of zebra mussels (ZM; *Dreissena polymorpha*) and quagga mussels (QM; *D. rostriformis*) from environmental DNA in three aquarium experiments (A–C) using our targeted high-throughput sequencing COIA assay, in relation to biomass (weight, g) and number of individuals (N). Water samples taken on days 2, 7, and 14. Haplotypes are lettered and colored. Roman numerals denote significant concordance and/or differences in relative proportion of sequence reads for each aquarium using contingency G-tests

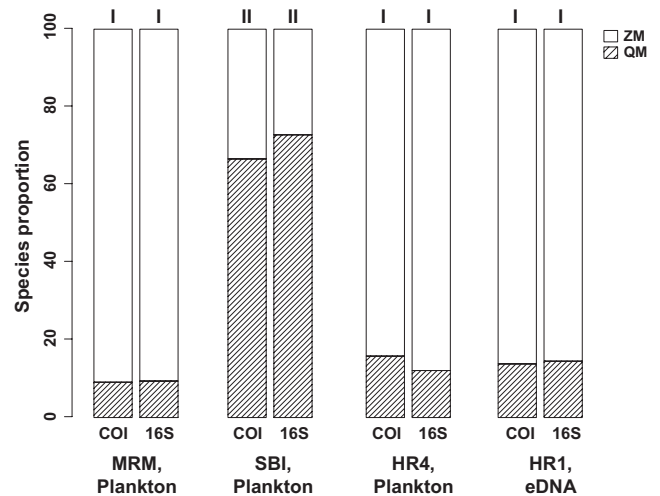
### 3 | RESULTS

#### 3.1 | HTS of mock communities and aquaria

Number of HTS reads passing cluster quality filtering ranged from 14,329,380 to 15,237,283 per run (Table A5). Regression analyses revealed significant matches ( $p < 0.05$ ) between the observed and expected sequence read percentages for the mock community samples with the COIA primers (Figure 3a; Table 3), corresponding to the expected slope = 1.00. In contrast, the COIB marker displayed a weaker relationship between observed and expected ASV proportions for mock communities 3–5 ( $p > 0.05$ ; Figure 3b; Table 3), with the slopes of 3–7 significantly differing from 1.00 (ranging from 0.53 to 1.42). As the COIB marker yielded inconsistent results with the mock communities, it was not used for further analysis with the aquaria or environmental samples. However, the HTS results between the two markers did not significantly differ for determining overall species composition in six of the eight mock communities, and both showed similar trends between observed and expected ASV proportions.

None of the negative controls from the three aquaria amplified, indicating that the water, aquaria, and aeration systems were free of DNA contamination. Both species and all of their expected haplotypes were identified with HTS from each aquarium on all dates (Figure 4). Biomass of the dreissenids significantly differed among all aquaria ( $***p = 0.001$ ), whereas the numbers of individuals per species significantly differed just between aquarium #1 and the other two (#1 vs. 2, #1 vs. 3:  $***p < 0.001$ ). Numbers of individuals

per species versus their relative biomass significantly differed for #3 alone ( $***p < 0.001$ , Figure 4). For #3, the zebra mussels were relatively small in size in comparison with the quagga mussels; thus, the



**FIGURE 5** Species proportions of zebra mussels (ZM; *Dreissena polymorpha*) and quagga mussels (QM; *D. rostriformis*) from plankton and eDNA field experiments using our targeted high-throughput sequencing COIA assay (this paper) and our previously published 16S assay (Klymus et al., 2017). Sampling locations were as follows: Maumee River's mouth (MRM) in Lake Erie (July 1, 2016), South Bass Island (SBI) in Lake Erie (July 1, 2016) and Hudson River (HR4, at North Germantown, NY on May 31, 2016, and HR1 at Stuyvesant, NY, on August 6, 2016). Roman numerals designate concordance among samples at sites using contingency G-tests

relative biomass of the latter was much greater than their relative numbers (Table 2).

Species proportions in the HTS reads from aquarium #1 at days 2 and 7 did not significantly differ from the relative numbers of individuals per species or from biomass (Figure 4a). On day 14, aquarium #1's sequence reads significantly differed from both its numbers of individuals per species ( $***p < 0.001$ ) and biomass ( $***p < 0.001$ ). Sequence reads from aquarium #2 significantly differed from numbers of individuals per species and biomass on all three days (Figure 4b). A quagga mussel died in aquarium #2, which went unnoticed. It likely died early in the experiment, increasing the amount of quagga mussel DNA throughout the duration. Sequence reads per species in aquarium #3 did not differ from the number of individuals per species at day 2, but differed from their respective biomass ( $***p < 0.001$ ; Figure 4c). However, at days 7 and 14, its sequence reads did not differ from biomass per species, but varied from the numbers of individuals per species ( $***p < 0.001$ ). This result resembled that of aquarium #1.

By the end of the experiment, the relative proportions of all sequence reads per species significantly differed from those at day 2 (day 2 vs. 14:  $***p < 0.001$  for all three aquaria; Figure 4), revealing a trend that more closely matched biomass at day 14, rather than the numbers of individuals (as occurred on day 2). The change to resembling biomass per species was observed at day 7 for aquaria #2 and #3, but at day 14 in #1.

### 3.2 | HTS of field eDNA water and plankton samples

The proportions of sequence reads obtained per species did not statistically differ between the new COIA (COI) assay (this study) and results using our previously published mollusk mtDNA 16S RNA (16S) assay (from Klymus et al., 2017), for the three plankton collection locations (ZM:QM proportions at the Maumee River's mouth (MRM), Lake Erie: 16S-9:91% versus COI-9:91%; South Bass Island (SBI), Lake Erie: 16S-73:27% versus COI-67:33%; North Germantown, Hudson River (HR4):16S-16:84% versus COI-12:88%; Figure 5). The read proportions from each assay also did not statistically differ in the eDNA water sample (Stuyvesant, Hudson River (HR3): 16S-14:86% versus COI-15:85%; Figure 5). The COIA was dreissenid specific, as 100% of the HTS results corresponded to zebra and quagga mussels. In contrast, the proportion of dreissenid sequences was much lower for the 16S assay (ranging from 0.36% to 80.25%).

Additionally, the proportions of species-specific sequence reads from the Hudson River eDNA water samples were statistically similar to the species ratios of the benthic adults determined by the Cary Institute's traditional morphological identifications (Figure 6). In both data sets and at all locations, zebra mussels greatly outnumbered quagga mussels. The Cary Institute's proportions were 95.0%–96.5% (mean = 95.7%) zebra mussels from morphological counts and 91.2%–92.3% (mean = 91.7%) zebra mussels using estimated biomass. Their species' proportions did not significantly

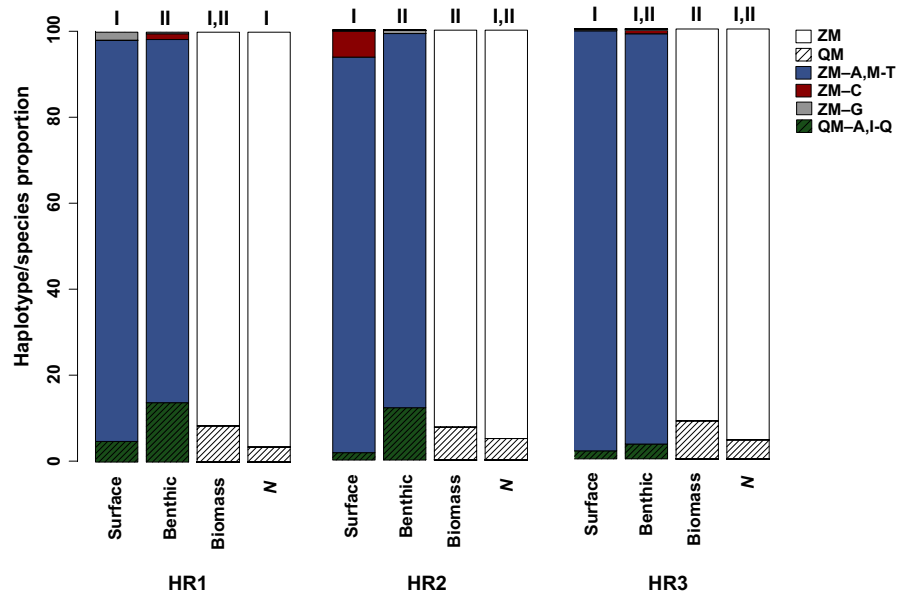
differ among the three Hudson River locations. Our HTS reads from eDNA water (surface and benthic samples) similarly yielded 83.5%–97.7% (mean = 92.1%) zebra mussels. Total read proportions (surface + benthic eDNA) did not significantly differ among the locations. Overall, the average ZM:QM proportions (surface + benthic eDNA) did not significantly differ between the HTS results and the Cary Institute's morphological counts or from the estimated biomass at any site.

For the eDNA water collected near the benthos, the HTS species' proportions differed between Stuyvesant versus Stockport ( $**p = 0.006$ ) and Coxsackie versus Stockport ( $*p = 0.01$ ), whereas none of the three surface samples differed. Species proportions significantly varied between the surface and benthic eDNA at Stuyvesant ( $*p = 0.02$ ) and Coxsackie ( $**p = 0.001$ ). The HTS species proportions (93.9%–97.7%, mean = 96.5%) from the surface eDNA water more closely matched the Cary Institute's morphological species counts (no significant differences at any site) than did their biomass estimates (significant variations at Stockport,  $*p = 0.02$  and Coxsackie,  $*p = 0.04$ ; Figure 6). The benthic eDNA sequence reads (83.5%–95.5% zebra mussels, mean = 87.8%) more closely matched the Cary Institute's biomass estimates (no significant differences at any site), than their adult mussel counts (significant difference for Stuyvesant,  $*p = 0.02$ ; Figure 6).

Physical counts of *Dreissena veliger* larvae made under the microscope from the plankton samples were consistently lower from the Maumee River's mouth than at the South Bass Island and Hudson River locations (mean MRM = 11.9/L (2,261/sample); SBI = 55.8/L (10,602); HR = 35.7/L (6,783/sample)). The numbers of veligers significantly differed between the Lake Erie sites (Maumee River's mouth and South Bass Island) throughout the sampling season ( $***p < 0.001$ ; Figure 7a,b). The relative HTS proportions of zebra mussels to quagga mussels (ZM:QM) in the plankton differed spatially across the Hudson River locations, revealing increasing proportions of quagga mussel veligers from upstream (north) to downstream (south; 0%–35%, mean = 20%). These species proportions significantly differed between the northern site at Stuyvesant (100% ZM) and the southern location at Poughkeepsie (70:30% ZM:QM,  $***p < 0.001$ ). Two plankton samples collected three days apart at Poughkeepsie yielded similar species proportions (64:36%, 67:33%), indicating short-term temporal consistency, whereas the August 2, 2016, sample significantly differed from the earlier ones ( $***p < 0.001$ ; Figure 7c).

