Quantifying early risks of species invasions: factors regulating south to north bivalve colonization of novel habitats

Zhiqiang Xia
University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation
https://scholar.uwindsor.ca/etd/7745

This online database contains the full-text of PhD dissertations and Masters’ theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.
Quantifying early risks of species invasions: factors regulating south to north bivalve colonization of novel habitats

By

Zhiqiang Xia

A Dissertation

Submitted to the Faculty of Graduate Studies through the Great Lakes Institute for Environmental Research in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy at the University of Windsor

Windsor, Ontario, Canada

2019

© 2019 Zhiqiang Xia
Quantifying early risks of species invasions: factors regulating south to north bivalve colonization of novel habitats

by

Zhiqiang Xia

APPROVED BY:

J. Bossenbroek, External Examiner
University of Toledo

S.J. Loeb
Department of Chemistry and Biochemistry

A. Zhan
Great Lakes Institute for Environmental Research

K. Drouillard
Great Lakes Institute for Environmental Research

H.J. MacIsaac, Co-Advisor
Great Lakes Institute for Environmental Research

G.D. Haffner, Co-Advisor
Great Lakes Institute for Environmental Research

May 31, 2019
DECLARATION OF CO-AUTHORSHIP / PREVIOUS PUBLICATION

I. Co-Authorship

I hereby declare that this dissertation incorporates materials that the result of joint research. G.D. Haffner and H.J. MacIsaac contributed to all chapters by providing opportunities and facilities necessary to complete the research along with intellectual guidance. In Chapters 3, 4, 5 and 6, A. Zhan contributed by providing facilities for laboratory analysis. In Chapters 5 and 6, X. Cao contributed by providing samples and laboratory test. In Chapters 2 and 4, M.L. Johansson contributed to data analysis. In Chapter 2, E. DeRoy contributed to data analysis. In Chapter 3, L. Zhang and Y. Gao contributed to sample collection and laboratory analysis, respectively. In Chapter 6, T. Hoxha contributed to data analysis. In all cases, the key ideas, data collection and interpretation, and writing of all chapters were performed by the author, Zhiqiang Xia.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my dissertation, and have obtained written permission from each of the co-author(s) to include the above material(s) in my dissertation.

I certify that, with the above qualification, this dissertation, and the research which it refers, is the product of my own work.

II. Previous Publication

This dissertation includes four original papers that have been previously published/submitted for publication in peer-reviewed journals, as follows:

<table>
<thead>
<tr>
<th>Dissertation chapter</th>
<th>Publication title/full citation</th>
<th>Publication status</th>
</tr>
</thead>
</table>
I certify that I have obtained a written permission from the copyright owner to include the above published material in my dissertation. Chapters published, or accepted for publication, in the above journals have been modified slightly for consistency. I certify that the above material describes work completed during my registration as graduate student at the University of Windsor.

I certify that, to the best of my knowledge, my dissertation does not infringe upon anyone’s copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my dissertation.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my dissertation committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.
ABSTRACT

Management of species invasions ideally requires early detection of species at low abundance, which is often challenging for traditional methods. Environmental DNA (eDNA) provides a promising tool with enhanced sensitivity relative to traditional methods. As an emerging method, however, detection of species at low abundance based on eDNA needs to be optimized to improve detection rate and reduce false negatives. I conducted a meta-analysis, the results of which suggested the significance of using the highly sensitive PCR method and extensive sampling (i.e., replicates sampling and large water volume) to improve detection rate in eDNA-based, low-abundance species detection programs. Needs for improved assay sensitivity screening, testing and reporting were also identified to reduce false negatives and to inform future uses. I developed and optimized an eDNA-based early detection method for the invasive bivalve *Limnoperna fortunei* (golden mussel) and applied it to investigate the spatial-temporal distribution of golden mussel DNA in the central route of South to North Water Diversion Project (SNWDP) in China. I found that improved detection could be achieved by optimizing sensitivity of the method used either through screening primer pairs or PCR methods. A primer pair with a lower limit of detection (LoD) achieved earlier and lower abundance detection of the target species relative to those with higher LoD. Water samples containing re-suspended matter from the bottom layer were better for detection than those exclusively collected from the surface layer, and only sampling the latter caused false negatives. Quantitative PCR yielded higher detection rates than conventional PCR, while the quantification efficiency was reduced in field water samples as compared to total
genomic DNA. Replicate sampling was critical to reduce false negative detections. The majority of positive detections of golden mussel DNA in the main canal of SNWDP were concentrated in warm months, and the occurrence of positive detections was significantly related to minimum daily air temperature, consistent with the expected spawning season of the species. Golden mussel DNA was detected as far as ~1150 km from the putative source of the individuals, indicating long-distance transport of veligers during spawning season. Finally, I tested the functional response and size-selective clearance of the golden mussel to project their potential impacts. Results indicated that golden mussels have a type I functional response, with clearance rate inversely related to food concentration. Presence of golden mussels suppressed suspended matter concentration, the extent of which was dependent on animal abundance, particle size, and their interactions. Golden mussels packaged fine suspended particles into coarser ones, and capture efficiency was inversely related to particle size. Given the suitable habitat and continuous water flow in the main canal of the SNWDP, it seems inevitable that it will be colonized by golden mussels. Abundance mitigation should be considered for the main canal, while containment and dispersal limitation should be prioritized to prevent further spread and reduce overall impact.
DEDICATION

To my mom, Xiaodao Zhang,

for her altruistic love.

悠悠寸草心，难报三春晖。
ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisors, Dr. Douglas Haffner, and Dr. Hugh MacIsaac, for providing me the opportunity to pursue my Ph.D. I acknowledge their support and guidance throughout the enclosed work. I am grateful for their knowledge and expertise in a variety of areas, and their willingness to share research experiences. I thank them for every challenge in my research, which makes me step forward in understanding the nature of science. I also want to acknowledge my graduate committee members, Dr. Aibin Zhan, Dr. Kenneth G. Drouillard, and Dr. Stephen Loeb for their support and insightful advice for my research. Special thanks to Dr. Aibin Zhan for providing laboratory facilities and for enduring long-distance travel to attend my committee meetings and defense. I thank my external examiner, Dr. Jonathan Bossenbroek, for his time and constructive comments. I thank Dr. Kathryn A. Pfaff for being the chair of the defense.

I am grateful to Dr. Lei Zhang and Dr. Junlong Gu, for their support in arrangement of sampling; two major chapters of this dissertation would not have been possible without their support. I thank Yangchun Gao, Xuena Huang and Xinkao Cao for their technical support. I am grateful to my co-authors and several anonymous reviewers for their instructive comments on manuscripts that underlie the chapters of this thesis. I am thankful to Dechao Shen, Xianqi Shen, Xiaosong Zhao, and Ke Hu, for their assistance in the field. I am grateful to Dr. Mattias L. Johansson, Dr. Steve Crookes, Yiyong Chen, Dr. Yuzhan Yang, Emma De Roy, Tedi Hoxha and Justin Barker for many inspiring discussions about my research. I want to thank Deanna Ferguson, Erin Hillis, Anne M.
McLeod, Justin Barker, Joshua Finn, for assistance in my writings. I am also very 
grateful to many colleagues and staff at GLIER, particularly Mary Lou Scratch, Dr. 
Christina Semeniuk, Christine Weisener, Kendra Thompson-Kumar, and members of the 
Haffner, MacIsaac, and Zhan Labs, for their help and support. I want to thank many 
friends in Windsor, particularly Baoyou Zhang, Yueyang Wu, Shifeng Yao, Jinyuan Li, 
for their support which made many of my stressful days easier.

This research was financially support by Natural Sciences and Engineering 
Research Council (NSERC, Canada) Discovery Grants and Canada Research Chairs to 
Douglas Haffner and Hugh MacIsaac, Mitacs Globlink (Canada) to Douglas Haffner, 
Ontario-China Research Innovation Fund to Douglas Haffner and Lei Zhang, and 
National Natural Science Foundation of China to Aibin Zhan. I thank the China 
Scholarship Council for financial support.

Finally, I would like to thank my parents, Xiaoda Zang and Xianlian Xia, and my 
sister, Jing Xia, for their unconditional support, continuous love and encouragement. I 
would also like to thank every other member of my big family for their support and 
encouragement.
TABLE OF CONTENTS

DECLARATION OF CO-AUTHORSHIP / PREVIOUS PUBLICATION ...................... iii
ABSTRACT ......................................................................................................................... v
DEDICATION .................................................................................................................. vii
ACKNOWLEDGEMENTS ............................................................................................. viii
LIST OF TABLES ........................................................................................................... xiii
LIST OF FIGURES ........................................................................................................... xv

CHAPTER 1: GENERAL INTRODUCTION ............................................................... 1
  1.1 Biological invasions: processes and influencing factors ........................................ 1
  1.2 Impacts of invasive species and management ...................................................... 4
  1.3 Environmental DNA as a tool for low abundance species detection in aquatic ecosystems ................................................................. 6
  1.4 Invasions of Limnoperna fortunei in freshwater ecosystems ................................ 8
  1.5 Dissertation objectives ..................................................................................... 10
  1.6 References ......................................................................................................... 13

CHAPTER 2: OPTIMIZING ASSAY SENSITIVITY IMPROVES ENVIRONMENTAL DNA-BASED DETECTION FOR LOW ABUNDANCE SPECIES .............................. 23
  2.1 Introduction ....................................................................................................... 23
  2.2 Materials and Methods ..................................................................................... 25
  2.3 Results ............................................................................................................... 29
  2.4 Discussion ........................................................................................................ 31
  2.5 References ......................................................................................................... 39
  2.6 Supplementary Information ............................................................................. 54

CHAPTER 3: EARLY DETECTION OF A HIGHLY INVASIVE BIVALVE BASED ON ENVIRONMENTAL DNA (EDNA) ................................................................. 62
  3.1 Introduction ....................................................................................................... 62
  3.2 Materials and Methods ..................................................................................... 64
3.3 Results ......................................................................................................................70
3.4 Discussion ................................................................................................................71
3.5 References ................................................................................................................77
3.6 Supplementary Information......................................................................................88

CHAPTER 4: CONVENTIONAL VERSUS REAL-TIME QUANTITATIVE PCR FOR RARE SPECIES DETECTION ..............................................................89
4.1 Introduction ..............................................................................................................89
4.2 Materials and Methods .............................................................................................92
4.3 Results ......................................................................................................................97
4.4 Discussion ................................................................................................................99
4.5 References ..............................................................................................................104
4.6 Supplementary Information....................................................................................117

CHAPTER 5: ENVIRONMENTAL DNA-BASED DETECTION OF AN INVASIVE BIOFOULING BIVALVE IN LARGE WATER DIVERSION CANALS ....................122
5.1 Introduction ............................................................................................................122
5.2 Materials and Methods ...........................................................................................124
5.3 Results ....................................................................................................................129
5.4 Discussion ..............................................................................................................132
5.5 References ..............................................................................................................138
5.6 Supplementary Information....................................................................................151

CHAPTER 6: FUNCTIONAL RESPONSE AND SIZE-SELECTIVE CLEARANCE OF SUSPENDED MATTER BY AN INVASIVE MUSSEL, AND IMPLICATIONS IN A LARGE WATER DIVERSION PROJECT ...............................................................157
6.1 Introduction ............................................................................................................157
6.2 Materials and Methods ...........................................................................................160
6.3 Results ....................................................................................................................166
6.4 Discussion ..............................................................................................................169
6.5 References ..............................................................................................................174
6.6 Supplementary Information....................................................................................189

CHAPTER 7 GENERAL DISCUSSION ........................................................................190
7.1 eDNA-based early detection of aquatic species .....................................................190
7.2 Colonization risks, potential impacts, and management implications of golden mussels in SNWDP .................................................................193
7.3 Future work ..................................................................................................195
7.4 References .....................................................................................................198

APPENDIX A: ADDITIONAL PUBLICATIONS .........................................................203
APPENDIX B: PUBLICATION PERMISSIONS .........................................................211
VITA AUCTORIS .....................................................................................................213
LIST OF TABLES

CHAPTER 2

Table 2.1  Summary of a reduced generalized linear model of detection rate of eDNA from environmental water samples as a function of the PCR method used, the number of sample replicates, and sample volume. 47

Table 2.2  Suggestions of solutions to screen robust assay and summary of some routine means to improve eDNA-based low abundance species detection in literature. 48

Table S2.1  Workflow to generate the dataset analyzed in this study, showing the successional steps, descriptors and the number of results of each. Note that every step was conducted within the results of the previous step. 54

Table S2.2  P-values of the covariates (t-test) of three sequential generalized linear models (f0, f1 and f2), and the variance inflation factor (VIF) of the full model (f0). Model f1 was collinearity removed, and model f2 was stepwise variable selection implemented. 56

Table S2.3  Assay developed for detecting aquatic animal species listed in “the 100 worst invasive species in the world” from eDNA samples. 57

Table S2.4  Current report formats for detection limit of assays in literature and suggested report formats for single species detection from eDNA samples. 58

CHAPTER 3

Table 3.1  Detectability of primer pair B against water samples from surface layer and mixed water column, respectively, collected from both laboratory aquarium tanks and Shisanling Reservoir. +: positive amplification, -: negative amplification; numbers in brackets refer to the lowest dilution fraction from the original eDNA extracts that could be successfully amplified by primer B. ND: not detected. 84

Table S3.1  Sequences and amplicon size of the 13 primers and their detection limit. 88

CHAPTER 4

Table S4.1  Sampling scheme of laboratory water samples. Time interval was 12 hours for days with two sampling points and 24 hours for days with one sampling point. 117
Table S4.2  Determination of limit of detection (LoD) of cPCR (conventional PCR, upper) and qPCR (quantitative PCR, lower). Concentration of total genomic DNA used to determine the LoD for each method was highlighted. Note that cPCR had 10 replicates for each concentration, and qPCR had five replicates for each concentration.