In western Lake Erie, the relative HTS proportions of ZM:QM in plankton averaged 70:30% (ranging from 37%–97% ZM) at the Maumee River's mouth and 34:66% (22%–41% ZM) at South Bass Island, which differed significantly in depth (with the Maumee River's mouth being shallower) and were located ~36 km apart. Overall HTS proportions of ZM:QM significantly differed between the two locations ( $***p < 0.001$ ), matching their adult communities. The species proportions significantly differed between the two sites on all common dates (June 3,  $***p < 0.001$ ; June 17,  $***p < 0.001$ ; July 1,  $***p < 0.001$ ), except July 19.

**FIGURE 6** Species and haplotype proportions of zebra mussels (ZM; *Dreissena polymorpha*) and quagga mussels (QM; *D. rostriformis*) from field eDNA samples using our targeted high-throughput sequencing COIA assay, in relation to estimated biomass (weight, g) and number of individuals ( $n$ ) in the adult community of the Hudson River. eDNA water samples collected from the surface and the benthos at three sites (Stuyvesant, NY (HR1), Coxsackie, NY (HR2), and Stockport, NY (HR3)). Haplotypes are lettered and colored. Roman numerals denote significant concordance and/or differences in relative proportion of sequence reads per site using contingency G-tests



Species composition of the veliger larvae significantly varied temporally at the Maumee River's mouth (mean  $**p = 0.008$ ), with fewer zebra mussels and more quagga mussels at the earliest (ZM:QM 57:43% June 3, 2016) and later dates (37:63% July 19, 2016). Zebra mussel veligers were overwhelmingly prevalent at the two intermediate dates (97:03% June 17; 90:10% July 1; Figure 7a). Dreissenid veligers were absent in the August and September microscope counts of the Maumee River's mouth samples, and DNA extractions from the plankton did not amplify. No significant temporal differences occurred in the HTS species proportions of ZM:QM larvae at the South Bass Island location, with quagga mussels always being more prevalent (Figure 7b).

HTS species compositions of the larvae, averaged across all sampling dates per site, did not significantly differ between the Hudson River (all four locations added together) and the Maumee River's mouth, reflecting the common overall prevalence of zebra mussels at both. Samples of larvae from both locations significantly differed from South Bass Island, where quagga mussels dominated (MRM vs. SBI:  $***p < 0.001$ ; HR vs. SBI:  $***p < 0.001$ ). NMDS likewise depicted the South Bass Island larval community as distinct from the other two (MRM vs. SBI:  $*p = 0.02$ ; HR vs. SBI:  $***p < 0.001$ ; HR eDNA vs. SBI:  $**p = 0.002$ ; Figure 8a). Pairwise KDI tests supported the NMDS findings, with the lowest KDI values for each plankton sample typically corresponding to the others from the same location (Table A6). When samples were grouped by location, the lowest KDI values occurred between the Hudson River and Maumee River's mouth (KDI = 0.10), and the largest between the Hudson River versus South Bass Island (KDI = 0.58; Table A6).

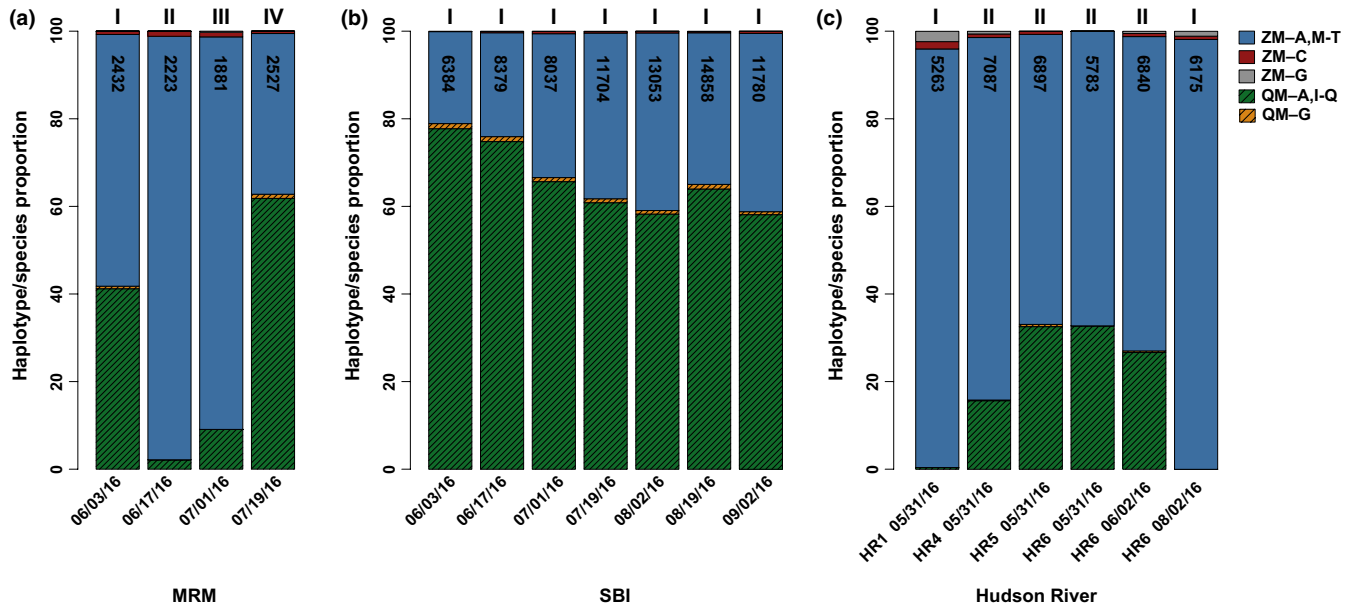
### 3.3 | Haplotypic diversity

Two rare zebra mussel haplotypes (ZM-C and ZM-G) and one rare quagga mussel haplotype (QM-G) were detected in the HTS analyses of eDNA water and plankton, with all three haplotypes previously

known in GenBank. Haplotype ZM-C was present in all but one Lake Erie plankton sample on August 2 (0.87% at MRM and 0.33% at SBI) and occurred throughout all Hudson River sites (0.78% in plankton and 1.37% in eDNA; Figures 7 and 8b). We found the rare haplotype ZM-G throughout the Hudson River (four sites at 0.76% in plankton and all locations at 0.65% in eDNA), which was scarce in Lake Erie (one MRM sample at 0.07% and four from SBI at 0.04%). The QM-G haplotype was discerned in plankton at both Lake Erie locations (0.39% at MRM and 0.98% at SBI) and in the Hudson River (0.18%), but was absent from our Hudson River eDNA samples (Figure 8b).

## 4 | DISCUSSION

This study developed and tested a new HTS metabarcoding assay that simultaneously differentiated among dreissenid species and their intraspecific haplotypes from single samples containing DNA from thousands of individuals. Since a single dreissenid mussel adult can release >1,000,000 gametes in a spawning season (Borcherding, 1991), and subsequent larval stages persist throughout the water column from early May through September (Borcherding, 1991; Sprung, 1993), water and plankton provide an opportunity to rapidly survey DNA from vast numbers of individuals, facilitating early and accurate detection of AIS. Estimated numbers of veliger larvae analyzed per our single plankton samples ranged from 2,220 to 14,860 individuals, demonstrating the utility of our assay and approach to assess the genetic variation of thousands of individuals. Thus, our metabarcoding approach offers greater ability to effectively discern species and population genetic diversity versus time-intensive, single individual processing with traditional sampling and sequencing. This HTS assay also accurately reveals species proportions within varying environmental samples (both water and plankton), as our sequencing results matched those estimated using our laboratory's previously published mollusk 16S assay (Klymus et al., 2017). That



**FIGURE 7** Species and haplotype proportions of zebra mussels (ZM; *Dreissena polymorpha*) and quagga mussels (QM; *D. rostriformis*) from field plankton samples collected from (a) Maumee River's mouth (MRM) in Lake Erie, (b) South Bass Island (SBI) in Lake Erie, and (c) the Hudson River at Stuyvesant, NY (HR1), North Germantown, NY (HR4), Rhinecliff, NY (HR5), and Poughkeepsie, NY (HR6), using our targeted high-throughput sequencing COIA assay. Haplotypes are lettered and colored. Roman numerals denote significant concordance and/or differences in relative proportion of sequence reads for each location using contingency G-tests. Numbers in columns are comparative dreissenid veliger larvae count estimates from plankton tows (three replicates each, counted under cross-polarized light microscope)

16S assay, unlike our new COI assay, does not resolve intraspecific variation. Furthermore, the 16S assay amplifies a wide range of taxa and thus might fail to detect a small dreissenid population.

The COI HTS assay was successful for evaluating representative haplotypic proportions of zebra and quagga mussels in laboratory and field tests. Sequencing results for the laboratory mock communities demonstrated strong relationship between observed and expected proportions. Aquaria experiments housing differing species compositions yielded results similar to their known species' biomass proportions. Although the HTS results from the aquaria initially resembled the numbers of individuals per species at day 2, sequences from all three tanks best reflected biomass by day 14. This relationship may have lagged while the mussels acclimated to the aquaria by adjusting their filtration rates. Considering that the sequence read output per species more closely matched biomass after longer acclimation times (days 7 and 14), eDNA sampled from field collections also should accurately reflect the respective biomass of species compositions, illustrating widespread ecological application of our assay. Experiments with fishes (Takahara, Minamoto, Yamanaka, Doi, & Kawabata, 2012) and amphibians (Evans et al., 2016; Thomsen et al., 2012) likewise found that eDNA quantification from mesocosms correlated well with biomass of the target organisms.

Our HTS assay accurately characterized dreissenid mussel community-level species compositions, along with their population genetic variability, from field-collected water. HTS proportions from their eDNA revealed similar compositions to the adult dreissenid mussel communities, matching morphological counts and estimated

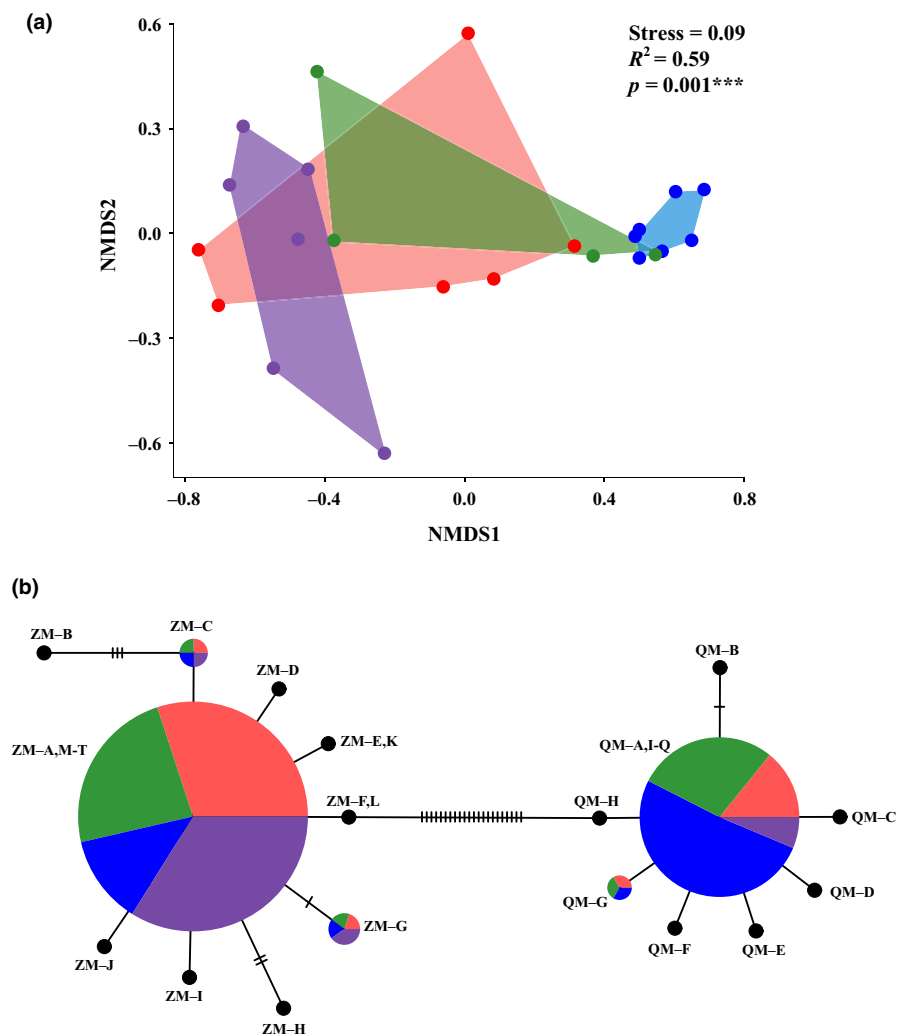
biomass. In the Hudson River, HTS species determinations for the surface water closely resembled the morphological identification counts of the benthic adults, whereas the benthic water showed greater quagga mussel proportions, which more closely resembled the biomass estimates. Similarly, by day 14 of the aquaria experiments, quagga mussel HTS proportions were slightly greater than expected in all three aquaria, due to their larger biomass. This trend might be attributed to the larger sizes of the quagga mussel individuals, and possible physiological differences between the species. For example, a study by Diggins (2001) found that quagga mussels had higher filtration rates (>37%) than zebra mussels of the same size (20 mm), which could lead to higher rates of eDNA shedding, but remains to be experimentally verified. In our study, when the surface and benthic eDNA results were averaged together, their species' proportional sequence reads did not differ from either their morphological counts or estimated biomass at any field location. These results suggest that slight differences in DNA concentrations can occur throughout the water column, and water collected from near the benthos (where most adults live) would best represent the true dreissenid biomass composition. A qPCR study of zebra mussel eDNA also found correspondence to benthic biomass of adults (Amberg, Merkes, Stott, Rees, & Erickson, 2019). As in the present study, eDNA surveys in the Rhine River using species-specific quantitative (q)PCR assays matched the respective species proportions from field surveys (De Ventura et al., 2017). Their approach differed from ours as they required two different probes to differentiate between the species and were unable to evaluate haplotypic (intraspecific) variation.

Sequencing results with our HTS assay accurately depicted the two allopatric dreissenid invasions, with the Hudson River and South Bass Island in Lake Erie sites being the most divergent. These results matched the morphological species compositions of the two invasions, with the quagga mussel dominating at South Bass Island (Karatayev et al., 2014) and the zebra mussel most prevalent in the Hudson River (Strayer & Malcom, 2013). As expected, our assay showed that the zebra mussel dominated the plankton from the Maumee River's mouth, since that region has been described as having the sole zebra mussel refugium in Lake Erie (Karatayev et al., 2014). Additionally, our assay displayed temporal resolution at the Maumee River's mouth, whose larval species composition varied throughout the sampling season. Throughout June, its species composition more closely resembled that from the Hudson River, but then changed to resemble South Bass Island on the final sampling date (July 19). This may have been due to larval transport, as Lake Erie experienced strong westward winds during that time period (wind data from the NOAA National Data Buoy Center (NDBC) at buoy SBI01 located at South Bass Island, and buoy THR01 at the Maumee River's mouth; <http://www.ndbc.noaa.gov>). These winds may have caused an influx of quagga mussel veliger larvae traveling from central Lake Erie to the west. For the Hudson River, HTS

of eDNA water and plankton agreed in species composition and haplotypic diversity representation, indicating that both sampling techniques yield similar results. These complementary collection methods thus surveyed a large breadth of dreissenid inter- and intra-specific variation in situ.

The quagga mussel has been outcompeting the zebra mussel in many invaded European rivers, increasing by an estimated 26% per year in Germany and the Netherlands (Heiler et al., 2013). Large fluctuations in their densities, with quagga mussels eventually displacing zebra mussels, were observed in the northern Dnieper River (Ukraine; Zhulidov et al., 2010). A similar takeover by the quagga mussel might substantially alter ecological processes in the Hudson River. For example, the quagga mussel's displacement of the zebra mussel in the lower Great Lakes increased total dreissenid biomass, as the quagga mussel more readily spreads across soft substrates, magnifying overall ecological impacts (Nalepa, Fanslow, & Lang, 2009). Moreover, selective filtering and phosphorous excretion appear to be contributing to increased harmful algal blooms in the lower Great Lakes and may significantly differ between the two dreissenid species (Conroy et al., 2005; Vanderploeg et al., 2001). Therefore, it is important to effectively discern the species compositions of dreissenid communities spatially and temporally. This can be

**FIGURE 8** (a). Nonmetric multidimensional scaling (NMDS) plot (Bray–Curtis distance metric of relative abundance) showing the relationships among dreissenid mussel communities from field samples using our targeted high-throughput sequencing COIA assay. Purple = Hudson River eDNA water, Red = Hudson River plankton, Green = Maumee River's mouth, Lake Erie plankton, Blue = South Bass Island, Lake Erie plankton. (b) Network showing relationships among zebra mussel (ZM; *Dreissena polymorpha*) and quagga mussel (QM; *D. rostriformis*) haplotypes for the mtDNA COI gene region used for our COIA assay. Hatch marks designate the number of nucleotide differences. Circle sizes represent the proportion of sequence reads attributed to that haplotype per sample location. Black circles denote haplotypes from NIH NCBI GenBank that were not found in our environmental samples



effectively accomplished through analysis of their abundant planktonic larvae, which were previously indistinguishable at the species and population levels. Thus, our assay can facilitate effective and accurate detection, and monitoring, of dreissenid communities via the simultaneous metabarcoding of thousands of veligers from plankton.

Although our assay solely was used here for zebra and quagga mussel populations, it could further be applied to distinguish the presence and species identities of the four other *Dreissena* species, as well as the two related—and also highly invasive—*Mytilopsis* species (*M. leucophaeata* and *M. sallei*). In marine and estuarine harbors worldwide, invasive *Mytilopsis* spp. exert similar ecological and economic fouling problems as do zebra and quagga mussels (Laine, Mattila, & Lehikoinen, 2006; Sousa, Gutiérrez, & Aldridge, 2009). In addition to the well-known zebra and quagga mussels, two *Dreissena* species native to the Balkans (*D. blanci* and *D. carinata*) recently expanded their ranges and may become invasive throughout Europe (Wilke et al., 2010). Furthermore, *D. anatolica* is endemic to just a few lakes in Turkey (Marshall & Stepien, 2019; Stepien et al., 2013; van der Velde et al., 2010) and therefore is at risk from competition pressures by future expansions of congeneric invaders. Our assay thus has apparent wide application to track and monitor the distributional trajectories of these taxa, by providing both species identification and their respective intra-specific population data.

Due to the much larger numbers of individuals surveyed (thousands vs. ~100), rarer haplotypes are more likely to be analyzed using HTS than with conventional sampling and sequencing of separate individuals. In our results, the rare haplotypes identified from surface and benthic collections at the same location sometimes differed; we thus recommend that future eDNA studies survey a variety of depths, with triplicate samples, to facilitate thorough evaluation of haplotypic diversity. Samples taken from different depths could be combined during DNA extraction, thereby yielding an enhanced representation of the entire water column, increasing the likelihood of surveying rare haplotypes. Site occupancy models also could be employed to evaluate possible failure to detect rare haplotypic variants (Dorazio & Erickson, 2017).

All haplotypes identified here for the zebra and quagga mussels are previously known sequences deposited in GenBank. This produces strong evidence that our screening procedure successfully removed any erroneous haplotypes resulting from possible PCR/sequencing error. Haplotype ZM-C previously was identified only in the Mohawk River, NY, a tributary of the Hudson River (Molloy et al., 2010), but here occurred in all of our Hudson River sites (plankton and eDNA), as well as in all but one of our Lake Erie plankton samples. ZM-G previously was described from the Dnieper River (a native habitat; Soroka, 2010) and Lake Superior in the Great Lakes (Grigorovich, Kelly, Darling, & West, 2008), yet appears to be widespread, since we found it in Lake Erie and throughout the Hudson River. Previous investigations described haplotypes QM-G as being widespread in the Dnieper River (where the quagga mussel natively occurs) and throughout

its invasive range in Europe (Romania, Germany, France, and The Netherlands), as well as in Lake Huron and Lake Cayuga, NY (Marescaux et al., 2016). That haplotype was not previously known from Lake Erie, where Marescaux et al. (2016) surveyed just ten individuals. The veliger larvae collection and HTS analysis approach used here is much more likely to elucidate the presence of rare haplotypes, such as these, due to the tremendous number of larvae sampled and analyzed.