CHAPTER 5

Table 5.1  Summary of linear (or multiple linear) regressions of eDNA quantity (in Log_{10} scale) as a function of distance from source (dist., km) in main canal (Main, June and September in the first phase of sampling), and of distance from source, ≤1 μm suspended particles (pa., # of particles per mL), and water temperature (temp., °C) in sub-canal. ‡ indicates variables used in the multiple regression. ** indicates significance at 0.01 level and *** at 0.001 level.

Table S5.1  Standard curves in qPCR reactions.

Table S5.2  Summary of modelling results of eDNA presence/absence detections and target DNA concentrations (June and September) as a function of average minimum and maximum air temperature, and distance to source reservoir in main canal.

Table S5.3  Summary of environmental conditions of sampling sites in sub-canal.

Table S5.4  Summary of linear regressions of golden mussel DNA concentrations in sub-channel against concentrations of suspended particles with varying sizes. * indicates significance at 0.05 level.

CHAPTER 6

Table 6.1  Summary of the generalized linear model of change percentage (%) of suspended particles on day 1 of experiment 2 as a function of animal abundance (Abundance, # of individual), size of suspended particles (Size, μm), and their interactions (Abundance × Size), showing results of all significant variables. Significance level: *** 0.001, ** 0.01, · 0.1.
LIST OF FIGURES

CHAPTER 2

Figure 2.1  Number of articles on species detection from aquatic environmental DNA samples using species-specific assays (A) and distribution of study cases for each taxonomic group therein (B), showing results of a total number of 223 cases from 140 articles recorded in Web of Science. 50

Figure 2.2  Box and whisker plots of detection rate of a subset (n=65) of cases that detected target DNA from field, showing distribution by taxonomic group (A), PCR method employed (B), and assay sensitivity screening state (C). The width of each box is proportional with the square root of the sample size for the corresponding category. Horizontal bars (bottom to top) of each box represent the 1st quartile, median, and the 3rd quartile, respectively. The number of cases of each box (left to right) is 29, 12, 10, 6, 5, and 3 for A; 20, 43, and 2 for B; 60 and 5 for C. 51

Figure 2.3  Summary of the sources, status of sensitivity screening, limit of detection reporting of assays, and PCR method employed in 140 articles on species detection using eDNA methods. No studies were identified in 2009 or 2010. 52

Figure 2.4  Number of citations in other studies for 17 designed assays of 12 aquatic animal species on “the 100 worst invasive species in the world” (Global Invasive Species Database, 2000). Grey bars represent assays that had sensitivity screened to optimize detection when they were developed, while white bars were unscreened. An asterisk indicates that the limit of detection (LoD) of the assay was reported, otherwise unreported. The dashed line indicates that the assay has accumulated no uses in other studies as of December 31, 2017 (Web of Science). 53

CHAPTER 3

Figure 3.1  Map of the study sites in the field test of the two most sensitive primers, showing water bodies inhabited (solid squares) or uninhabited (open circles) by the golden mussel. 85

Figure 3.2  Detection results of the laboratory validation, showing detection rate as a function of exposure time and number of mussels present. 86
Figure 3.3  Species detection in samples collected from water bodies inhabited (or not) by the golden mussel. Results of Shisanling Reservoir consist of the first sampling (sites 1-3) and the second (field application) sampling (sites 4-5). Dash line indicates no detection.

CHAPTER 4

Figure 4.1  Map of sampling sites in the three irrigation channels (A, n = 6; B, n = 4; C, n = 7), showing the location of each site and detection results of golden mussel by both cPCR and qPCR. Arrows indicate the direction of water flow. Three replicate samples were collected per site, and sampling was carried out from downstream to upstream. Inset indicates location of the study area (asterisk), and dotted line indicates boundary of Henan Province and Hubei Province, China.

Figure 4.2  Detection rate of (a) all replicate samples from laboratory aquaria (n = 33) and field (n = 51) and (b) a subset of the former (n = 12) and latter (n = 15) in which Cq values were assigned to samples owing to skewed melting curves.

Figure 4.3  Mean (±S.D.) of three lowest quantities (solid circle) and their valid Cq (quantification cycle) values (bar) of target DNA detected from total genomic DNA, laboratory samples, and field samples, respectively, using qPCR. Cq refers to the number of cycles required for fluorescent signals to reach a threshold. Different letters indicate significant differences (p < 0.05).

Figure 4.4  False negative rate using one, two, and three replicates in the laboratory (upper) and in the field (lower) using cPCR (grey bar) and qPCR (white bar). Dashed line indicates 100% positive detections for laboratory samples (upper) and positive detections for field samples (lower) determined by qPCR when three replicates were used.

Figure 4.5  Linear regression of Cq against distance to water source (gate A) for water samples collected from channel A (upper, P <0.001), channel C (middle, P = 0.03), and combination (lower, P < 0.001). Each circle indicates a replicate showing positive detection of golden mussel by qPCR, and the thicker circles indicate two overlapped replicates. Note that only one or two replicates were available for some sites.

Figure S4.1  Melting curve of qPCR amplifying high concentration of genomic DNA (upper) and No-Template-Control (NTC, lower), showing
melting temperature 79 - 80 ºC for *Limnoperna fortunei* and 77 - 78 ºC for NTC, respectively.

**Figure S4.2** Standard curve of qPCR, demonstrating the linear regression relationship between Cq and quantity of total genomic DNA (Log Quantity) used in each reaction. Each point is the mean (± SD) of five replicates, except for the lowest concentration which only had three valid Cq values.

**Figure S4.3** Melting curve of qPCR amplifying unknown samples, showing positive amplification of *Limnoperna fortunei* but skewed toward NTC. Cq 33, 34, and 35 were assigned to amplifications similar with A, B, and C, respectively.

**CHAPTER 5**

**Figure 5.1** Map of study area, showing location of central route of the South to North Water Diversion Project (SNWDP) in China (A, dashed rectangle), sampling sites in the main canal of SNWDP (B, red dots), and in the sub-channel (B, green dots), respectively. Both sites M1 and M2 (~30 km apart) were within the Danjiangkou Reservoir, and the asterisk indicates the proximate location of the sub-channel.

**Figure 5.2** Quantity of golden mussel DNA (Mean ± S.E.) in 500 mL water sample in the main canal of the SNWDP and the average daily minimum low air temperature of the neighboring city at each site. Horizontal dashed lines indicate limit of detection ($1.2 \times 10^{-7}$ ng), while the vertical solid line separates the first (left panel: sites M1-M6) and second (right panel: sites M1-3 & M5) phases of sampling. Panels of sampling sites (M1 to M6) are ordered from south to north in the SNWDP. An asterisk indicates a negative detection.

**Figure 5.3** Logistic regression of probability of golden mussel DNA from the main canal as a function of average daily minimum air temperature. Actual detection results are presented as circles. The shaded band indicates ± one standard error of the prediction.

**Figure 5.4** Relationship between concentration of golden mussel DNA and (A) distance from the source, concentration of suspended particles (≤1 µm), and their interactions; and (B) water temperature. Unit of ≤1 µm suspended particles is # of particles per mL. The solid line indicates fitted prediction and shaded band indicates ± one standard error of the prediction.
Figure S5.1  Melting curves of qPCR products, showing melting peaks of both tissue-derived genomic DNA, no-template control (NTC), positive and negative detections of target species from water samples.

Figure S5.2  Sequence alignment of four samples (sequences 2-5) from the main canal mitochondrial COI PCR amplicons with known *L.fortunei* (GeneBank Accession #: HQ843796.1).

CHAPTER 6

Figure 6.1  Clearance rate (Mean ± S.E.) of golden mussels in aquaria (24 °C) at varying algae concentrations. Different letter indicates significance at 0.1 level (Fisher’s LSD test).

Figure 6.2  Fitted functional response (with 95% confidence interval, gray area) of golden mussels. Open circles indicate the actual ingestion rate.

Figure 6.3  Chlorophyll a concentration in mesocosms in which golden mussel aggregates with different number of individuals were (A) daily fed or (B) not fed by adding additional algae throughout the experiment. Test water was aerated tap water in (A) and raw lake water in (B). Error bars (A) indicate standard errors of the average concentration in no-animal control mesocosms.

Figure 6.4  Change percentage (Mean ± S.E.) of concentration of suspended particulates in mesocosms with golden mussels relative to no-animal controls, showing results of each size category at the end (d7) of the experiment. Error bars indicate standard error of three replicates for each sample.

Figure 6.5  Change percentage (Mean ± S.E.) of concentration of suspended particulates in mesocosms with golden mussels related to no-feeding experiment. Error bars indicate standard error of three replicates for each sample.

Figure 6.6  Capture efficiency (Mean ± S.E.) of suspended particles of varying size, showing results of removal rate during the first day. Error bars indicate standard error of three replicates for each sample. Note that the data of size >25 μm in treatments mussel abundance =31 and =100 was not shown because the removal rate (capture efficiency) was significantly related to the interaction between animal abundance and particle size (Table 6.1).
Figure S6.1  \( \log_{10} \) transformed concentration (Mean ± S.E.) of suspended particulate of different sizes in the initial lake water used in experiment 2.
It is a natural phenomenon that species disperse to occupy novel environments. Even though a limited number of species possess exceptional dispersal ability and can move far from their home range, natural dispersal typically occurs within a limited geographic range owing to geographical barriers that isolate species and communities over long periods of time. However, as a result of expanded human activities – such as free trade and globalization – the degree of species dispersal on the earth has been greatly accelerated over both space and time (Elton 1958; Wilson et al. 2009; Lockwood et al. 2013; Simberloff et al. 2013; Roques et al. 2016). Many ecosystems, especially those subjected to extensive human activities, are exposed to enormous numbers of nonindigenous species (NIS) (Elton 1958; Simberloff et al. 2013). Many NIS (e.g., crops and livestock) are intentionally introduced to new environment to meet human demands and are thus considered beneficial (Elton 1958; Simberloff et al. 2013), while others are introduced accidentally and may damage the new environment (Elton 1958; Simberloff et al. 2013; Lockwood et al. 2013). Invasive species are those NIS that exert ecological, health and/or economic impacts on the novel environment or those that have strong potential to disperse widely (Richardson et al. 2000; Blackburn et al. 2011; Lockwood et al. 2013).

1.1 Biological invasions: processes and influencing factors

A typical human-aided process for species being moved out of natural geographic ranges and becoming invasive in novel environment comprises several stages (i.e.,
transport, introduction, establishment, and spread). A number of barriers have been identified that can prevent a species from passing onto a next stage and prevent successful invasions. Specifically, species must overcome geographic barriers to be successfully introduced to novel ecosystem. Upon being introduced, the NIS population must overcome factors that affect survival and reproduction in the new environment. The NIS also interacts with other species in the novel environment, which could increase or decrease survival and/or reproduction.