#### 4.1 | The present approach versus other eDNA techniques

Some previous studies found that qPCR is a useful tool for estimating the number of dreissenid mussel sequence copies in a plankton or environmental sample (Amberg et al., 2019; De Ventura et al., 2017; Ram et al., 2011). However, those each required two separate markers (one for zebra mussels and another for quagga mussels) and neither discerned population genetic variability. Ram et al. (2011) used mt16S RNA for zebra mussel and mtDNA COI for quagga mussel, with the two targeted regions differing by 181 nt in length. Amberg et al. (2019) designed COI primers for the zebra mussel, with their future aspirations stated to develop a probe to differentiate quagga mussel eDNA. De Ventura et al. (2017) employed two overlapping regions of mtDNA COI, which had different primers and differed in length by 100 nt. Thus, their respective targeted regions would have differential amplification, and could degrade at different rates in the environment, rendering the two products of the two species not directly comparable. This likely would lead to inaccurate estimates of their relative abundances (see Dejean et al., 2011) in both methods. Two other qPCR assays for dreissenids (Gingera et al., 2017; Peñarrubia et al., 2016) lacked the ability to distinguish between the species. Egan et al. (2015) used light transmission spectroscopy to test water from the hulls of ships from Great Lakes harbors, and were able to detect eDNA from the quagga mussel, but not the zebra mussel, from localities where both were known to be present. Our approach overcomes such issues by using a higher-resolution HTS metabarcoding approach, amplifying the same gene region for both species.

The present targeted metabarcoding HTS assay yields important intraspecific population-level information for dreissenid mussels, beyond the scope of the earlier published methods. For example, the Klymus et al. (2017) 16S assay accurately differentiates between the two species, but that gene region evolves too slowly to provide population genetic resolution. Our study results here discerned haplotypic diversity of zebra and quagga mussels that was undetected by conventional sequencing of adults (Marescaux et al., 2016; Stepien et al., 2013). Similar intraspecific approaches using the mtDNA control region have been developed for the whale shark *Rhincodon typus* (Sigsgaard et al., 2016) and harbor porpoise *Phocoena phocoena* (Parsons et al., 2018), and with the cytochrome *b* region for the invasive silver carp *Hypophthalmichthys molitrix* (Stepien et al., 2019). Our assay is designed to accurately estimate haplotypic frequencies, as well as relative species proportions, which will be useful for

elucidating invasion pathways, along with documenting population relationships across temporal and spatial scales.

In summary, this diagnostic approach identifies and determines the relative representation of closely related species and their populations from plankton and eDNA, facilitating community- and population-level detection and ecological comparisons. The new targeted HTS metabarcoding dreissenid assay yields key information about the relative proportions of inter- and intra-specific variation. Tremendous numbers of individuals can be rapidly and simultaneously assessed, discerning and quantifying the relative abundances of species and their population-level haplotypic diversity among locations. Present-day morphological techniques for early detection of dreissenid invasions from plankton have been deemed insufficient for western U.S. reservoirs (Counihan & Bollens, 2017), with eDNA suggested as a useful alternative (Hosler, 2017). In comparison, our method, which jointly evaluates inter- and intra-specific diversity of plankton and eDNA water samples, allows for analysis of the temporal and spatial trajectories of dreissenid invasions, including potential displacements of zebra mussel populations by quagga mussels. Our assay and approach thus show broad and widespread application as an effective monitoring and research tool for understanding biological invasions and resultant ecological changes.

#### ACKNOWLEDGMENTS

This is contribution 4743 from NOAA Pacific Marine Environmental Laboratory (PMEL). The research was funded by the USEPA Great Lakes Restoration Initiative GL-00E01289, for "Invasive Invertebrate Species Prevention, Detection, and Control: A New Next Generation Sequencing Assay" and GL-00E01898, for "Invasive Species Prevention of fish and mollusks from bait, outfitter, and pet stores using metagenetic methods, supply chain analyses, and public education" to C.A.S. The laboratory of C.A.S. (Genetics & Genomics Group) relocated from the University of Toledo to NOAA PMEL and the University of Washington Joint Institute for Science of the Atmosphere and Ocean (JISAO) in October 2016, and funds were subawarded to the latter. N.T.M. additionally was partially supported by the Hudson River Foundation Bain Fellowship and the International Association for Great Lakes Research scholarship. We thank J. Hageman, J. Chaffin, J. Ross, M. Balcer, and the Cary Institute of Ecosystem Studies (D. Strayer, H. Malcom, and D. Fischer) for sample collections. We especially appreciate logistic help provided by S. McBride, as well as from T. Ackerman, G. Andres, F. Averick, D. Butterfield, F. Calzonetti, K. Czajkowski, T. Fisher, D. Green, J. Guyton, A. Izzi, D. Moorhead, C. Sabine, and W.V. Sigler. We also thank present and past members of the Genetics & Genomics Group headed by C.A.S. (K. Andrews, A. Elz, K. Klymus, E. Kramer, and M. Snyder) for laboratory and/or editorial assistance, and M. Neilson for assistance in obtaining photographs from USGS of zebra and quagga mussels.

#### CONFLICT OF INTEREST

None Declared.

#### AUTHOR CONTRIBUTIONS

C.A.S. conceived the study, wrote the grant, and obtained the funding. N.T.M. conducted sample collection, laboratory analyses, and developed the figures and tables. Data analysis, manuscript drafting, and manuscript revising and editing were done together by N.T.M. and C.A.S.

#### DATA ACCESSIBILITY

New haplotypes discerned in the present study are deposited in NCBI GenBank (accession numbers MK358469–70). Raw sequence data are provided in the NCBI GenBank repository under BioProject PRJNA513276 with accession numbers SAMN10699433–SAMN10699476.

#### ORCID

Carol A. Stepien  <https://orcid.org/0000-0002-5544-4333>

#### REFERENCES

- Ackerman, J. D., Sim, B., Nichols, S. J., & Claudi, R. (1994). A review of the early life history of zebra mussels (*Dreissena polymorpha*): Comparisons with marine bivalves. *Canadian Journal of Zoology*, *72*, 1169–1179. <https://doi.org/10.1139/z94-157>
- Amberg, J. J., Merkes, C. M., Stott, W., Rees, C. B., & Erickson, R. A. (2019). Environmental DNA as a tool to help inform zebra mussel, *Dreissena polymorpha*, management in inland lakes. *Management of Biological Invasions*, *10*. [http://www.reabic.net/journals/mbi/2019/Accepted/MBI\\_2019\\_Amberg\\_et\\_al\\_correctedproof.pdf](http://www.reabic.net/journals/mbi/2019/Accepted/MBI_2019_Amberg_et_al_correctedproof.pdf)
- Benson, A. J. (2013). Chronological history of zebra and quagga mussels (*Dreissenidae*) in North America, 1988–2010. In T. F. Nalepa, & D. W. Schloesser (Eds.), *Quagga and zebra mussels: Biology, impacts, and control* (2nd ed., pp. 9–31). Boca Raton, FL: CRC Press.
- Borcherding, J. (1991). The annual reproductive cycle of the fresh water mussel *Dreissena polymorpha* Pallas in lakes. *Oecologia*, *87*, 208–218. <https://doi.org/10.1007/BF00325258>
- Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: A unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, *16*, 176–182. <https://doi.org/10.1111/1755-0998.12428>
- Brown, J. E., & Stepien, C. A. (2010). Population genetic history of the dreissenid mussel invasions: Expansion patterns across North America. *Biological Invasions*, *12*, 3687–3710. <https://doi.org/10.1007/s10530-010-9763-2>
- Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME Journal*, *11*, 2639–2643. <https://doi.org/10.1038/ismej.2017.119>
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*, 581–583. <https://doi.org/10.1038/nmeth.3869>
- Claxton, W. T., & Boulding, E. G. (1998). A new molecular technique for identifying field collections of zebra mussel (*Dreissena polymorpha*) and quagga mussel (*Dreissena bugensis*) veliger larvae applied to eastern Lake Erie, Lake Ontario, and Lake Simcoe. *Canadian Journal of Zoology*, *76*, 194–198. <https://doi.org/10.1139/z97-175>