Even successfully established populations may be constrained at the original establishment site if further dispersal is prevented (Blackburn et al. 2011; Lockwood et al. 2013). Whether or not the NIS invades adjacent (or other) environments is determined by the same group of factors that affected original colonization, though the importance of each factor may differ substantially. Key factors affecting spread include the potential of the recipient ecosystem being invaded (i.e., invasibility), traits of NIS (i.e., invasiveness); introduction effort (i.e., propagule pressure) and their interactions. Invasibility of an environment is a function of its abiotic and biotic conditions, which are characterized but not limited to physical factors such as temperature, precipitation, moisture, pH, salinity, light, nutrient, landscape (Richardson et al. 2000; Simberloff et al. 2013), as potentially by biological features such as presence/absence of predators, competitors, pathogens, and suitable food resources (Williamson & Fitter 1996). Recipient ecosystems with similar environmental conditions to the donor ecosystems usually favor establishment of the introduced NIS (Herborg et al. 2007; Duncan 2016). Many hypotheses have been developed to explain factors affecting invasion success (or invasion failure) (Catford et al. 2009). For example, the enemy release hypothesis suggests that the absence of
enemies in a recipient environment allows NIS to reallocate defense resources to other tissues, enhancing growth, reproduction and/or survival (Keane & Crawley, 2002; Colautti et al. 2004; Jeschke 2014). The biotic resistance hypothesis, on the other hand, suggests that ecosystems with high biodiversity are more stable than less diverse ones and thus less vulnerable (i.e., resistant) to invasion (Elton 1958). A disturbance, however, can alter both biotic and abiotic conditions of ecosystems and have more complex implications with respect to invasion success (Hobbs & Huenneke 1992).

Invasiveness of NIS is often determined by intrinsic characteristics. For instance, the \(_r\)-select life history, high plasticity in genotype or phenotype, broad environmental tolerance, and behavioral syndromes have all been identified as beneficial to invasion success (Pyšek & Richardson 2007; Pyšek et al. 2009; Chapple et al. 2012). The propagule pressure hypothesis suggests that invasion success is determined by introduction effort, which focuses on both propagule size (i.e., the number of individuals of NIS released into new environment in a single event) and propagule number (i.e., the number of release events) (Simberloff 2009; Lockwood et al. 2013). Accordingly, multiple introductions with multiple propagules will aid invasions by overcoming stochastic and/or density-dependent effects (e.g., Allee effect) (Lockwood et al. 2005; Taylor & Hastings 2005). It has to be kept in mind, however, that no single hypothesis can explain all empirical observations in invasion ecology and that often there exists a need for synthesized hypotheses (Catford 2009).
1.2 Impacts of invasive species and management

Invasive species have been recognized as a key driver of global environmental change (Tylianakis et al. 2008; Simberloff et al. 2013) with substantial impacts on ecosystems (Gallardo et al. 2016; Schirmel et al. 2016), economies (Holmes et al. 2009; Pratt et al. 2017) and human health (Hulme 2014). Upon being introduced into novel environments, NIS begin to modify the recipient ecosystem through direct or indirect interactions with either the abiotic or biotic components. System responses can occur on many levels from genes, individuals, populations, communities, to ecosystems (Lockwood et al. 2013; Simberloff et al. 2013). For example, invasive species can affect the genetic integrity of native species through hybridization and introgression (Mooney & Cleland 2001); drive behavioral changes in native animals (Langkilde 2009; Shine 2011); cause extinctions through predation or competition (Savidge 1987; Porter & Savignano 1990; Clavero & García-Berthou 2005); alter underwater light climates (Ricciardi 1998; Sousa et al. 2009; Karatayev et al. 2015); and alter nutrient cycling in both aquatic and terrestrial ecosystems (Hickman et al. 2010; Karatayev et al. 2015). Economic losses caused by invasive species are often tremendous, even though it is very difficult to accurately estimate (Pimentel et al. 2005; Colautti et al. 2006; Lovell et al. 2006; Olson 2006; Xu et al. 2006; Holmes et al. 2009).

Management aims to prevent or control invasive species to reduce detrimental influences. Risk assessment based on potential impacts should be conducted at the very first step as the most cost-effective strategy is to prevent the introduction of potential invaders (Pyšek & Richardson 2010; Lockwood et al. 2013; Blackburn et al. 2014). Prevention of some nuisance species is attainable if fundamental risk assessments are
conducted and proper preventive approaches (e.g., pre-border screenings) are implemented (Kolar & Lodge 2002; Pyšek & Richardson 2010). A critical approach to reduce NIS is to manage their transport vectors and pathways, with the goal of reducing the propagule pressure (see above) and colonization pressure (CP: the number of species introduced into novel ecosystems) (Simberloff 2009; Pyšek & Richardson 2010; Lockwood et al. 2013). A well-documented example of this aim is the management of ballast water (e.g., MacIsaac et al. 2002; Paolucci et al. 2015; Darling et al. 2018). Proposed methods such as ballast water exchange have been adopted in standard protocols to reduce PP and CP introduced to novel aquatic ecosystems (e.g., International Maritime Organization, 2004).

Introductions of NIS are often inevitable and a large number of hitchhikers (e.g., propagules picked up unintentionally at transport stage) can find their way to new environments. Thus, post-border measures, aiming at removing or controlling spread of the introduced propagules, should be considered. Post-invasion measures including eradication, containment, and mitigation can be considered according to the assessments of invader abundance and area affected (Simberloff 2003; Pyšek & Richardson 2010; Blackburn et al. 2011). Eradication is the complete removal of all propagules of an invasive species from the invaded environment (Zavaleta et al. 2001). This is a favored approach to expel invaders, but its feasibility is highly dependent on the affected landscape, area, and target taxon (Pyšek & Richardson 2010). By contrast, containment and mitigation do not attempt complete removal of invaders but rather tries to restrain their spread and abundance, respectively. This approach requires long-term effort and investment (Zavaleta et al. 2001). It should be noted, however, that some undesired
consequences, such as potential disturbance to the native food webs (see Pyšek & Richardson 2010) and increased stress on native endangered species have been reported during the implementation of invasive species management programs (Zavaleta et al. 2001; Lampert et al. 2014). In addition, such undesired consequences are expected to increase with time since the invasion (Pyšek & Richardson 2010). Therefore, it is critical to formulate invasive species management programs with balanced and comprehensive perspectives to trade-off conflicting goals, and in particular to take measures as early as possible (i.e., rapid response) (Pyšek & Richardson 2010).

1.3 Environmental DNA as a tool for low abundance species detection in aquatic ecosystems

Newly introduced NIS usually experience time lags before their population size and affected areas increase (Crooks & Soulé 1999; Lockwood et al. 2013; Rouget et al. 2016). Given the low abundance and limited affected area, the lag time, theoretically, offers an ideal window of opportunity to launch rapid response programs to remove introduced propagules or control their spread at relatively low cost and acceptable success (Simberloff 2003; Pyšek & Richardson 2010). The success of the rapid response management method, however, depends on quickly identifying the invasive threat (i.e., early detection) which is often challenging due to the rarity of individuals at early stages (Mehta et al. 2007; Harvey et al. 2009). This dilemma demands highly efficient diagnostic methods to advance the detection of potential invaders at an early stage.

For many centuries, species detections depended on ‘catch and look’ approaches, by which the targeted species were collected and identified. However, these methods are
often limited in the ability for early detection of invasive species because of inherent deficiencies in sampling and analyses. First, collection of rare, newly-introduced individuals is very difficult, especially in inaccessible ecosystems (e.g., aquatic ecosystems); thus extensive sampling effort is needed to catch target organisms (Grigorovich et al. 2003; Harvey et al. 2009). Secondly, conventional morphology-based taxonomic methods require expertise. The current decline in taxonomists renders species identification difficult (Hopkins & Freckleton 2002; Carlton et al. 2017). Thirdly, identification of some species with cryptic morphological traits or at certain life stages is almost impossible and misidentification is common (Hebert et al. 2003; Briski et al. 2011). Last but not the least, some endangered endemic species may co-habitat the environment with target species, and established traditional sampling methods (e.g., electrofishing, gear netting) can result in negative outcomes for them (Lewison et al. 2004; Wilson et al. 2014; Pikitch 2018).

Environmental DNA refers to the DNA that is shed into the environment by individuals. This DNA can be extracted from bulk environmental samples (Taberlet et al. 2012; Bohmann et al. 2014). In natural aquatic environments, DNA released from organisms tends to be combined with organic matters and persists on suspended particles or in sediments, but rarely as freely-dissolved DNA (Turner et al. 2014). eDNA provides an important alternative to discriminate species of interest by targeting a unique DNA segment (DNA barcode) with no need to collect organisms (Hebert et al. 2003). Either single species or whole communities can be profiled from the total DNA extracted from environmental samples by using species-specific and universal genetic markers. Coupled with advances in DNA sequencing technologies, eDNA has been adopted in a variety of
studies to profile communities or diagnose certain species (Ficetola et al. 2008; Thomsen et al. 2012). As an emerging tool, eDNA is becoming increasingly popular for species detection since the first publication on macro-organisms in 2008 (Ficetola et al. 2008). To date, eDNA-based methods have been used for single invasive and native endangered species detection (Ficetola et al. 2008; Jerde et al. 2011; Fukumoto et al. 2015), biodiversity monitoring (Thomsen et al. 2012), diet analysis (Shehzad et al. 2012), and to estimate abundance (Lodge et al. 2012; Pilliod et al. 2013). A typical application of eDNA methods for species detection comprises multiple steps, including assay development and validation, sample collection and processing, and data analysis and interpretation (Goldberg et al. 2016). It should be kept in mind that for many uses, eDNA is a developing tool and technical challenges need to be identified and considered to formulate effective protocols. Well-developed protocols for the above stages are important to improve detection efficiency of the methods and reduce the potential of both false positive (type I error) and false negative (type II error) results.

1.4 Invasions of *Limnoperna fortunei* in freshwater ecosystems

Many aquatic ecosystems have been invaded by invasive species (Cohen & Carlton 1998; Gallardo et al. 2016; Carlton et al. 2017). Filter feeders such as sponges, tunicates, and bivalves represent important potential invaders (Jeschke et al. 2004). Invasive bivalves (e.g., zebra and quagga mussels, golden mussel) not only profoundly impact invaded ecosystems but also compromise local economies via biofouling (e.g., Gili & Coma 1998; Ricciardi 1998; Boltovskoy & Correa 2015; Linares et al. 2017). The golden mussel (*Limnoperna fortunei*) is a small size freshwater bivalve native to Southeast Asia.
It spread widely from there and invaded South America in the early 1990s (Ricciardi 1998). It is a sessile filter-feeder with adult animals attached to hard underwater surfaces. Through the clearance of suspended particles in the water column, the golden mussel acts as an ecosystem engineer in invaded ecosystems, similar to the *Dreissena* mussels in the Great Lakes (MacIsaac 1996; Ricciardi 1998; Darrigran & Damborenea 2011). In most cases, adult animals form dense clumps through byssal threads, a characteristic that renders the golden mussel a nuisance fouling species in many water facilities such as pipelines for municipal water supply, and hydropower plants (Ricciardi 1998; Boltovskoy 2015). Ballast water-mediated, planktonic larvae introductions were suggested as the major vector for transoceanic invasions of golden mussel, while hull fouling, dam construction and water diversions projects can facilitate their inland spread (Ricciardi 1998; Gois et al. 2015; Zhan et al. 2015). Broad tolerance of environmental conditions predicts the potential for this species to obtain a global distribution (Ricciardi 1998), especially in environments with similar seasonality and temperature, which often results in successful invasions being confined to specific latitudes (Kramer et al. 2017).

The South to North Water Diversion Project (SNWDP) in China (central route) was created to transport water from the Yangtze River basin (central China) to the North to mitigate water scarcity. The newly-constructed canal spans several major water basins with increasing latitude from the source to recipient water bodies, allowing the south aquatic biota to move north (Zhan et al. 2015), extending their south-to-north distribution in China. Relative to other long-distance dispersal pathways (e.g., ballast water-mediated continental dispersal), interbasin water transfer projects can lead to much higher propagule pressure, placing recipient aquatic systems at higher invasion risk (Lockwood
et al. 2005; Wilson et al. 2009). The SNWDP provides a unique case to study the ability of the golden mussel to move from its sub-tropical native ranges to cooler temperate environments. As trillions of propagules can be introduced throughout the year, the risk of spread to temperate reservoirs in northern China has been predicted (Zhan et al. 2015, Appendix A). As a result, the receiving water bodies of the project become more vulnerable to species invasion following the open of the project.

1.5 Dissertation objectives

This dissertation aims to develop a comprehensive optimization and evaluation of environmental DNA-based species detection for taxa at low abundance, using the golden mussel in a large water diversion project – central route of the South to North Water Diversion Project (SNWDP) in China – as a unique study model. First, I review the existing studies on eDNA-based low abundance species to explore optimization strategies of eDNA-based low abundance species detection. Next, I develop an eDNA-based method to detect the golden mussel and optimize the detection sensitivity by screening primer pairs used, PCR method used, water sample sources, and the number of sample replicates to reduce false negative results (type II errors). I then apply the developed method to investigate the temporal and spatial dynamics of golden mussel DNA in main canal of the central route of the SNWDP to map periods of high risk for this species’ dispersal. Finally, investigated the suspension feeding behavior of the golden mussel to project potential impacts of this species on invaded environment, and to form further advice about management. Results of this study can expand current insights in applications of eDNA tools for species detection, providing checkpoints for protocol
optimization to improve efficacy of eDNA methods, and thereby facilitate better management of invasive species.