- Collas, F. P. L., Koopman, K. R., Hendriks, A. J., van der Velde, G., Verbrugge, L. N. H., & Leuven, R. S. E. W. (2014). Effects of desiccation on native and non-native molluscs in rivers. *Freshwater Biology*, 59, 41–55. <https://doi.org/10.1111/fwb.12244>
- Conroy, J. D., Edwards, W. J., Pontius, R. A., Kane, D. D., Zhang, H., Shea, J. F., ... Culver, D. A. (2005). Soluble nitrogen and phosphorus excretion of exotic freshwater mussels (*Dreissena* spp.): Potential impacts for nutrient remineralisation in western Lake Erie. *Freshwater Biology*, 50, 1146–1162. <https://doi.org/10.1111/j.1365-2427.2005.01392.x>
- Counihan, T. D., & Bollens, S. M. (2017). Early detection monitoring for larval dreissenid mussels: How much plankton sampling is enough? *Environmental Monitoring and Assessment*, 189, 98–112. <https://doi.org/10.1007/s10661-016-5737-x>
- Darling, J. A., & Mahon, A. R. (2011). From molecules to management: Adopting DNA-based methods for monitoring biological invasions in aquatic environments. *Environmental Resources*, 111, 978–988. <https://doi.org/10.1016/j.envres.2011.02.001>
- De Ventura, L., Kopp, K., Seppälä, K., & Jokela, J. (2017). Tracing the quagga mussel invasion along the Rhine river system using eDNA markers: Early detection and surveillance of invasive zebra and quagga mussels. *Management of Biological Invasions*, 8, 101–112. <https://doi.org/10.3391/mbi.2017.8.1.10>
- Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., & Miaud, C. (2011). Persistence of environmental DNA in freshwater ecosystems. *PLoS ONE*, 6, e23398. <https://doi.org/10.1371/journal.pone.0023398>
- DFO Department of Fisheries and Oceans, Canada (2018). Aquatic invasive species. <http://dfo-mpo.gc.ca/species-especes/ais-eae/index-eng.html>
- Diggins, T. P. (2001). A seasonal comparison of suspended sediment filtration by quagga (*Dreissena bugensis*) and zebra (*D. polymorpha*) mussels. *Journal of Great Lakes Research*, 27, 457–466. [https://doi.org/10.1016/S0380-1330\(01\)70660-0](https://doi.org/10.1016/S0380-1330(01)70660-0)
- Dorazio, R. M., & Erickson, R. A. (2017). eDNAoccupancy: An R package for multiscale occupancy modeling of environmental DNA data. *Molecular Ecology Resources*, 18, 368–380. <https://doi.org/10.1111/1755-0998.12735>
- Egan, S. P., Grey, E., Olds, B., Feder, J. L., Ruggiero, S. T., Tanner, C. E., & Lodge, D. M. (2015). Rapid molecular detection of invasive species in ballast and harbor water by integrating environmental DNA and light transmission spectroscopy. *Environmental Science & Technology*, 49(7), 4113–4121. <https://doi.org/10.1021/es5058659>
- Elbrecht, V., Vamos, E. E., Steinke, D., & Leese, F. (2018). Estimating intra-specific genetic diversity from community DNA metabarcoding data. *PeerJ*, 6, e4644. <https://doi.org/10.7717/peerj.4644>
- Evans, N. T., Olds, B. P., Renshaw, M. A., Turner, C. R., Li, Y., Jerde, C. L., ... Lodge, D. M. (2016). Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Molecular Ecology Resources*, 16, 29–41. <https://doi.org/10.1111/1755-0998.12433>
- Fadrosh, D. W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R. M., & Ravel, J. (2014). An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome*, 2, 6–13. <https://doi.org/10.1186/2049-2618-2-6>
- Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters*, 4, 423–425. <https://doi.org/10.1098/rsbl.2008.0118>
- Folmer, O. M., Black, W., Hoeh, R., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3, 294–299.
- Gallardo, B., Clavero, M., Sánchez, M. I., & Vilà, M. (2016). Global ecological impacts of invasive species in aquatic ecosystems. *Global Change Biology*, 22, 151–163. <https://doi.org/10.1111/gcb.13004>
- Gingera, T. D., Bajno, R., Docker, M. F., & Reist, J. D. (2017). Environmental DNA as a detection tool for zebra mussels *Dreissena polymorpha* (Pallas, 1771) at the forefront of an invasion event in Lake Winnipeg, Manitoba, Canada. *Management of Biological Invasions*, 8, 287–300. <https://doi.org/10.3391/mbi.2017.8.3.03>
- Grigorovich, I. A., Kelly, J. R., Darling, J. A., & West, C. W. (2008). The quagga mussel invades the Lake Superior basin. *Journal of Great Lakes Research*, 34, 342–350. [https://doi.org/10.3394/0380-1330\(2008\)34\[342:TQMITL\]2.0.CO;2](https://doi.org/10.3394/0380-1330(2008)34[342:TQMITL]2.0.CO;2)
- Hebert, P. D. N., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London*, 270, 313–321. <https://doi.org/10.1098/rspb.2002.2218>
- Hebert, P. D. N., Muncaster, B. W., & Mackie, G. L. (1989). Ecological and genetic studies on *Dreissena polymorpha* (Pallas): A new mollusc in the Great Lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, 46, 1587–1591. <https://doi.org/10.1139/f89-202>
- Heiler, K. C., Bij de Vaate, A., Ekschmitt, K., von Oheimb, P. V., Albrecht, C., & Wilke, T. (2013). Reconstruction of the early invasion history of the quagga mussel (*Dreissena rostriformis bugensis*) in Western Europe. *Aquatic Invasions*, 8, 53–57. <https://doi.org/10.3391/ai.2013.8.1.06>
- Hosler, D. M. (2011). Early detection of dreissenid species: Zebra/quagga mussels in water systems. *Aquatic Invasions*, 2, 217–222. <https://doi.org/10.3391/ai.2011.6.2.1>
- Hosler, D. M. (2017). Where is the body? Dreissenid mussels, raw water testing, and the real value of environmental DNA. *Management of Biological Invasions*, 8, 335–341. <https://doi.org/10.3391/mbi.2017.8.3.07>
- Jerde, C. L., Mahon, A. R., Chadderton, W. L., & Lodge, D. M. (2011). “Sight-unseen” detection of rare aquatic species using environmental DNA. *Conservation Letters*, 4, 150–157. <https://doi.org/10.1111/j.1755-263X.2010.00158.x>
- Johnson, L. E. (1995). Enhanced early detection an enumeration of zebra mussel (*Dreissena* spp.) veligers using cross-polarized light microscopy. *Hydrobiologia*, 312, 139–146. <https://doi.org/10.1007/BF00020769>
- Karatayev, A. Y., Burlakova, L. E., Pennuto, C., Ciborowski, J., Karatayev, V. A., Jutte, P., & Clapsadl, M. (2014). Twenty five years of changes in *Dreissena* spp. populations in Lake Erie. *Journal of Great Lakes Research*, 40, 550–559. <https://doi.org/10.1016/j.jglr.2014.04.010>
- Keller, R. P., Lodge, D. M., Lewis, M. A., & Shogren, J. F. (2009). *Bioeconomics of invasive species: Integrating ecology, economics, policy, and management*. Oxford: Oxford University Press.
- Klymus, K. E., Marshall, N. T., & Stepien, C. A. (2017). Environmental DNA (eDNA) metabarcoding assays to detect invasive invertebrate species in the Great Lakes. *PLoS ONE*, 12, e0177643. <https://doi.org/10.1371/journal.pone.0177643>
- Krjutškov, K., Andreson, R., Magi, R., Nikopensius, T., Khrunin, A., Mihailov, E., ... Metspalu, A. (2008). Development of a single tube 640-plex genotyping method for detection of nucleic acid variations on microarrays. *Nucleic Acids Research*, 36, e75. <https://doi.org/10.1093/nar/gkn357>
- Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29, 1–27. <https://doi.org/10.1007/BF02289565>
- Kulczynski, S. (1927). Die Pflanzenassoziationen der Pienenen. *Bulletin International De L'academie Polonaise Des Sciences Et Des Letters, Classe Des Sciences Mathematiques Et Naturelles. Serie B, Supplement II*, 2, 57–203.
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular*



- Biology and Evolution*, 33, 1870–1874. <https://doi.org/10.1093/molbev/msw054>
- Kvach, Y., & Stepien, C. A. (2008). Metazoan parasites of introduced round and tubenose gobies in the Great Lakes: Support for the “enemy release hypothesis”. *Journal of Great Lakes Research*, 34, 23–35. [https://doi.org/10.3394/O380-1330\(2008\)34\[23:MPOIRA\]2.CO;2](https://doi.org/10.3394/O380-1330(2008)34[23:MPOIRA]2.CO;2)
- Laine, A. O., Mattila, J., & Lehtikoinen, A. (2006). First record of the brackish water dreissenid bivalve *Mytilopsis leucophaeata* in the northern Baltic Sea. *Aquatic Invasions*, 1, 38–41. <https://doi.org/10.3391/ai.2006.1.1.9>
- Marescaux, J., von Oheimb, K. C., Etoundi, E., von Oheimb, P. V., Albrecht, C., Wilke, T., & Van Doninck, K. (2016). Unraveling the invasion pathways of the quagga mussel (*Dreissena rostriformis*) into Western Europe. *Biological Invasions*, 18, 245–264. <https://doi.org/10.1007/s10530-015-1005-1>
- Marshall, N. T., & Stepien, C. A. (2019). The family Dreissenidae. In K. Cummings, & C. Lydeard (Eds.), *Freshwater mollusk families of the world*. Baltimore, MD: John Hopkins Press.
- May, B., & Marsden, J. E. (1992). Genetic identification and implications of another invasive species of dreissenid mussel in the Great Lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, 49, 1501–1506. <https://doi.org/10.1139/f92-166>
- MNRF Ontario's Ministry of Natural Resources and Forests (2018). Ontario's invading species awareness program. <http://www.invasingspecies.com>
- Molloy, D. P., Giamberini, L., Burlakova, L. E., Karatayev, A. Y., Cryan, J. R., Trajanovski, S. L., & Trajanovska, S. P. (2010). Investigation of the endosymbionts of *Dreissena stankovici* with morphological and molecular confirmation of host species. In G. van der Velde, S. Rajagopal, & A. Bij de Vaate (Eds.), *The zebra mussel in Europe* (pp. 227–237). Kerkwerve: Backhuys Publishers.
- Nalepa, T. F., Fanslow, D. L., & Lang, G. A. (2009). Transformation of the offshore benthic community in Lake Michigan: Recent shift from the native amphipod *Diporeia* spp. to the invasive mussel *Dreissena rostriformis bugensis*. *Freshwater Biology*, 54, 466–479. <https://doi.org/10.1111/j.1365-2427.2008.02123.x>
- Nichols, S. J., & Black, M. G. (1994). Identification of larvae: The zebra mussel (*Dreissena polymorpha*), quagga mussel (*Dreissena rostriformis bugensis*), and Asian clam (*Corbicula fluminea*). *Canadian Journal of Zoology*, 72, 406–417. <https://doi.org/10.1139/z94-057>
- NOAA National Oceanic and Atmospheric Association (2018). Great Lakes Aquatic Nonindigenous Species Information System (GLANSIS). <https://www.glerl.noaa.gov/res/Programs/glanis/glan-sis.html>
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlenn, D., ... Wagner, H. (2018). Package ‘Vegan’. *Community ecology package, version, 2(9)*. <https://cran.r-project.org>, <https://github.com/vegandevs/vegan>
- Parsons, K. M., Everett, M., Dahlheim, M., & Park, L. (2018). Water, water everywhere: Environmental DNA can unlock population structure in elusive marine species. *Royal Society Open Science*, 5, 180537. <https://doi.org/10.1098/rsos.180537>
- Peñarrubia, L., Alcaraz, C., Bij de Vaate, A., Sanz, N., Pla, C., Vidal, O., & Viñas, J. (2016). Validated methodology for quantifying infestation levels of dreissenid mussels in environmental DNA (eDNA) samples. *Scientific Reports*, 6, (1–9). <https://doi.org/10.1038/srep39067>
- Pimentel, D. (2014). *Biological invasions: Economic and environmental costs of alien plant, animal, and microbe species*. Boca Raton, FL: CRC Press.
- Pimentel, D., Zuniga, R., & Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*, 52, 273–288. <https://doi.org/10.1016/j.ecolecon.2004.10.002>
- R Core Team (2015). *R: A language for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Radulovici, A. E., Archambault, P., & Dufresne, F. (2010). DNA barcodes for marine biodiversity: Moving fast forward? *Diversity*, 2, 450–472. <https://doi.org/10.3390/d2040450>
- Ram, J. L., Karim, A. S., Acharya, P., Jagtap, P., Purohit, S., & Kashian, D. R. (2011). Reproduction and genetic detection of veligers in changing *Dreissena* populations in the Great Lakes. *Ecosphere*, 2, 1–16. <https://doi.org/10.1890/ES10-00118.1>
- Robinson, A., Burgman, M. A., & Cannon, R. (2011). Allocating surveillance resources to reduce ecological invasions: Maximizing detections and information about the threat. *Ecological Applications*, 21, 1410–1417. <https://doi.org/10.1890/10-0195.1>
- Sigsgaard, E. E., Nielsen, I. B., Bach, S. S., Lorenzen, E. D., Robinson, D. P., Knudsen, S. W., ... Møller, P. R., & Thomsen, F., (2016). Population characteristics of a large whale shark aggregation inferred from seawater environmental DNA. *Nature Ecology and Evolution*, 1, 1–4. <https://doi.org/10.1038/s41559-016-0004>
- Simberloff, D. (2013). *Invasive species: What everyone needs to know*. Oxford: Oxford University Press.
- Sokal, R. R., & Rohlf, F. J. (1994). *Biometry: The principles and practice of statistics in biological research* (3rd ed.). New York, NY: W. H. Freeman and Company.
- Soroka, M. (2010). Genetic differentiation of *Dreissena polymorpha* from East-European countries. In G. van der Velde, S. Rajagopal, & A. Bij de Vaate (Eds.), *The zebra mussel in Europe* (pp. 133–144). Kerkwerve: Backhuys Publishers.
- Sousa, R., Gutiérrez, J. L., & Aldridge, D. C. (2009). Non-indigenous invasive bivalves as ecosystem engineers. *Biological Invasions*, 11, 2367–2385. <https://doi.org/10.1007/s10530-009-9422-7>
- Sprung, M. (1993). The other life: An account of present knowledge of the larval phase of *Dreissena polymorpha*. In T. F. Nalepa, & D. W. Schloesser (Eds.), *Zebra mussels: Biology, impacts, and control* (1st ed., pp. 39–53). Boca Raton, FL: CRC Press.
- Stepien, C. A., Grigorovich, I. A., Gray, M. A., Sullivan, T. J., Yerga-Woolwine, S., & Kalayci, G. (2013). Evolutionary, biogeographic, and population genetic relationships of dreissenid mussels, with revision of component taxa. In T. F. Nalepa, & D. W. Schloesser (Eds.), *Quagga and zebra mussels: Biology, impacts, and control* (2nd ed., pp. 403–444). Boca Raton, FL: CRC Press.
- Stepien, C. A., Hubers, A. N., & Skidmore, J. L. (1999). Diagnostic genetic markers and evolutionary relationships among invasive dreissenid and corbiculoid bivalves in North America: Phylogenetic signal from mitochondrial 16S rDNA. *Molecular Phylogenetics and Evolution*, 13, 31–49. <https://doi.org/10.1006/mpev.1999.0666>
- Stepien, C. A., Snyder, M. R., & Elz, A. E. (2019). Invasion genetics of the silver carp *Hypophthalmichthys molitrix* across North America: Differentiation of fronts, introgression, and eDNA metabarcode detection. *PLoS ONE*, in press.
- Strayer, D. L., & Malcom, H. M. (2013). Long-term Change in the Hudson River's Bivalve Populations. In T. F. Nalepa, & D. W. Schloesser (Eds.), *Quagga and zebra mussels: Biology, impacts, and control* (2nd ed., pp. 71–81). Boca Raton, FL: CRC Press.
- Strayer, D. L., Powell, J., Ambrose, P., Smith, L. C., Pace, M. L., & Fischer, D. T. (1996). Arrival, spread, and early dynamics of a zebra mussel (*Dreissena polymorpha*) population in the Hudson River estuary. *Canadian Journal of Fisheries and Aquatic Sciences*, 53, 1143–1149. <https://doi.org/10.1139/f96-038>
- Takahara, T., Minamoto, T., Yamanaka, H., Doi, H., & Kawabata, Z. (2012). Estimation of fish biomass using environmental DNA. *PLoS ONE*, 7, e35868. <https://doi.org/10.1371/journal.pone.0035868>
- Thomsen, P., Kielgast, J. O. S., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., ... Willerslev, E. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology*, 21, 2565–2573. <https://doi.org/10.1111/j.1365-294X.2011.05418.x>
- Turner, C. R., Barnes, M. A., Xu, C. C., Jones, S. E., Jerde, C. L., & Lodge, D. M. (2014). Particle size distribution and optimal capture of aqueous