In chapter two, to call attention to sensitivity optimization of genetic markers used in eDNA methods for species detections at low abundance, I review existing studies focusing on eDNA-based species detection from water samples. I identify the current research limit on assay screening to optimize detection sensitivity, limit of detection (LoD) test to provide complete information for future uses. I also investigate several factors associated with both laboratory and field practices affecting detection rate.

In chapter three, I test the hypothesis that primer pair screening can optimize detection sensitivity and improve detection probability of eDNA methods for low abundance species. I develop a conventional end-point PCR-based eDNA method for detecting golden mussel in water samples. I screen robust primer pairs by testing their LoD, the concentration below which we obtain false negative results. I then validate them by detecting target species in laboratory aquariums with varying abundance at different time points. Finally, the selected primer pairs are validated in field water samples. By testing the consistency of primer pairs’ performance with different water samples, I identified a highly sensitive primer pair that can be used in further applications.

In chapter four, I test the hypothesis that the use of more sensitive PCR methods in the laboratory and replicate sampling in the field can improve detection probability of eDNA methods. I compare detection performance of a newly-developed, real-time quantitative PCR with conventional end-point PCR developed in the previous chapter. Specifically, I quantify detection probability and false negative rate when using different
PCR methods and varying number of sample replicates. I also investigate the eDNA concentration changes with water flow in a channel.

In chapter five, I test the hypothesis that the eDNA-based method can be used to predict periods with high risk of spread by the golden mussel. I apply qPCR developed in chapter three to explore the temporal and spatial dynamics of golden mussel eDNA in the source reservoir and SNWDP (central route) channel. I also model the attenuation of target DNA with flow distance in both the SNWDP canal and a smaller scale irrigation channel. Coupled with the temporal and spatial dynamics of target DNA in the main SNWDP channel, high-risk seasons of golden mussel invasion are identified.

In chapter six, I test the hypothesis that the golden mussel has strong clearance effects on the suspension matter in aquatic ecosystems. I investigate the filter-feeding behaviour of the species, focusing on clearance rate and ingestion rate at varying food supply levels, functional response, massive clearance of suspended matter, and size distribution changes of suspended particles following mussel introductions to the water column. I conclude by discussing the potential impacts of golden mussels in the SNWDP.

Finally, in chapter seven, I summarize the major contributions made by this dissertation and discuss take-home messages for post-invasion management of golden mussels, including the prediction of further spread.
1.6 References


CHAPTER 2: OPTIMIZING ASSAY SENSITIVITY IMPROVES ENVIRONMENTAL DNA-BASED DETECTION FOR LOW ABUNDANCE SPECIES

2.1 Introduction

Fast and accurate identification of low-abundance species, such as newly introduced non-indigenous species (NIS) or endemic endangered species, has attracted the attention of ecologists and policymakers alike, as it underpins management (Mehta et al. 2007; Bohmann et al. 2014; Trebitz et al. 2017). Many traditional methods are challenged in detecting the presence of species at very low abundance (e.g., Harvey et al. 2009; Darling & Mahon 2011; Hoffman et al. 2011; Zhan & MacIsaac 2015).

Environmental DNA (eDNA) has emerged as a promising alternative method to target these species in a non-invasive manner. It has the dual advantages of high environmental prevalence and ease of extraction from bulk environmental samples (Taberlet et al. 2012; Bohmann et al. 2014). By targeting eDNA extracted from environmental samples with properly-selected assays, it is possible to discriminate a single species or profile an entire community (Bohmann et al. 2014). The former, which is the focus of the present study, is achieved by using species-specific primer pairs (and/or probes) (hereafter assays) to amplify single species through PCR and (or) standard sequencing to confirm specificity. The latter, by contrast, relies on the use of cross-species (i.e., universal) assays to amplify multiple species simultaneously and next generation sequencing to specify species (Bohmann et al. 2014). Available evidence indicates that eDNA has higher sensitivity (i.e., a lower false negative rate) than traditional survey methods (Jerde et al. 2011; Zhan et al. 2019). Optimizing assay sensitivity improves environmental DNA-based detection for low abundance species. Molecular Ecology Resources (revision submitted).
et al. 2013; Tréguier et al. 2014). As a result, eDNA has been used to detect a variety of low abundance species, especially in aquatic ecosystems (e.g., Goldberg et al. 2013; Tréguier et al. 2014; Jerde & Mahon 2015; Adrian-Kalchhauser & Burkhardt-Holm 2016; Port et al. 2016).

Despite its popularity, eDNA-based species detection is an evolving approach that is affected by a number of technical issues that impact users’ confidence in results obtained (Roussel et al. 2015; Ficetola et al. 2016). False negatives, for instance, lead to incorrect assessments for ‘undetected’ endangered species or non-native species, which may cause delayed conservation or prevention strategies, respectively (Xiong et al. 2016; Furlan & Gleeson 2017). A typical application of eDNA for species detection comprises multiple stages including selection of assay(s), retrieval of target DNA from the environment, PCR amplification, and data analysis (Liu & Cordes 2004; Goldberg et al. 2016; Xiong et al. 2016; Furlan & Gleeson 2017). Problems at any of these stages may impede detection of low-abundance species. Studies have investigated the importance of some critical considerations when dealing with samples such as collection (Moyer et al. 2014; Mächler et al. 2016), preservation and transport (Takahara et al. 2015), DNA extraction (Deiner et al. 2015), and removal of PCR inhibitors (McKee et al. 2015) to improve enrichment of target DNA from environmental samples. By contrast, much less attention has been paid to the selection of robust assays, which is critical to detect trace target DNA extracted. Prior studies have addressed the importance of gene selection to species detection on a community basis (Deagle et al. 2014; Zhan et al. 2014), while markedly fewer studies have been reported for single species, which should be targeted by species-specific assays (Wilcox et al. 2013; Pedersen et al. 2015).
In the present study, we investigate the current state of factors influential to applications of eDNA methods for low abundance species detection by reviewing existing studies. We focus particularly on efforts devoted to both laboratory (i.e., assay screen, PCR method selection, and PCR reaction protocols) and field practices (i.e., number of sample replicates and sample volume), as these factors have been demonstrated to influence detection probability. We then identify limitations that result from a dearth of information regarding LoD and sensitivity optimization of assays used for eDNA-based species detection. These two issues may result in use of suboptimal assays, thereby inflating false negative errors. We demonstrate this by using eDNA detection studies for species on the “100 of the world’s worst invasive alien species” (Global Invasive Species Database, 2000).

2.2 Materials and Methods

Literature surveys

To generate a quality data set to address the aforementioned issues, we focused on studies using species-specific assays to detect animal species from water samples as they represent the most extensive application of eDNA-based methods for detection of low abundance species. Specifically, we searched the ISI Web of Science (WoS) using “eDNA” or “environmental DNA” and “species detect*” as keywords from 2008 through 2017, which encompasses the first decade since macro-organism detection via eDNA was first used (Ficetola et al. 2008). We narrowed our survey to 11 WoS categories as these categories encompass the majority of articles in aquatic ecosystems. Through paper-by-paper identification, we removed articles that were beyond the scope of this study such as
those exclusively focusing on species barcoding or detecting algae and bacteria, etc. (Table S2.1, supplementary information), resulting in 140 retained articles.

For each article, criteria pertaining to taxonomic group of target species, source of assay used (i.e., designed in study, cited from existing study, or combined when multiple species were included in a single article), assays’ specificity and sensitivity screening state at their development [i.e., screened if any effort was devoted to examining specificity (e.g., in silico BLAST test, in vitro test with tissue-derived DNA of target and non-target species), search for a highly sensitive assay (e.g., testing multiple assays or any other preliminary measures), or unscreened if no such measures were considered], and limit of detection report state (i.e., reported or unreported) of assay used, number of sample replicates (i.e., number of replicate water sample per sampling site, see below the data set description for more details) and sample volume (volume of water collected for each replicate sample), PCR method (i.e., conventional PCR (cPCR, including nested PCR), quantitative PCR (qPCR, including dye-based and probe-based qPCR), or droplet digital PCR (ddPCR)), Template ratio (volume ratio of eDNA extract as template in a PCR reaction), PCR cycles (the number of PCR cycles running in a PCR reaction), and PCR replicates (i.e., the number of replicate PCR running for each sample) were extracted when available. Articles that studied multiple species were broken down to single species, resulting in a case-based data set of 223 cases (Table S2.1, supplementary information). For cases that used cited assays, corresponding information was incorporated from the original studies.

To address the potential that unscreened assays may be used in future studies, we reviewed articles detecting aquatic animal species listed on the “100 of the world’s worst
invasive alien species” (Global Invasive Species Database, 2000). By using the keywords for constructing the above article data set, together with species name (i.e., either Latin or common name), we identified 17 articles in which 12 aquatic animal species on the list were detected via eDNA approaches using originally designed assays (three species with more than one study). Based on these records, we calculated the number of times each original assay was used in other studies. To achieve this, we reviewed all articles citing each of the 17 original articles and only retained studies using the assay.

**Data set description**

Based on the case-based data set, we identified factors associated with laboratory and field practices expected to influence detection rate of target species from field water samples. Specifically, cases could be categorized into two major groups according to study types. First, the study system consisted of a network of connected or potentially connected sampling sites within one larger water body (e.g., a large lake or a major river and its tributaries). In such studies, water samples were collected from multiple sites and information regarding the number of subsamples per site was available. Secondly, systems consisted of independent but contiguous or adjacent water bodies with sampling conducted in each (e.g., isolated ponds or rivers in different basins). In this case, each water body was sampled at multiple sites, though information regarding subsamples was either not available or inconsistent among water bodies, leading to difficulties in data collection. Therefore, we focused on the former study type in this study. In these cases, detection results were dominantly reported as the proportion of sites detected positive. We adopted this as the detection rate of the respective case when it was reported, or
calculated it if it was not directly reported but calculation was possible; cases were not included in analysis when the detection rate was hard to identify. Moreover, only water bodies with putative or confirmed presence of target species (i.e., inhabited sites) were considered, and any ones with target species absent (i.e., uninhabited sites) or those selected as negative controls were excluded prior to analysis. For cases with multiple sampling trips across different time intervals, the average detection rate was used. For one case in which the sampling covered the full spawning season (i.e., Gingera et al. 2016), several trips with 100% detection of the target species within the spawning season were rejected because the target species was too abundant. Given that some cases focused on validation of eDNA methods and tested their methods using water samples with abundant target species, which produced 100% detection, these cases were also excluded from the analysis. We retained 65 cases following the above screenings of the data set, and each case had a detection rate greater than 0% and less than 100%.

**Data analysis**

We conducted data exploration following Zuur et al. (2010) and used a generalized linear model (GLM) to identify the relative importance of the aforementioned factors on species detection rate. The variance inflation factor (VIF) was used to detect multicollinearity among covariates. We set a VIF threshold of 3.0 to discover influential signals which may be weak in the present data set compiled from a broad range of studies (Zuur et al. 2010). Covariates which exceeded this threshold were dropped from the resultant models. As a result, we included three fixed effects (i.e., taxonomic group, PCR method, assay sensitivity screening state) and five covariates (i.e., sample volume,
number of sample replicates, template ratio, number of PCR cycles, and number of PCR replicates) in the initial model. We also included an interaction term between sample volume and the number of sample replicates in the initial model because we expected that studies with large sample volume might limit the number of sample replicates. We Log_{10}(X+1) transformed all covariates except for the template ratio, which ranged between 0.02 and 0.4. We then used the ‘step()’ function in R package ‘MASS’ (Venables & Ripley 2002) for an Akaike information criterion (AIC)-based stepwise variable selection (default ‘both’ direction) to identify which variables contributed the greatest relative importance to detection rate. All statistical analyses were conducted using R (version 3.5.2) (R Core Team, 2018). R code for the GLM used can be found in supplementary information.

2.3 Results

*Increasing popularity of eDNA methods for species detection*

The number of eDNA-based publications on detection of species from water samples using species-specific assays has increased rapidly (Fig. 2.1 A). eDNA detection was most commonly employed with fish, accounting for 124 of 223 cases examined (55.6%), followed by amphibians (14.3%), mussels (11.2%), crustaceans (7.6%), other invertebrates (4.0%), and others (e.g., turtles, aquatic mammals) (Fig. 2.1 B).

*Factors influencing detection rate*

The average detection rate of the 65 cases was 39.6 ± 3.6 % (Mean ± S.E.) though it varied by taxonomic groups. Specifically, detection rate in fish (46.4 ± 6.2%) and crustaceans (48.5 ± 5.1%) was higher than that for amphibians (25.5 ± 5.7%) and mussels
(29.4 ± 9.3%) though no statistical difference ($P > 0.05$) was found among groups (Fig. 2.2 A). Detection rate differed among studies with different PCR methods ($\chi^2 = 8.64, df = 2, P = 0.013$, Kruskal-Wallis rank sum test). Specifically, studies conducted using emerging droplet digital PCR (ddPCR) (47.6 ± 9.5%) and quantitative PCR (qPCR) (46.7 ± 4.7%) had the highest detection rates, with the latter significantly higher than for conventional PCR (cPCR) (23.5 ± 4.5%) ($P = 0.014$, Fig. 2.2 B). Likewise, studies using assays with prior sensitivity screening (60.6 ± 7.6%) had a marginally significant higher detection rate than those without screening (37.9 ± 3.8%) ($P = 0.056$, Mann-Whitney U test, Fig. 2.2 C).

The variable ‘taxon’ was dropped from the initial GLM model due to collinearity (VIF = 3.81), retaining seven variables and one interaction term in the model prior to stepwise variable selection (Table S2.2, supplementary information). Three variables including PCR method, Log(sample volume), and Log(sample replicates) demonstrated significantly ($P < 0.05$) positive relationships with detection rate, explaining 29.1% of the deviance (Table 2.1). No significant interactions were observed between Log(sample volume) and Log(sample replicates) in the reduced model.

**Source, sensitivity screen and report of assays**

The majority (66.4%; consistently over 60% in most years) of studies used newly-designed assays for target species (Fig. 2.3 A). We found that all studies considered the specificity of their assays which were screened, at least, by conducting an *in silico* BLAST (basic local alignment search tool). By contrast, only 7.9% of studies reported sensitivity screening to optimize the assay used, and surprisingly, the pattern did not
improve in recent years (Fig. 2.3 B). Furthermore, only 47.1% of studies reported the LoD of their assays, which tended to fluctuate over time (Fig. 2.3 C). cPCR was widely used until 2013, after which it was largely supplanted by qPCR. qPCR was used in 73.6% studies and increased over time. Droplet digital PCR was used in only a small number of studies (Fig. 2.3 D), though it was very sensitive.

Detection of the ‘worst invasive alien species’

Twenty aquatic animal species are listed in the “100 of the world’s worst invasive alien species” (Global Invasive Species Database, 2000). As of December 31, 2017, only 12 of these species have been recorded in Web of Science as having been detected at least once each from eDNA samples. Brown trout, bullfrog, and common carp were each represented by more than one different assays, resulting in a total of 17 species-specific assay (Table S2.3, supplementary information). Three (17.6%) of these assays were screened to achieve low detection limits and thereby optimizing detection sensitivity, and LoD was reported for nine (52.9%) of them. Seven (41.2%) of the assays were subsequently used in other studies, even though only one (14.3%) had been screened previously for sensitivity and only three had LoD reported (Fig. 2.4).

2.4 Discussion

Detection of newly introduced NIS or endangered, endemic species is typically problematic given the species’ very low population abundance. eDNA serves as a promising tool for detection of these species and is rapidly growing in popularity in consequence (Fig. 2.1 A). eDNA-based species detection relies on both successful enrichment of target DNA from bulk environmental samples and robust visualization
tools in laboratory (Darling & Mahon 2011; Goldberg et al. 2016). Given its low concentration, detection of eDNA requires careful study design and execution in both field and laboratory to avoid missing target DNA and producing false negatives.

**Assay sensitivity optimization is necessary to improve eDNA applications**

The use of a highly sensitive assay (i.e., low LoD) is critical to prevent an inflated false negative rate, as would be expected with an extremely low concentration of target DNA (Darling & Mohan 2011; Furlan et al. 2016). Existing studies have demonstrated that using an optimal assay following sensitivity screening of multiple ones can reduce false negatives (e.g., Adrian-Kalchhauser & Burkhardt-Holm 2016; Ma et al. 2016, Xia et al. 2018a). Results presented here illustrate that studies that used sensitivity-screened assays had a higher detection rate than those that did not (Fig. 2.2 C). Despite this, an overwhelming 92.1% of surveyed articles seemingly overlooked the crucial step of using sensitivity-screened assays in their studies (Fig. 2.3 B). Moreover, 52.9% of these studies did not report the LoD of their assays (Fig. 2.3 C). The absence of knowledge on LoD prevents the researcher from identifying the boundary at which false negative rate increases. This missing information can, in turn compromise eDNA-based species detection programs.

Two recent studies to detect fishes using eDNA methods, and the American Fisheries Society’s standard sampling assessment (i.e., gillnetting, boat electrofishing, and snorkeling), reported overall lower detection rates with eDNA than with standard methods (Perez et al. 2017; Ulibarri et al. 2017). In addition, Ulibarri et al. (2017) challenged the use of eDNA as a detection tool, partially owing to the lack of methods
optimization, considering both sampling methods and assay design. We suggest that both eDNA and traditional methods require optimized sampling to achieve the highest possible detection capability and to compare their relative utility. Besides, some studies reported interspecific differences in detectability with eDNA methods (e.g., Tréguier et al. 2014; Forsström & Vasemägi 2016), which is consistent with results across studies in our analysis (i.e., Fig. 2.2 A). One obvious explanation for divergence among taxonomic groups is that there may be less eDNA available in those species with low detectability (e.g., Forsström & Vasemägi 2016). However, the relative paucity of LoD information clouds this issue, as we observed cases of high detection probability in otherwise low-detectability groups (e.g., mussels, Fig. 2.2 A).

False positives and false negatives are critical concerns in the application of eDNA (or other) tools for rare species detection (Darling & Mahon 2011; Zhan & MacIsaac 2015). Careful selection of candidate assays with high specificity and high detection sensitivity can reduce the risk of both problems (Wilcox et al. 2013; Roussel et al. 2015; Clarke et al. 2017). Even though our primary focus was sensitivity, the remarkable divergence between attention paid to assay specificity and assay sensitivity (screen rates of 100% vs. 7.9%, respectively) highlights the risk of false negatives owing to insufficient knowledge of the latter. Here, the proportion of studies with unscreened assay sensitivity may be overestimated because it is possible that the authors of some studies did, in fact, conduct sensitivity screening or test their assay’s LoD and simply did not report these values in published articles. However, future users of previously-designed assays require this information. Thus we encourage authors to provide both LoD and sensitivity information in all studies.
**Current efforts benefit future uses**

We observed that designed assays were used more frequently than cited ones (Fig. 2.3). Over 40% of the original assays in the ‘worst aquatic invaders’ example have been utilized in at least one other study (Fig. 2.4). Given that eDNA methods for macro-organism detection were first used only a decade ago (Ficetola et al. 2008), that the number of reports of eDNA has been increasing rapidly, and that large numbers of important species have yet to be detected using eDNA, it is likely that even more studies will use literature-based assays in the future. The proliferation of species being barcoded and shared through public platforms and databases such as the Barcode of Life Database System (BOLD) and NCBI provides further testament to the growing databases containing assay information. Almost 47% of studies reported the LoD of their assays used, though formatting varied across studies (Table S2.4, supplementary information). Diverse characterization of the LoD and reporting formats could also impede their use in future studies (see Bustin et al. 2009) because they render difficult comparisons of the assays’ efficiency for a single species. For example, existing studies may report LoD as the number of larvae used, even though larval biomass can vary. A similar issue applies to studies reporting the quantity of genomic DNA, which is comprised of both mitochondrial and nuclear DNA. Even though target genes are commonly mitochondrial DNA- or ribosomal DNA-based, their proportions can widely vary across tissue types, age, and taxa (D’Erchia et al. 2015). Therefore, for future studies of species that have yet to be detected by eDNA methods, we suggest that sensitivity of newly designed assays be optimized, tested, and reported (e.g., Veldhoen et al. 2016; MacDonald & Sarre 2017). In
addition, the testing and reporting of LoD of any newly designed assays should be formatted in a standard manner such as the number of copies of the target fragment, which can be derived using gBlocks gene fragments (i.e., synthesized DNA fragments).

**Possible solutions to screen robust assays**

The evolution of PCR methods reflects the pursuit of highly sensitive eDNA tools, as an increasing number of studies have turned to qPCR or, more recently, ddPCR (Fig. 2.3 D), both of which are more sensitive than traditional cPCR (e.g., Doi et al. 2015; Hunter et al. 2016). High sensitivity of an assay is typically expressed as low LoD, which has a theoretical minimum at one copy per PCR reaction (Bustin et al. 2009). Though it may be challenging to achieve the theoretical low LoD, eDNA assays should be screened to achieve a value as low as possible to maximize detection probability. To screen robust assays, one should keep several things in mind. First, assay design should ensure specificity and maximize sensitivity (Wilcox et al. 2013). Secondly, designing multiple candidate assays for sensitivity screening may be necessary, as limit of detection is unknown until tested. Different genes with multiple assays based on each can be considered to increase the probability of finding highly sensitive ones. For animal species, mitochondrial genes are most widely used because of their abundant biological copies. Different assay design platforms (Table 2.2) can also be considered to generate different high-score assays, which may exhibit varying amplification efficiency as these design tools often stem from different algorithms (e.g., Burpo 2001). Even though further studies are required, the above options can allow multiple assays to be generated and screened to achieve a low LoD (Table 2.2). While detection sensitivity should benefit
from these efforts, it should be noted that this may not be applicable for all species owing to lack of reference sequences. Furthermore, any selected assays should be fully validated using both standard target DNA and/or environmental samples containing target species to confirm specificity and sensitivity before formal use (Wilcox et al. 2013). It is important to keep in mind that highly sensitive assays are sensitive to trace levels of target eDNA regardless of its source, so strict quality control is needed to avoid contamination, which may cause false positives (Darling & Mahon 2011). We did not examine specific means to ensure assay specificity except for in silico BLAST as it is beyond the primary aim of the study. However, as a requirement to reduce false positive risk, specificity of assays should be examined by additional effort such as in vitro and in vivo test (Darling & Mahon 2011; Table 2.2).

**Routine strategies for detection optimization**

Some critical steps should be considered to optimize eDNA-based detection (Table 2.2). As expected, detection rate is positively related to the number of sample replicates and the volume per sample (Table 2.1). This is consistent with both eDNA (e.g., Turner et al. 2014; Xia et al. 2018b) and traditional methods for rare species detection (e.g., Harvey et al. 2009; Hoffman et al. 2011), suggesting the need for extensive sampling. Indeed, reliable eDNA applications require repeat sampling, mainly due to rarity, degradation, and clumped distribution of eDNA in aquatic environments (e.g., Thomsen et al. 2012; Tréguier et al. 2014; Hunter et al. 2015; Furlan et al. 2016; Wilcox et al. 2018; Xia et al. 2018b). Increasing either the number of sampling replicates or the water volume of each sample can increase the probability to capture target DNA. Collecting
larger sample volumes will increase filtration difficulties, and possibly include more inhibitors that challenge later steps. Uses of multi-filters or large-pore-size filters and PCR inhibitor-removal measures can be considered to solve these issues (Table 2.2). Results across studies also highlight the importance of using highly sensitive PCR method (Table 2.1; Fig. 2.2 B), consistent with studies on certain species (e.g., Doi et al. 2015; Xia et al. 2018a). Use of highly-sensitive PCR methods can maximize the probability to amplify even trace amounts of DNA. However, the paucity of LoD information and inconsistent report formatting (e.g., Table S2.4, supplementary information) for studies that do include it impeded our assessment of the relationship between LoD and detection rate.

Though other laboratory factors did not exhibit significant relationships with detection rate, this does not preclude their importance in affecting eDNA-based species detection. For example, running an increased number of PCR replicates has been demonstrated to increase detection probability of low-concentration, target DNA (e.g., Furlan et al. 2016). However, we caution against using increased amounts of eDNA extract in PCRs because of the prevalence of inhibitors in these samples. Reduced detection efficiency at higher template ratio has been reported (e.g., Takahara et al. 2015).

To summarize, rare species detection may benefit from the replacement of traditional sampling approaches with eDNA methods, although technical issues still remain. We argue that the sensitivity of assays should be optimized to achieve robust detection for trace amounts of target eDNA and that studies must report the screening processes used and the LoD. Extensive sampling effort and highly sensitive PCR
methods should be used to improve eDNA-based, species detection for taxa at low abundance. These efforts will collectively instill confidence in sensitivity assessments and in the reliability of eDNA-based low abundance species detection.
2.5 References


**Table 2.1** Summary of a reduced generalized linear model of detection rate of eDNA from environmental water samples as a function of the PCR method used, the number of sample replicates, and sample volume.