- microbial eDNA. *Methods in Ecology and Evolution*, 5, 676–684. <https://doi.org/10.1111/2041-210X.12206>
- USGS United States Geological Survey (2018). *Nonindigenous aquatic species (NAS) information*. <https://www.nas.er.usgs.gov/html>
- van der Velde, G., Rajagopal, S., & Bij de Vaate, A. (2010). From zebra mussels to quagga mussels: An introduction to the Dreissenidae. In G. van der Velde, S. Rajagopal, & A. Bij de Vaate (Eds.), *The zebra mussel in Europe* (pp. 1–10). Kerkwerve: Backhuys Publishers.
- Vanderploeg, H. A., Liebig, J. R., Carmichael, W. W., Agy, M. A., Johengen, T. H., Fahnenstiel, G. L., & Nalepa, T. F. (2001). Zebra mussel (*Dreissena polymorpha*) selective filtration promoted toxic *Microcystis* blooms in Saginaw Bay (Lake Huron) and Lake Erie. *Canadian Journal of Fisheries and Aquatic Sciences*, 58, 1208–1221. <https://doi.org/10.1139/f01-066>
- Ward, J. M., & Ricciardi, A. (2007). Impacts of *Dreissena* invasions on benthic macroinvertebrate communities: A meta-analysis. *Diversity and Distributions*, 13, 155–165. <https://doi.org/10.1111/j.1472-4642.2007.00336.x>
- Wilke, T., Schultheiß, R., Albrecht, C., Bornmann, N., Trajanovski, S., & Kevrekidis, T. (2010). Native *Dreissena* freshwater mussels in the Balkans: In and out of ancient lakes. *Biogeosciences*, 7, 3051–3065. <https://doi.org/10.5194/bg-7-3051-2010>
- Wu, L., Wen, C., Qin, Y., Yin, H., Tu, Q., Van Nostrand, J. D., ... Zhou, J. (2015). Phasing amplicon sequencing on Illumina Miseq for robust environmental microbial community analysis. *BMC Microbiology*, 15, 125. <https://doi.org/10.1186/s12866-015-0450-4>
- Zaiko, A., Minchin, D., & Olenin, S. (2014). 'The day after tomorrow': Anatomy of an 'r' strategist aquatic invasion. *Aquatic Invasions*, 9, 145–155. <https://doi.org/10.3391/ai.2014.9.2.03>
- Zhulidov, A. V., Kozhara, A. V., Scherbina, G. H., Nalepa, T. F., Protasov, A., Afanasiev, S. A., ... Pavlov, D. F. (2010). Invasion history, distribution, and relative abundances of *Dreissena bugensis* in the Old World: A synthesis of data. *Biological Invasions*, 12, 1923–1940. <https://doi.org/10.1007/s10530-009-964-y>

**How to cite this article:** Marshall NT, Stepien CA. Invasion genetics from eDNA and thousands of larvae: A targeted metabarcoding assay that distinguishes species and population variation of zebra and quagga mussels. *Ecol Evol.* 2019;9:3515–3538. <https://doi.org/10.1002/ece3.4985>

## APPENDIX A

**TABLE A1** Single nucleotide polymorphisms (SNPs) that differentiate zebra mussel (ZM; *Dreissena polymorpha*) and quagga mussel (QM; *D. rostriformis*) haplotypes (lettered) for the COIA region (from NIH NCBI GenBank and our records). Haplotypes used for our mock communities are in **bold italics**. Primer regions are shaded grey.

Species-Haplotype	GenBank Representative	1	26
<b>ZM-A,M-T</b>	KX231273	A G T G T T Y T K A T T C G T T T R G A G C T W A G K G C	A C C T G G A A G C G T C C T T G G T G A T T T T C
ZM-B	JQ435817		C
<b>ZM-C</b>	DQ333701		
ZM-D	U47653		
ZM-E,K	EU484433		
ZM-F,L	DQ840124		
<b>ZM-G</b>	EU484448		
ZM-H	AM748997		
ZM-I	DQ840126		
ZM-J	AM748996		
<b>QM-A,I-Q</b>	KJ881409		T . . . . . A . . . . . A T . . . . . A . . . . . A . . . . .
QM-B	U47650		T . . . . . A . . . . . A . . . . . A T . . . . . A . . . . .
QM-C	MK358469		T . . . . . A . . . . . A G . . . . . A . . . . . A . . . . .
QM-D	AF510507		T . . . . . A . . . . . A . . . . . A T . . . . . A . . . . .
QM-E	KJ881417		T . . . . . A . . . . . A T . . . . . A . . . . . A . . . . .
<b>QM-F</b>	KJ881410		T . . . . . A . . . . . A T . . . . . A . . . . . A . . . . .
<b>QM-G</b>	KJ881412		T . . . . . A . . . . . A T . . . . . A . . . . . A . . . . .
QM-H	MK358470		T . . . . . A . . . . . A T . . . . . A . . . . . A . . . . .
		27	81
<b>ZM-A,M-T</b>	KX231273	A T T A T A T A A T T T A A T T G T C A C C A C T C A T G G G C T T G T T A T A A T T T T T T T T T C T A G T A	
ZM-B	JQ435817		G
<b>ZM-C</b>	DQ333701		
ZM-D	U47653		C
ZM-E,K	EU484433		
ZM-F,L	DQ840124		
<b>ZM-G</b>	EU484448		
ZM-H	AM748997		
ZM-I	DQ840126		
ZM-J	AM748996		
<b>QM-A,I-Q</b>	KJ881409		T . . . . . T . . . . . C . . . . . A . . . . . C . . . . . T . . . . .
QM-B	U47650		T . . . . . T . . . . . C . . . . . A . . . . . C . . . . . T . . . . .
QM-C	MK358469		T . . . . . T . . . . . C . . . . . A . . . . . C . . . . . T . . . . .
QM-D	AF510507		C . . . . . T . . . . . T . . . . . C . . . . . A . . . . . C . . . . . T . . . . .
QM-E	KJ881417		C T . . . . . T . . . . . C . . . . . A . . . . . C . . . . . T . . . . .
<b>QM-F</b>	KJ881410		T . . . . . T . . . . . C . . . . . A . . . . . C . . . . . T . . . . .
<b>QM-G</b>	KJ881412		T . . . . . T . . . . . C . . . . . A . . . . . C . . . . . T . . . . .
QM-H	MK358470		T . . . . . T . . . . . C . . . . . A . . . . . C . . . . . T . . . . .
		82	136
<b>ZM-A,M-T</b>	KX231273	A T A C C T A T A A T A A T G G G G G G A T T C G G A A A T T G A T T G G T A C C A A T A A T A C T G A G T C	
ZM-B	JQ435817		G
<b>ZM-C</b>	DQ333701		
ZM-D	U47653		C
ZM-E,K	EU484433		C
ZM-F,L	DQ840124		C
<b>ZM-G</b>	EU484448		
ZM-H	AM748997		A
ZM-I	DQ840126		
ZM-J	AM748996		
<b>QM-A,I-Q</b>	KJ881409		G . . . . . A . . . . . T . . . . . T . . . . . C . . . . . G . . . . . C . . . . . G . . . . . A . . . . .
QM-B	U47650		G . . . . . A . . . . . T . . . . . T . . . . . A . . . . . C . . . . . G . . . . . C . . . . . G . . . . . A . . . . .
QM-C	MK358469		G . . . . . A . . . . . T . . . . . T . . . . . C . . . . . G . . . . . C . . . . . G . . . . . A . . . . .
QM-D	AF510507		G . . . . . A . . . . . T . . . . . T . . . . . C . . . . . G . . . . . C . . . . . G . . . . . A . . . . .
QM-E	KJ881417		G . . . . . A . . . . . T . . . . . T . . . . . C . . . . . G . . . . . C . . . . . G . . . . . A . . . . .
<b>QM-F</b>	KJ881410		G . . . . . A . . . . . T . . . . . T . . . . . C . . . . . G . . . . . C . . . . . G . . . . . A . . . . .
<b>QM-G</b>	KJ881412		G . . . . . A . . . . . T . . . . . T . . . . . C . . . . . G . . . . . C . . . . . G . . . . . A . . . . .
QM-H	MK358470		G . . . . . A . . . . . T . . . . . T . . . . . C . . . . . G . . . . . C . . . . . G . . . . . A . . . . .
		137	169
<b>ZM-A,M-T</b>	KX231273	T C C T G A T A T A G G T T T C C C T C G T C T T A A T A A T	G T W A G W T T T T G G G T T Y T A C C T R T C
ZM-B	JQ435817		A
<b>ZM-C</b>	DQ333701		C
ZM-D	U47653		
ZM-E,K	EU484433		
ZM-F,L	DQ840124		
<b>ZM-G</b>	EU484448		C . . . . . T . . . . .
ZM-H	AM748997		C . . . . . T . . . . .
ZM-I	DQ840126		T . . . . .
ZM-J	AM748996		C . . . . .
<b>QM-A,I-Q</b>	KJ881409		T . . . . . A . . . . .
QM-B	U47650		T . . . . . A . . . . .
QM-C	MK358469		T . . . . . A . . . . .
QM-D	AF510507		T . . . . . A . . . . .
QM-E	KJ881417		T . . . . . A . . . . .
<b>QM-F</b>	KJ881410		T . . . . . A . . . . .
<b>QM-G</b>	KJ881412		T . . . . . A . . . . .
QM-H	MK358470		T . . . . . A . . . . .