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Est. coefficient (S.E.)</th>
<th>t-value</th>
<th>P-value</th>
<th>Deviance explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-0.3386 (0.2226)</td>
<td>-1.52</td>
<td>0.1348</td>
<td></td>
</tr>
<tr>
<td>PCR method: qPCR</td>
<td>0.2519 (0.0760)</td>
<td>3.31</td>
<td>0.0018**</td>
<td>29.1%</td>
</tr>
<tr>
<td>Log (Sample Replicates)</td>
<td>0.2826 (0.1293)</td>
<td>2.19</td>
<td>0.0338*</td>
<td></td>
</tr>
<tr>
<td>Log (Sample Volume)</td>
<td>0.1383 (0.0622)</td>
<td>2.22</td>
<td>0.0312*</td>
<td></td>
</tr>
</tbody>
</table>
Table 2.2 Suggestions of solutions to screen robust assay and summary of some routine means to improve eDNA-based low abundance species detection in literature.

<table>
<thead>
<tr>
<th>Workflow / Check-point</th>
<th>General points</th>
<th>Suggestions or examples</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development of a candidate assay pool</td>
<td>Multiple genes</td>
<td>COI, Cyt-B, D-loop, 16S rDNA, 18S rDNA</td>
<td>Takahara et al. 2012; Turner et al. 2014; Zhan et al. 2014; Deiner et al. 2015</td>
</tr>
<tr>
<td>Development</td>
<td>Comprehensive reference sequences</td>
<td>As many target species sequences as possible; Include closely related species and sympatric species</td>
<td>Darling &amp; Mahon 2011; Wilcox et al. 2013</td>
</tr>
<tr>
<td>Development</td>
<td>Multiple assay design tools;</td>
<td>Primer Premier, Primer3+, Primer BLAST (NCBI), DNAsar, OLIGO7, Vector NTI Advance, Integrated DNA technologies, Eurofins Genomics</td>
<td>Burpo 2001; Present study;</td>
</tr>
<tr>
<td>Development</td>
<td>Multiple candidate assays per gene</td>
<td>Shifting primer binding site; Varying amplicon size (e.g., 100-300 bp)</td>
<td>Ma et al. 2016; Xia et al. 2018b</td>
</tr>
<tr>
<td>Assay validation: in silico test</td>
<td>Primer pair: maximize mismatches with non-target species, and ideally perfect match with target sequence; Mismatches at 3’ end of primers;</td>
<td>Darling &amp; Mahon 2011; Wilcox et al. 2013</td>
<td></td>
</tr>
<tr>
<td>Assay validation: sensitivity</td>
<td>Consult primer quality control tool (e.g., MF EPprimer)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
and sensitivity

<table>
<thead>
<tr>
<th>in vitro specificity test; LoD test</th>
<th>Specificity validation with target species, close and sympatric species; LoD test with standard target DNA fragment (e.g., gBlock genes) and report as # of copies per reaction.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>in vivo specificity test</th>
<th>Field water samples with/without target species; sequence positive amplification products to confirm target species.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Sample collection, proper transport &amp; filtration, and preservation; efficient extraction</th>
<th>Extensive sampling; Multiple (e.g., 4-6) replicates per site; multiple sampling trips if possible; Sampling controls; Transport on ice; Large water volume filtration with large pore size (e.g., 1.5 or 2.0 μm) filter or multi-filter; Non-frozen before filtration; PCI extraction method.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Renshaw et al. 2014; Turner et al. 2014; Deiner et al. 2015; Takahara et al. 2015; Hunter et al. 2019</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PCR analysis</th>
<th>Reduce inhibition; Dilute eDNA extracts (e.g., 10×) or use commercial inhibitor removal kit; Use commercial Environmental Master Mix, or add BSA (bovine serum albumin) in PCR to reduce inhibition; Use qPCR or ddPCR; Run multiple technical PCR replicates (e.g., 6) for each sample; Include negative controls. Sequence at least a portion of positive detections.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deiner et al. 2015; Doi et al. 2015; McKee et al. 2015; Takahara et al. 2015; Furlan et al. 2016</td>
</tr>
</tbody>
</table>
Figure 2.1 Number of articles on species detection from aquatic environmental DNA samples using species-specific assays (A) and distribution of study cases for each taxonomic group therein (B), showing results of a total number of 223 cases from 140 articles recorded in Web of Science.
Figure 2.2 Box and whisker plots of detection rate of a subset (n=65) of cases that detected target DNA from field, showing distribution by taxonomic group (A), PCR method employed (B), and assay sensitivity screening state (C). The width of each box is proportional with the square root of the sample size for the corresponding category. Horizontal bars (bottom to top) of each box represent the 1st quartile, median, and the 3rd quartile, respectively. The number of cases of each box (left to right) is 29, 12, 10, 6, 5, and 3 for A; 20, 43, and 2 for B; 60 and 5 for C.
Figure 2.3 Summary of the sources, status of sensitivity screening, limit of detection reporting of assays, and PCR method employed in 140 articles on species detection using eDNA methods. No studies were identified in 2009 or 2010.
**Figure 2.4** Number of citations in other studies for 17 designed assays of 12 aquatic animal species on “the 100 worst invasive species in the world” (Global Invasive Species Database, 2000). Grey bars represent assays that had sensitivity screened to optimize detection when they were developed, while white bars were unscreened. An asterisk indicates that the limit of detection (LoD) of the assay was reported, otherwise unreported. The dashed line indicates that the assay has accumulated no uses in other studies as of December 31, 2017 (Web of Science).
### 2.6 Supplementary Information

**Table S2.1** Workflow to generate the dataset analyzed in this study, showing the successional steps, descriptors and the number of results of each. Note that every step was conducted within the results of the previous step.

<table>
<thead>
<tr>
<th>Step</th>
<th>Descriptors</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: WoS topic search</td>
<td>Environmental DNA or eDNA, species detect*</td>
<td>2576 articles</td>
</tr>
<tr>
<td></td>
<td>Ecology, Environmental Sciences, Multidisciplinary Sciences, Evolutionary Biology, Marine and Freshwater Biology, Biodiversity Conservation, Zoology, Fisheries, Oceanography, Water Resources, Limnology</td>
<td></td>
</tr>
<tr>
<td>2: WoS categories</td>
<td></td>
<td>1197 articles</td>
</tr>
<tr>
<td>3: WoS publication years</td>
<td>2008-2017</td>
<td>887 articles</td>
</tr>
</tbody>
</table>
| 4: Paper-by-paper identification | 1, Excluding articles focusing on species barcoding  
2, Excluding articles using universal assays.  
3, Excluding articles exclusively focusing on algae, parasites or pathogens, bacteria.  
4, Excluding articles using towed samples. | 140 articles             |
| 5: Case-by-case (species) identification | Any articles including more than one species were broken down to individual species (cases) study. | 223 cases                |
R Code for GLM

eDNA<-read.csv(file.choose(),header=T, sep="","")   ##load the data
attach(eDNA)  ## load vif() function
library(car)   ## load vif() function
xx<-na.omit(eDNA)   ## omit the missing values
f0<-glm(detection.rate ~ taxon + PCR + screen.type + Log.sample.v +
        Log.sample.repl + Log.PCR.cyc + Log.PCR.repl + template.ratio, data=xx)
## full model
summary(f0)
vif(f0)   ## check variance inflation factor
f1<-glm(detection.rate ~ PCR + screen.type + Log.sample.v +
        Log.sample.repl + Log.PCR.cyc + Log.PCR.repl +
        template.ratio + Log.sample.v*Log.sample.repl, data=xx))   ## taxon
## is removed from f0
step(glm(detection.rate ~ PCR + screen.type + Log.sample.v +
        Log.sample.repl + Log.PCR.cyc + Log.PCR.repl +
        template.ratio + Log.sample.v*Log.sample.repl, data=xx))   ## step(f1)
f2<-glm(detection.rate ~ PCR + Log.sample.v + Log.sample.repl, data =
        xx)   ## new model following step(f1)
summary(f2)
par(mfrow=c(2,2))
plot(f2)   ## check residuals
Table S2.2 *P*-values of the covariates (*t*-test) of three sequential generalized linear models (f0, f1 and f2), and the variance inflation factor (VIF) of the full model (f0).

Model f1 was collinearity removed, and model f2 was stepwise variable selection implemented.

<table>
<thead>
<tr>
<th>Covariate</th>
<th><em>P</em>-value (f0)</th>
<th>VIF</th>
<th><em>P</em>-value (f1)</th>
<th><em>P</em>-value (f2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.3715</td>
<td></td>
<td>0.3872</td>
<td>0.1350</td>
</tr>
<tr>
<td>Taxon</td>
<td>0.00513-0.03389*</td>
<td>3.808</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR method</td>
<td>0.00297 ** (qPCR)</td>
<td>2.589</td>
<td>0.0132 (qPCR)</td>
<td>0.0018**</td>
</tr>
<tr>
<td>Sensitivity screen</td>
<td>0.64748 (screened)</td>
<td>1.883</td>
<td>0.5705</td>
<td></td>
</tr>
<tr>
<td>Log (Sample Volume)</td>
<td>0.00338**</td>
<td>1.532</td>
<td>0.7657</td>
<td>0.0312*</td>
</tr>
<tr>
<td>Log(Sample Replicates)</td>
<td>0.69478</td>
<td>1.820</td>
<td>0.6790</td>
<td>0.0338*</td>
</tr>
<tr>
<td>Log(PCR cycles)</td>
<td>0.20755</td>
<td>1.816</td>
<td>0.4365</td>
<td></td>
</tr>
<tr>
<td>Log(PCR replicates)</td>
<td>0.03122*</td>
<td>2.007</td>
<td>0.1762</td>
<td></td>
</tr>
<tr>
<td>Template ratio</td>
<td>0.73632</td>
<td>1.871</td>
<td>0.8529</td>
<td></td>
</tr>
<tr>
<td>Interaction term</td>
<td>0.5338</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table S2.3 Assay developed for detecting aquatic animal species listed in “the 100 worst invasive species in the world” from eDNA samples.

<table>
<thead>
<tr>
<th>Species name (common name)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Eriocheir sinensis</em> (Chinese mitten crab)</td>
<td>Mahon et al. 2011</td>
</tr>
<tr>
<td><em>Cercopagis pengoi</em> (Fishhook waterflea)</td>
<td>Gorokhova 2006</td>
</tr>
<tr>
<td><em>Pomacea canaliculata</em> (Golden apple snail)</td>
<td>Cooke et al. 2012</td>
</tr>
<tr>
<td><em>Carcinus maenas</em> (Green crab)</td>
<td>Jones et al. 2008</td>
</tr>
<tr>
<td><em>Asterias amurensis</em> (Northern pacific seastar)</td>
<td>Deagle et al. 2003</td>
</tr>
<tr>
<td><em>Dreissena polymorpha</em> (Zebra mussel)</td>
<td>Mahon et al. 2011</td>
</tr>
<tr>
<td><em>Rana catesbeiana</em> or <em>Lithobates catesbeianus</em> (Bullfrog) a</td>
<td>Ficetola et al. 2008</td>
</tr>
<tr>
<td><em>Rana catesbeiana</em> or <em>Lithobates catesbeianus</em> (Bullfrog) b</td>
<td>Strickler, et al. 2015</td>
</tr>
<tr>
<td><em>Rana catesbeiana</em> or <em>Lithobates catesbeianus</em> (Bullfrog) c</td>
<td>Veldhoen et al. 2016</td>
</tr>
<tr>
<td><em>Salmo trutta</em> (Brown trout) a</td>
<td>Gustavson et al. 2015</td>
</tr>
<tr>
<td><em>Salmo trutta</em> (Brown trout) b</td>
<td>Carim et al. 2016</td>
</tr>
<tr>
<td><em>Cyprinus carpio</em> (Common carp) a</td>
<td>Takahara et al. 2012</td>
</tr>
<tr>
<td><em>Cyprinus carpio</em> (Common carp) b</td>
<td>Turner et al. 2014</td>
</tr>
<tr>
<td><em>Cyprinus carpio</em> (Common carp) c</td>
<td>Eichmiller, Bajer &amp; Sorensen 2014</td>
</tr>
<tr>
<td><em>Micropterus salmoides</em> (Largemouth bass)</td>
<td>Perez et al. 2017</td>
</tr>
<tr>
<td><em>Oreochromis mossambicus</em> (Mozambique tilapia)</td>
<td>Robson et al. 2016</td>
</tr>
<tr>
<td><em>Oncorhynchus mykiss</em> (Rainbow trout)</td>
<td>Wilcox et al. 2015</td>
</tr>
</tbody>
</table>
Table S2.4 Current report formats for detection limit of assays in literature and suggested report formats for single species detection from eDNA samples.

<table>
<thead>
<tr>
<th>PCR methods</th>
<th>cPCR</th>
<th>qPCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current formats</td>
<td>0.5-0.01 ng; 7.25 $\times 10^{-11}$ ng μl$^{-1}$; 1 zooid; 5 D-hinge-stage larvae; 50 copies</td>
<td>1 individual per 100 m$^2$; 2 copies; 3 copies per reaction; 1 larva; 0.013-3.2 pg μl$^{-1}$; &lt;10 copies 4μl$^{-1}$</td>
</tr>
<tr>
<td>Suggested formats</td>
<td># of copies</td>
<td># of copies</td>
</tr>
</tbody>
</table>
References for Figure 2.4


CHAPTER 3: EARLY DETECTION OF A HIGHLY INVASIVE BIVALVE BASED ON ENVIRONMENTAL DNA (EDNA) ²

3.1 Introduction

Non-indigenous invasive species (NIS) are a leading cause of ecological, economic and human health harm, and thus have received the attention of both scientists and policymakers (e.g., Walsh et al. 2016). Decisions regarding management should be made early to enhance the likelihood of success, though this is largely dependent upon early detection (Brown et al. 2016; Holden et al. 2016; Xiong et al. 2016). Traditionally, morphological trait-based methods were utilized to identify species. However, these methods have limitations when dealing with cryptic species (Heinrichs et al. 2011), and species with ontogenetic stages or exhibiting extreme phenotypic plasticity (e.g., Kekkonen & Hebert 2014). These methods are also increasingly challenged by the long-term erosion in systematics expertise (e.g., Xiong et al. 2016). Furthermore, detecting rare NIS may require very extensive sampling to maximize species retrieval (e.g., Harvey et al. 2009).

The advent of the DNA barcode created a novel way to conduct species identification that bypasses morphological traits, allowing researchers lacking taxonomic expertise to nevertheless identify species (Hebert et al. 2003; Hebert et al. 2004; Ikeda et al. 2016). Environmental DNA (eDNA) refers to DNA shed into the environment, and its presence in aquatic systems allows for efficient and sensitive identification of target species from bulk water samples (e.g., Bohmann et al. 2014). For rare species, such as

newly introduced NIS or endangered species, eDNA usually yields a higher detection rate than traditional methods such as trapping and seining (Dejean et al. 2012; Dougherty et al. 2016; Schmelzle & Kinziger 2016). In addition, eDNA-based metabarcoding usually recovers more species in a community than morphological methods (Zhan & MacIsaac 2015; Valentini et al. 2016).

The bio-fouling golden mussel (*Limnoperna fortunei*) is native to Southeast Asia but has spread widely there (Nakano et al. 2015) as well as in eastern South America (see Boltovskoy 2015). The species poses serious biofouling problems to water supply and drainage systems and is an ecosystem engineer in many invaded ecosystems, altering planktonic and benthic communities and changing nutrient cycling (see Boltovskoy 2015). The species is similar to *Dreissena* species (i.e., zebra and quagga mussels) in terms of physiological traits and ecological impacts, though it possesses broader tolerance to many environmental conditions and thus has the potential of a wider distribution (Karateyev et al. 2007). The species has many avenues for dispersal, including as veliger larvae in actively transported water (ballast, live wells; Ricciardi 1998) or in currents and water diversion channels (Zhan et al. 2015), while adults may disperse on external surfaces of vessels or floating debris (see Boltovskoy 2015). Despite the negative impacts of this species in both Asia and South America, heretofore early detection has not been well prioritized though it had been molecularly identified (Pie 2006; Pie et al. 2017).

In this study, we developed an eDNA-based early detection method and conducted a sensitivity test using a serial dilution of total genomic DNA. Subsequently, we explored species detectability in laboratory and field using our most sensitive primers. We also optimized a field sampling strategy to minimize false negatives.
3.2 Materials and Methods

To obtain robust primers to detect the golden mussel from environmental samples, we first designed 13 primer pairs based on the mitochondrial COI gene. Subsequently, we conducted a sensitivity test using a serial dilution of the total genomic DNA to screen these primer pairs. The two most sensitive primer pairs were tested further based on laboratory aquarium experiments for sensitivity validation using controlled numbers of animals. Finally, we chose our most sensitive primers for detecting this mussel based on environmental samples collected from a variety of water bodies using an optimized sampling method.

Primer design and laboratory sensitivity validation

To design species-specific primers, we downloaded all 56 available mitochondrial COI sequences (>500 bp) of *L. fortunei* in the NCBI GenBank (https://www.ncbi.nlm.nih.gov/) and retrieved representative COI sequences with high similarity (>70%) to *L. fortunei*. These sequences were aligned in MEGA (version 6.06) and inspected to determine conserved regions for our target species (*L. fortunei*) but sufficiently variable in related species to avoid cross-species amplification and false positive results. Given that DNA in aquatic ecosystems is usually highly degraded, we designed primers with fragment length ranging from 127 to 299 bp using Primer Premier 5 (PREMIER Biosoft) based on suggestions from other studies (Bohmann et al. 2014; Dougherty et al. 2016). To ensure species specificity, all 13 primer pairs were compared with available mitochondrial COI sequences of five mollusk species (family Unionidae:...
*Unio douglasiae* and *Sinanodonta woodiana*; family Viviparidae: *Bellamya purificata*, family Semisulcospiridae: *Semisulcospira cancellata*; family Cyrenidae: *Corbicula fluminea*) which may be recovered in our sampling areas. Also, we tested the primers against the total genomic DNA extracted from tissues of these species and results produced no positive amplification. As *L. fortunei* is the only freshwater species in genus *Limnoperna* and the order Mytiloida, we did not consider other related species in the order Mytiloida when conducting the species-specificity test.

The sensitivity of a primer pair was characterized by the limit of detection (LoD: the lowest amount of genomic DNA that can be amplified) in a 25 μL PCR reaction. A lower LoD equates to higher sensitivity for the species. We performed 10-fold serial dilutions of genomic DNA from an initial concentration of 40.0 ng μL⁻¹ to achieve a series of concentrations from 4 to 4×10⁻⁸ ng μL⁻¹. A total of 10 replicates were carried out for each concentration. To reduce biased PCR amplification, we defined the detection limit of each primer pair as the lowest amount of genomic DNA that could be successfully amplified in five or more replicates.

Animals used for the laboratory validation test were collected from the Pengxi River, a tributary of the Yangtze River near Chongqing, China (Fig. 3.1). A submerged brick covered with golden mussels was collected and brought in a cooler with wet towels back to the laboratory and acclimated at 26°C in a 60 L aquarium tank. We prepared single animals and animal aggregates with five and 15 medium-sized (shell length ~15 mm) individuals by carefully cutting byssal threads, following which single animals were allowed to re-attach on a glass slide during acclimation. During acclimation, animals were fed every two days with 50 mL commercial *Chlorella sp.* (~10⁸ cells L⁻¹). Three
treatments with one, five and 15 animals and five replicates for each treatment were used in the laboratory validation. The average wet weight of animals for the three treatments was 0.994, 4.301 and 10.554 g, respectively. Prepared animal aggregates (or slides with one animal on each) were taken from the acclimation tank and flushed by double-distilled water to eliminate chemical traces prior to deployment in experiment tanks (10 L target-DNA-free water therein). Water used in the acclimation tank and experiment tanks was well-aerated tap water, and the water was detected as target DNA-free throughout the whole experiment. Healthy animals typically began to filter within five minutes after transfer. A 250 mL water sample for mussel eDNA was collected from each tank after 20, 40, 60, 90 and 120 minutes exposure in the tanks and an additional sample was collected at 10 and 30 min, respectively, for the treatment with 15 animals. Animals were not fed during the two-hour experiment. The tank was mixed thoroughly before water collection and each sample was filtered onto a 0.45 μm pore size cellulose acetate microporous membrane filter. Each filter was preserved in a 2 mL centrifuge tube and stored under -20°C until DNA extraction. Detection time (i.e., time between animal deployment and the point of first detection from the aquarium water sample) of each replicate was recorded, while detection time of each treatment was reported when ≥50% of replicates had successful detections.

We conducted both laboratory and field experiments to optimize the sampling method before the field application. We tested the hypothesis that water samples containing particulate matter re-suspended from the bottom layer were more likely to test positive for the species than those without re-suspended particulate matter. We examined water samples collected by two sampling methods using the most sensitive primer pair:
1) collecting only surface water without disturbing the bottom layer; and 2) collection of surface water after the bottom layer was disturbed and particulates re-suspended.

In the laboratory experiment, we combined water from the five 15-animal treatment tanks, which experienced an additional 7d culture after the two-hour of laboratory validation experiment to accumulate eDNA into a new tank. We allowed the tank (~40 cm depth) to remain undisturbed for 5 d to let any particulate matter sink to the bottom. Subsequently, we carefully siphoned three 250 mL water samples from the surface layer (~5 cm) prior to collecting the same amount of water after thoroughly mixing the tank. We repeated the same two sampling methods at three field sites in Shisanling Reservoir (Fig. 3.1: sites 1-3). Specifically, we collected water samples from the surface layer (~25 cm) and then we disturbed the bottom of the same site, followed by immediate collection of the water containing re-suspended particulates at the depth ~50 cm above the bottom. We examined these samples using primer pair B by detecting golden mussels in their eDNA extracts from the water samples and serial dilutions thereof, with the expectation that more positive detections would be observed at a higher dilution rate from the samples collected from the disturbed water column than those without disturbance.

**Field application**

From June to July 2015, a total of 22 field sites in three water bodies where golden mussels have been reported [i.e., Pengxi River and Danjiangkou Reservoir in South China and Shisanling Reservoir (sites 4 and 5) near Beijing in North China; Fig. 3.1], plus nine sites in two systems where golden mussels have not been reported (i.e., Miyun Reservoir and Kunming Lake in Beijing; Fig. 3.1) were sampled for detecting golden
mussel. Sampling sites at each location were randomly determined. All water bodies were
lentic systems except Pengxi River, which was sampled from downstream-to-upstream to
avoid cross-contamination. For the other water bodies, we disturbed the bottom layer
before collecting water when depth permitted. Three 250 mL subsamples were collected
at each site and those samples were then shipped to the laboratory in a cooler with ice and
processed within 12 h of collection. Water was filtered and stored until DNA extraction.
To reduce false positives, we considered a positive detection if ≥2 of subsamples at a site
detected the species.

Pengxi River and Danjiangkou Reservoir are located in the native region of golden
mussels in South China. While we did not quantify the abundance of golden mussels at
these sites, the density is higher at the former than the latter. Shisanling Reservoir is
outside of the mussel’s native region but was previously invaded (Ye et al. 2011); thus
we conducted a field survey in May 2015 (prior to the formal sampling in July 2015) and
observed a very low density of animals on hard surfaces near the shoreline. Sites that
have not been colonized by golden mussels were sampled twice in August 2014 and May
2015 (Miyun Reservoir) by field observation, and verified by interviews of local
residents and reservoir managers.

**DNA extraction and PCR**

Genomic DNA was extracted from animals collected from Pengxi River (Fig. 3.1),
and extracted DNA was quantified by Nanodrop 2000 spectrophotometer (Thermo
Scientific). The eDNA on the filter was extracted following a modified protocol based on
a published method (Waters et al. 2000). Specifically, digested solution was transferred
to a new 2 mL tube before centrifugation and the air-dried DNA pellet was dissolved in 30 μL pure water; proteins were precipitated by 7.5 M ammonium acetate and DNA was precipitated by 100% ethanol.

We conducted gradient temperature PCR for each primer pair prior to formal sensitivity testing to determine the optimized annealing temperature for each primer pair. PCR was conducted in a 25 μL reaction volume containing 1× PCR buffer, 1 μL DNA extract, 0.05 mM of each dNTP, 0.4 mM of each primer, 2.0 mM of Mg\(^{2+}\), 1 unit of Taq Polymerase (Takara Bio Inc.). PCR amplification was performed in a Mastercycler (Eppendorf) with a thermal profile consisting of an initial denaturation at 95°C for 5 min, followed by 35 cycles of denaturation at 95°C for 30 s, annealing at 55°C for 35 s and extension at 72°C for 30 s, with a final extension at 72°C for 7 min. PCR products were visualized and analyzed using an automatic gelatin image analysis system, by loading 5 μL of PCR products in each well of 2% agarose gels stained with ethidium bromide. Positive reactions were identified for \textit{L. fortunei} by a single, distinct band at 197 bp and 286 bp for primer pairs B and I (Table S3.1, supplementary information), respectively.