**TABLE A2** Consensus sequences for six species of *Dreissena* and two closely related invasive *Mytilopsis* species for the COIA region from NIH NCBI GenBank and our records. Primer regions are shaded grey

Species	1	26
<i>Dreissena polymorpha</i>	<b>A G T G T T Y T K A T T C G T T T R G A G C T W A G K G C</b>	<b>A C C T G G A A G C G T C C T Y G G T G A T T T T C</b>
<i>D. anatolica</i>	.....	..... T .....
<i>D. blanci</i>	..... Y .....	T . . Y . . G . . . . . G . . . . . C . . . . .
<i>D. carinata</i>	..... Y M . Y . . . . . Y . . . . . R Y . . . . .	W . . Y . . G . . Y . . G . . . . . Y . . Y . . Y . .
<i>D. caputlacus</i>	..... C .....	T . . C . . . . . A . . A T . A . . G . . . . . A . .
<i>D. rostriformis</i>	.....	T . . R . . . . . A . . A K . A . . A . . . . . A . .
<i>Mytilopsis leucophaeata</i>	. G . . . . . A . . . . . A . . . . .	T . . . . . T . . T T . A . . A . . . . . A C . .
<i>M. sallei</i>	. G . . . . .	T . . . . . A . . T T . A . . G . . . . . A Y . .
	27	81
<i>Dreissena polymorpha</i>	A T T A T A T A A T T T A A T T G T C A C C A C T C A K G G G C T T G T T A T A A T T T T Y T T T C T A G T A	
<i>D. anatolica</i>	.....	..... T .....
<i>D. blanci</i>	. . . R . . . . . C . . T . . T . . . . . T . . . . . W . . . . . C . . T . . . . .	
<i>D. carinata</i>	. Y . . . . . R . Y . . Y . . Y . S . . . . . T . . . . . R . . . . . R . . . . . C . . T . . . . .	
<i>D. caputlacus</i>	. Y . G . . . . . C . R . . . . . T . . T . . . . . C . . . . . A . Y . . . . . T . . C . . T . . . . .	
<i>D. rostriformis</i>	. Y . . . . . Y Y . . . . . T . . . . . C . . . . . A . Y . . . . . T . . . . . T . . . . .	
<i>Mytilopsis leucophaeata</i>	T . . . . . G G . . . . . A . . A . G . . Y . . T T . A . . . . . T . . . . . T . . . . . T . . . . .	
<i>M. sallei</i>	T . . R . . . . . G R . . . . . K . . T . . . . . T . . T . G . . . . . T . . . . . T . . . . .	
	82	136
<i>Dreissena polymorpha</i>	A T A C C T A T M A T A A T G G G G G A T T S G G A A A Y T S A T T G R T A C C A A T A A T A C T G A G T C	
<i>D. anatolica</i>	..... A . . . . . A . . . . . C . . . . . T . G . . . . . G . . . . .	
<i>D. blanci</i>	..... A . . . . . A . . T . G . C . . . . . T . G G . . . . . G . T . . . . . T . A . . . . .	
<i>D. carinata</i>	. . R . . Y . . R . . . . . A . . Y . . G . . Y . . R . . . . . G R Y . R G . Y . . R . . R . . R . . A . . Y . .	
<i>D. caputlacus</i>	..... A . . . . . A . . C . . . . . T . . . . . C . G G . . . . . G . G . . . . . A . . . . .	
<i>D. rostriformis</i>	..... A . . G . . A . . T . . . . . T R . . . . . C . G G . . . . . G . C . . R . . . . . A . . . . .	
<i>Mytilopsis leucophaeata</i>	..... A . . G . . . . . T . . T . . T . . . . . T . G . . . . . A G . T . . . . . Y . A G C R G	
<i>M. sallei</i>	. . R . . . . . R . . G . . A . . T . . T . . T . . . . . C . G G Y . A G . T . . R . . R . . G T . A G C . G	
	137	169
<i>Dreissena polymorpha</i>	T C C T G A H A T A G G Y T T Y C C T C G Y C T T A A T A A T	<b>G T W A G W T T T T G G G T T Y T A C C T R T C</b>
<i>D. anatolica</i>	..... T . . . . . T . . T . . . . . T . . . . .	.....
<i>D. blanci</i>	..... T . . . . . D . . T . . . . .	..... Y
<i>D. carinata</i>	. . . Y . . T . . R . . G . . T . . Y . . T . . Y . . . . .	R . . . . . Y . . . . . Y
<i>D. caputlacus</i>	..... T . . . . . T . . T . . Y . . T T . A . . C . .	.....
<i>D. rostriformis</i>	..... T . . . . . T . . C . . . . . T T . A . . . . .	.....
<i>Mytilopsis leucophaeata</i>	..... T . . . . . A . . T . . . . . T T . A . . . . .	..... R . . . . . G . . . . . A
<i>M. sallei</i>	. . . Y . . T . . . . . K . . T . . . . . G . . . . .	..... A . . G . . G . . . . . T

**TABLE A3** Single nucleotide polymorphisms (SNPs) that differentiate zebra mussel (ZM; *Dreissena polymorpha*) and quagga mussel (QM; *D. rostriformis*) haplotypes (lettered) for the COIB region from NIH NCBI GenBank and our records. Haplotypes used for mock communities are in *bold italics*. Primer regions are shaded grey.

Species-Haplotype	GenBank Representative	1	35
ZM-A,C,E,P,T	KX231273	<b>GGATTYGGKGGKGGTTGAAC</b>	CTTATATCCTCCTTTATCTAGAGTTATAGGACATT
ZM-B,F,L,M	JQ435817		
ZM-G-I	EU484448		C
ZM-J	AM748996		
ZM-K	EU484435		
ZM-N	AM749001		
ZM-O	KY091877		
QM-A,E,F,G,H,J	KJ881409	A	A G C GA G G
QM-B,K	U47650	A	C A G C GA G G
QM-C	MK358469	A	G C GA G G
QM-D,I,N	AF510507	A	A G C GA G G
QM-L	JQ756297	A	A T G C GA G G
QM-M	JQ756298	A	G C GA G G
QM-O	KJ881415	A	A G C GA G G
QM-P	KJ881413	A	A G C GA G G
QM-Q	KJ881414	A	A G C GA G G

Species-Haplotype	GenBank Representative	36	90
ZM-A,C,E,P,T	KX231273	<b>CAGGCCCTGCGATAGATTTTTTGATTTTATCTCTTCATATTGGGGGAGCTTCTTC</b>	
ZM-B,F,L,M	JQ435817	G	
ZM-G-I	EU484448	G	
ZM-J	AM748996		
ZM-K	EU484435		
ZM-N	AM749001		
ZM-O	KY091877		
QM-A,E,F,G,H,J	KJ881409	G A C C A C A C T	
QM-B,K	U47650	G A C C A C A C T	
QM-C	MK358469	G A C C A C A C T	
QM-D,I,N	AF510507	G A C C A C A C T	
QM-L	JQ756297	G A C C A C A C T	
QM-M	JQ756298	G A C C A C A C T	
QM-O	KJ881415	G A C C A C A C T	
QM-P	KJ881413	G A C C A C A C T	
QM-Q	KJ881414	G A C C A C A C T	

Species-Haplotype	GenBank Representative	91	145
ZM-A,C,E,P,T	KX231273	<b>GATTATGGCTTCTATTAATTTTTATAGGACATGAGGTAATATACGTGCTGGATGT</b>	
ZM-B,F,L,M	JQ435817		
ZM-G-I	EU484448		
ZM-J	AM748996	G	
ZM-K	EU484435		
ZM-N	AM749001		C
ZM-O	KY091877		
QM-A,E,F,G,H,J	KJ881409	T A A A A A A C G	
QM-B,K	U47650	T A A A A A A C G	
QM-C	MK358469	T A A A A A A C G	
QM-D,I,N	AF510507	T A A A A A A C G	
QM-L	JQ756297	T A C A A A A C G	
QM-M	JQ756298	T A C A A A A C G	
QM-O	KJ881415	T A A A A A A C G	
QM-P	KJ881413	T A A A A A A C G	
QM-Q	KJ881414	T A A A A A A C G	

Species-Haplotype	GenBank Representative	146	175
ZM-A,C,E,P,T	KX231273	<b>CATCAATTTTACCGGGTCCCTTTGTTCTGCACWTCATTGGTGTKACMAGWTTY</b>	
ZM-B,F,L,M	JQ435817		
ZM-G-I	EU484448		
ZM-J	AM748996		
ZM-K	EU484435	A	
ZM-N	AM749001		
ZM-O	KY091877		T
QM-A,E,F,G,H,J	KJ881409	G T G T T T	
QM-B,K	U47650	G T G T T T	
QM-C	MK358469	G T G T T T	
QM-D,I,N	AF510507	G T G T T T	
QM-L	JQ756297	G T G T T T	
QM-M	JQ756298	G T G T T T	
QM-O	KJ881415	G T G T T T	
QM-P	KJ881413	G T G T T T	
QM-Q	KJ881414	G T G T T T	

**TABLE A4** Consensus sequences for six species of *Dreissena* and two closely related invasive *Mytilopsis* species for the COIB region from NIH NCBI GenBank and our records. Primer regions are shaded grey

Species	1	35
<i>Dreissena polymorpha</i>	<b>GGATTYGGKGGKGGTTGAAAC</b>	<b>CCTTATAYCCTCCTTTATCTAGAGTTATAGGACATT</b>
<i>D. anatolica</i>	. . . . .	. . . . . T . . . . . GA . . . . .
<i>D. blanci</i>	. . . . . <b>A</b> . . . . .	Y . R . T . C . . C . G . . . . . G . . . . . T . . . . .
<i>D. carinata</i>	. . . . . <b>R</b> . . . . . <b>Y</b> . . . . .	Y . R . . . . . Y . . . . . Y . . . . . R S . . . . . Y . . . . .
<i>D. caputlacus</i>	. . . . . <b>C</b> . . . . . <b>R</b> . . . . .	A Y . . . . . T . . . . . C . . . . . G . . . . . G . . . . .
<i>D. rostriformis</i>	. . . . .	A . . . . . A . Y G . . . . . C . GA . . . . . G . G . . . .
<i>Mytilopsis leucophaeata</i>	. . . . . <b>T</b> . . . . . <b>A</b> . . . . .	T . . . . . T . . . . . A . G . . . . . G . . . . . C .
<i>M. sallei</i>	. . . . .	. . . . . Y . . . . . T . R . . . . . R . T . A . . . . . G . C .
	36	90
<i>Dreissena polymorpha</i>	<b>CAGGSCCTGCGATAGATTTTTTGATTTTATCTCTTCATATTGGGGGAGCTTCTTC</b>	
<i>D. anatolica</i>	. . . . . G . . . . .	. . . . . K . . . . .
<i>D. blanci</i>	. . . . . A . G . . . . .	. . . . . A . . . . . R . . . . . R . . . . .
<i>D. carinata</i>	. . . . . R . R . R . . . . .	A . Y . . . . . Y . R . . . . . K . . . . . Y Y . Y . .
<i>D. caputlacus</i>	. . . . . G . . . . . C . . . . . C . . . . .	A . . . . . C . . . . . Y . . . . . R . . . . . C . C . .
<i>D. rostriformis</i>	. . . . . G . A . C . . . . .	Y . A . . . . . C . K . . . . . Y . . . . .
<i>Mytilopsis leucophaeata</i>	. T . A . . . . . A . G . Y . . . . .	A . . . . . A . T . G . . . . .
<i>M. sallei</i>	Y Y . . T . . . . . A . . . . .	R . . . . . G . . . . . T . G . . . . .
	91	145
<i>Dreissena polymorpha</i>	<b>GATTATGGCTTCKATTAATTTTTATAGGACATGAGGTAATATACGTGCYGGATGT</b>	
<i>D. anatolica</i>	. . . . . T . . . . .	. . . . . R . . . . . T . . . . .
<i>D. blanci</i>	A . . . . . T . . . . .	. . . . . R . . . . . Y . . . . . C . G . C
<i>D. carinata</i>	A . . . . . Y . . . . . Y . . . . .	R . R . . . . . R . . . . . Y . R . . . . . C . G . Y
<i>D. caputlacus</i>	T . . . . . A . S . T . . . . .	. . . . . A . . . . . R R . A . C . G . . . . . C . G . .
<i>D. rostriformis</i>	T . . . . . A . Y . T . . . . .	. . . . . R . . . . . A . . . . . S . C . G . .
<i>Mytilopsis leucophaeata</i>	T . . . . . A . . . . . T . . . . .	. . . . . A . T . . . . . A . . . . . T . G . .
<i>M. sallei</i>	A . . . . . K . T . . . . . Y . . . . .	. . . . . R . T . R . G . . . . . G . . . . . T . G . .
	146	175
<i>Dreissena polymorpha</i>	<b>CATCAATTTWACCGGTCCTTTGTTCTGY</b>	<b>ACWTCCTATTGGTGTKACMAGWTTY</b>
<i>D. anatolica</i>	. . . . . T . . . . .	. . . . . T . . . . .
<i>D. blanci</i>	. . . . . C . . . . . T . T . . . . . G . . . . . A . . . . .	. . . . . G . . . . . A . T . R . . . .
<i>D. carinata</i>	. . . . . R . . . . . Y T . Y . . . . . R . . . . . Y . . . . . R . . . . . Y . . . . . C . . . . .	. . . . . G . . . . . Y . . . . . R . . . . . Y . G . . . .
<i>D. caputlacus</i>	. . . . . G . . . . . T . Y . . . . . G . . . . . T . . . . .	. . . . . Y . . . . . C . . . . .
<i>D. rostriformis</i>	. . . . . G . . . . . T . T . . . . . R . . . . . T . T . . . . .	. . . . .
<i>Mytilopsis leucophaeata</i>	. . . . . G . . . . . T . T . . . . . A . . . . . A . . . . . T . T . . . . .	. . . . . C . R . . . . . A . Y . . . . .
<i>M. sallei</i>	. . . . . G . . . . . T . T . . . . . T . . . . . C . T . T . T . T . . . . .	. . . . . Y . . . . . A . . . . . G . . . . .

**TABLE A5** Number of reads, merged reads, and exact match trimmed reads (called) for mock community, plankton, and water samples. MC5A, MC2B, and MC7B were analyzed with the OBITools Python package, instead of the DADA2 pipeline, due to less sequencing depth

Sample	Marker	Type	Spacer	# Reads	# Merged	% Merged	# Called	% Called
MC 1A	COIA	Mock Community	E	173370	164241	94.73	143341	82.68
MC 2A	" "	" "	F	235448	225092	95.60	185391	78.74
MC 3A	" "	" "	G	254623	241128	94.70	205763	80.81
MC 4A	" "	" "	H	237268	228323	96.23	195985	82.60
MC 5A	" "	" "	None	15029	1520	10.11	1520	10.11
MC 6A	" "	" "	E	352913	308859	87.52	252268	71.48
MC 7A	" "	" "	F	120495	112876	93.68	98010	81.34
MC 8A	" "	" "	G	173457	155680	89.75	129662	74.75
MC 1B	COIB	Mock Community	H	477773	371366	77.73	332069	69.50
MC 2B	" "	" "	None	23769	5680	23.90	2367	9.96
MC 3B	" "	" "	E	345642	297879	86.18	246145	71.21
MC 4B	" "	" "	F	257624	227787	88.42	196872	76.42
MC 5B	" "	" "	G	408454	279678	68.47	245718	60.16
MC 6B	" "	" "	H	507302	379539	74.82	337146	66.46
MC 7B	" "	" "	None	14929	5075	33.99	2472	16.56
MC 8B	" "	" "	E	401220	332333	82.83	305703	76.19
SBI 1	COIA	Plankton	F	335700	327007	97.41	253259	75.44
SBI 2	" "	" "	E	253753	246079	96.98	206950	81.56
SBI 3	" "	" "	G	509518	486307	95.44	335896	65.92
SBI 3	Mol16S	Plankton	E	21790	15146	69.51	9972	45.76
SBI 4	COIA	Plankton	H	366913	355599	96.92	264612	72.12
SBI 5	" "	" "	E	242625	236097	97.31	186703	76.95
SBI 6	" "	" "	F	578672	563515	97.38	432889	74.81
SBI 7	" "	" "	G	376133	364953	97.03	251330	66.82
MRM 1	" "	" "	G	394627	381106	96.57	256764	65.06
MRM 2	" "	" "	H	256200	248145	96.86	188973	73.76
MRM 3	" "	" "	H	131387	125416	95.46	103257	78.59
MRM 3	Mol16S	Plankton	H	174169	159511	91.58	75791	47.51
MRM 4	COIA	Plankton	E	387123	375513	97.00	296571	76.61
StuyP	" "	" "	H	634443	533079	84.02	395934	62.41
NGermP	" "	" "	F	393192	384287	97.74	297889	75.76
NGermP	Mol16S	Plankton	F	32764	22557	68.85	16404	50.07
RhinP	COIA	Plankton	G	653035	584190	89.46	415135	63.57
PougP1	" "	" "	H	20351	19149	94.09	14521	71.35
PougP2	" "	" "	E	346763	334664	96.51	260406	75.10
PougP3	" "	" "	F	217612	213067	97.91	167888	77.15
StuyS	COIA	Water	F	273971	255911	93.41	197762	72.18
StuyB	" "	" "	E	268824	229668	85.43	180542	67.16
StuyB	Mol16S	Water	E	136140	124499	91.45	107323	78.83
CoxS	COIA	Water	F	529641	474330	89.56	367345	69.36
CoxB	" "	" "	E	615279	483397	78.57	383128	62.27
StockS	" "	" "	H	496502	418430	84.28	316129	63.67
StockB	" "	" "	G	396852	365499	92.10	255063	64.27
Aquarium 1	" "	" "	E	301851	277979	92.09	156520	51.85
Aquarium 2	" "	" "	F	399631	324482	81.20	192107	48.07
Aquarium 3	" "	" "	G	370290	328942	88.83	167948	45.36
Aquarium 4	" "	" "	H	310110	249129	80.34	150554	48.55
Aquarium 5	" "	" "	E	303221	276318	91.13	78501	25.89
Aquarium 6	" "	" "	F	290458	258414	88.97	148199	51.02
Aquarium 7	" "	" "	G	383593	326868	85.21	196116	51.13
Aquarium 8	" "	" "	H	429649	331568	77.17	212110	49.37
Aquarium 9	" "	" "	E	454504	420287	92.47	254968	56.10