To avoid cross-contamination, all bottles used for sample collection were new. During each sampling trip, two bottles were filled with deionized water and placed with sampling bottles as sampling blanks. Vessels were regularly maintained by cleaning hull and checked before sampling to ensure that no mussels colonized. Non-disposable tools (e.g., forceps and glassware) used in sample filtration and DNA extraction were immersed in 10% commercial bleach for 10 minutes to destroy residual DNA before reuse. Aerated tap water was filtered for laboratory validation experiment tanks and deionized water was filtered for sample filtration equipment (e.g., filters and glassware)
as equipment controls throughout laboratory analyses. In each batch of PCR, a positive (total genomic DNA) and negative control (deionized water) were applied to assure good practice.

### 3.3 Results

Detection limit varied widely across the 13 primer pairs tested, ranging from $4 \times 10^{-2}$ to $4 \times 10^{-6}$ ng of total genomic DNA (Table S3.1, supplementary information). Primer pair B was the most sensitive and could amplify genomic DNA as low as $4 \times 10^{-6}$ ng. Another three primer pairs (G, I and L) also had relatively high sensitivity ($4 \times 10^{-4}$ ng). These primer pairs exhibited a similar capacity to amplify aquarium eDNA in a pilot test prior to our laboratory, validation experiment, so we subsequently utilized primer pairs B and I for our formal validation experiment.

For primer pairs B and I, the time required for three out of five replicates detected was inversely related to animal abundance in aquaria (Fig. 3.2). Specifically, detection with primer pair B occurred in 60, 40 and 20 minutes, using one, five and 15 animals, respectively, while the less sensitive primer pair I required 90, 90 and 60 min. Primer pair B was more sensitive to amplification of aquarium eDNA samples than primer pair I and detection occurred 30, 50 and 40 min earlier than those for primer pair I with one, five and 15 animals, respectively.

Golden mussels were detected when using both sampling methods in the laboratory experiment (Table 3.1) by primer pair B. Specifically, eDNA extracted from the water surface layer could be detected at $10^{-2}$ dilution fraction, while samples from the mixed water column were detectable at $10^{-4}$ dilution fraction. Among samples collected in
Shisanling Reservoir, no positive detections were observed for surface layer samples, though mixed water samples recorded positive detections at two sites. Specifically, mixed samples from sites 2 and 3 detected presence of golden mussel at 20% and 10% dilution fraction, respectively (Table 3.1). All equipment controls and sampling blanks yielded negative tests throughout this study.

Primer pairs B and I detected golden mussels in natural waters inhabited by the species, while the samples never tested positive for all areas where the species has never been reported (Fig. 3.3). Similar to our laboratory validation, primer pair B was more sensitive than primer pair I in the field, yielding four more positive detections in Danjiangkou Reservoir, though both primer pairs yielded a 100% hit rate in Pengxi River (10 sites). In Danjiangkou Reservoir, primer pair B successfully detected the species at 70% of sites (i.e., sites 1-3 and 5-8), while primer pair I detected at only 30% of sites (i.e., sites 6-8). In Shisanling Reservoir, all five sites were examined by primer pairs B and I, though only sites 4 and 5 were sampled during the formal field sampling. Primer pair B detected the species at 40% of sites (i.e., sites 2 and 3) while primer pair I did not detect at any site. Also, we performed Sanger sequencing for PCR products randomly selected from sites 6 and 7 from Danjiangkou Reservoir using primer pairs B and I, both of which correctly identified the presence of golden mussels.

3.4 Discussion

The golden mussel is rapidly expanding its distribution in both South America and Southeast Asia (Boltovskoy 2015). Early detection is, therefore, an essential component of the species’ management, though this aspect has not been extensively studied.
Recently, Pie et al. (2017) developed a COI-based method to detect this species, which was sensitive to DNA levels as low as $2.25 \times 10^{-4}$ ng. In this study, we screened 13 primer pairs, the most sensitive of which (primer pair B) could detect down to $4 \times 10^{-6}$ ng. Indeed, primer pair B was highly sensitive to eDNA samples from both laboratory and field (i.e., river, reservoir, and lake). It is not surprising that the detection limit of the former was poor relative to primer pair B, as our study was specifically designed to identify a highly sensitive primer pair.

Enormous numbers of samples may be required to detect NIS when they are present at low population abundance in order to preclude false negatives (i.e., Harvey et al. 2009; Hoffman et al. 2011). Judicious selection of primers in this study dramatically improved detection sensitivity to low levels of eDNA. Screening primer pairs resulted in two benefits. First, sensitive primer pairs could be identified that allow for early detection in field applications (Fig. 3.2). This is critical for some groups of NIS which have long lag times before populations exceed traditional detection thresholds. Second, utilization of a sensitive primer pair on field-collected samples should reduce the likelihood of false negative results, a major limitation of traditional field sampling for rare species (e.g., Wilcox et al. 2016; Schultz & Lance 2016) or when species release only small amounts of extracellular DNA (Tréguier et al. 2014). For example, we detected the presence of golden mussels in Shisanling Reservoir using primer pair B though primer pair I failed to detect the species (false negative; Fig. 3.3). The use of highly sensitive primer pairs can reduce the occurrence of such false negatives. We did not, however, observe performance difference between the two primer pairs for Pengxi River (Fig. 3.3), where mussel abundance (and presumably eDNA concentration) is much higher. In addition, neither
primer pair produced positive detection in the negative control regions, indicating no false positives.

Both conventional PCR and quantitative real-time PCR (qPCR) have been widely utilized for species detection based on eDNA in environmental samples. Both methods can be used to infer species presence/absence, but the latter is more sensitive and informative and can provide additional abundance information of target DNA template (Wilcox et al. 2013; Balasingham et al. 2017). We used conventional PCR in this study as this method is cost-efficient and can be conducted in most laboratories. A critical concern using conventional PCR in eDNA studies is the potential for false positives, which may result from either low specificity of primers or from contamination (Ficetola et al. 2016). No false positives were recorded with these primer pairs in any tested water bodies from which *Limnoperna* has never been reported (Fig. 3.3), nor in any negative controls or blanks used during sample processing. This suggests that the high specificity of our primers, coupled with careful sample collection and handling protocols, precluded false positives. In addition, we randomly selected and sequenced several PCR products to confirm positive detections further. If the primers developed here were utilized in other areas, specificity tests against sympatric species could be conducted to assess the possibility of false positives. As conventional PCR is not quantitative and cannot provide abundance information, qPCR-based methods appear superior in this regard.

DNA may be degraded due to a variety of factors, notably high water temperature (Taberlet et al. 2012; Barnes et al. 2014). Inhibitors present in eDNA samples may also impede species detection (McKee et al. 2015). The detection limit of primers in this study was determined using the total genomic DNA, while the sensitivity against the water
samples was not known although we expected primer pair B to perform best. Results indicate that the capacity of a primer pair to detect species in environmental samples was influenced by its detection limit, indicating that screening primer pairs can increase sensitivity for detection purposes.

We utilized the mitochondrial COI gene, which has proven reliable in animal species discrimination (e.g., Pečnikar & Buzan 2014; Jiang et al. 2016). In addition, the abundance of mitochondrial genes is higher than nuclear ones, as there are usually multiple copies of the former in somatic cells. Consequently, mitochondrial genes have a higher probability being detected than nuclear ones in a single, complex animal eDNA sample (Taberlet et al. 2012). Degradation of eDNA requires that amplicon size of candidate primer pairs be constrained to a small size. Thus, designing more candidate primer pairs with relatively small amplicons, either from single or multiple genes (Pečnikar & Buzan 2014), may increase the probability of finding highly sensitive primers. However, further studies are required to explore the relationship between the number of candidate primer pairs and the highest sensitivity achieved.

To maximize the efficiency of eDNA-based methods in rare species surveillance, knowledge of target DNA sources is important to ensure that DNA is collected from environmental samples (e.g., Lacoursière-Roussel et al. 2016). The golden mussel is a filter-feeder that produces feces and pseudofeces, aggregates of which collect on the sediment surface (see Boltovskoy 2015). This material provides a good opportunity to collect eDNA. In addition, the microenvironment of the sediment surface allows the eDNA to be better preserved than upper layer water column due to organic matter and minerals (Tréguier et al. 2014; Turner et al. 2015) as well as reduced UV exposure.
Thus water samples containing stirred up sediments may yield a higher detection probability for target species than those collected strictly at the surface. Results from sampling of both aquaria and field sites demonstrated that water samples containing resuspended particulate matter were a better eDNA source than surface waters only (Table 3.1). In the field, areas with lower disturbance may experience a higher particle settling rate, thereby capturing available eDNA and resulting in lower eDNA levels in the surface layer (Turner et al. 2014). An optimized sampling method was developed for crayfish Procambarus clarkii surveillance in ponds with these issues in mind (Tréguier et al. 2014). One concern is that positive detections from sediment samples may reflect historic DNA from animals no longer present at the site or DNA transported from another location. However, eDNA is relatively short-lived in common aquatic environments (Turner et al. 2014; Strickler et al. 2015) and for sedentary organisms like golden mussels, we expect this issue to be negligible in most cases. Here we did not detect target DNA signals from the surface water samples, indicating absence (or extremely low density) of free-swimming veligers in the upper layer of the species.

Harnessing a sensitive primer can improve rare species detection (Zhan et al. 2013; Zhan et al. 2014). Our study demonstrated that sensitivity against genomic DNA varied among primers, even though all primers used were designed based on a single gene with similar amplicon size. However, the sensitivity of primer pairs used in eDNA methods applications are often not tested or stated (e.g., Dejean et al. 2012), with comparatively few studies conducting the preliminary step of screening these primer pairs before use (e.g., Dougherty et al. 2016). Indeed, many researchers utilize primer pairs recommended by primer design software (e.g., Rees et al. 2014) or those available in the literature (e.g.,
Egan et al. 2013). The lack of information pertaining to sensitivity may unwittingly expose these studies to an elevated false negative rate, which can be problematic if the goal of a project is the detection of a NIS or a threatened species. Goldberg et al. (2016) summarized a number of critical aspects that should be considered when designing eDNA-based methods to detect aquatic species. We add that careful screening of primers is an essential step that ought to be coupled with these considerations and be employed in all such studies.

In conclusion, eDNA-based methods are increasingly used in rare species detection, though the sensitivity of applied genetic markers and their corresponding primers is rarely tested. In this study, we found that the sensitivity of different primer pairs varied widely, and urge that primer sensitivity should be known (or tested) before use in the field. When combined with an optimized sampling strategy and good field and laboratory practices, the use of highly sensitive primers can reduce false negative results.
3.5 References


Egan, S. P., Barnes, M. A., Hwang, C. T., Mahon, A. R., Feder, J. L., Ruggiero, S. T.,
combining environmental DNA with light transmission spectroscopy. Conservation
Letter, 6, 402-409.
environmental DNA and metabarcoding? Molecular Ecology Resources, 16, 604-
607.
Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M.
A., … Taberlet, P. (2016). Critical considerations for the application of
environmental DNA methods to detect aquatic species. Methods in Ecology and
Evolution, 7, 1299-1307.
Harvey, C. T., Qureshi, S. A., & MacIsaac, H. J. (2009). Detection of a colonizing,
aquatic, non-indigenous species. Diversity and Distributions, 15, 429-437.
identifications through DNA barcodes. Proceedings of the Royal Society B:
Biological Sciences, 270, 313-321.
Ten species in one: DNA barcoding reveals cryptic species in the neotropical
skipper butterfly Astraptes fulgerator. Proceedings of National Academy of
Sciences (USA), 101, 14812-14817.
Wissemann, V. (2011). Formalizing morphologically cryptic biological entities:
new insights from DNA taxonomy, hybridization, and biogeography in the leafy


Table 3.1 Detectability of primer pair B against water samples from surface layer and mixed water column, respectively, collected from both laboratory aquarium tanks and Shisanling Reservoir. +: positive amplification, -: negative amplification; numbers in brackets refer to the lowest dilution fraction from the original eDNA extracts that could be successfully amplified by primer B. ND: not detected

<table>
<thead>
<tr>
<th>Sample source</th>
<th>Surface layer</th>
<th>Mixed water column</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aquarium tank (replicate 1)</td>
<td>+ (10^{-2})</td>
<td>+ (10^{-4})</td>
</tr>
<tr>
<td>Aquarium tank (replicate 2)</td>
<td>+ (10^{-2})</td>
<td>+ (10^{-5})</td>
</tr>
<tr>
<td>Aquarium tank (replicate 3)</td>
<td>+ (10^{-2})</td>
<td>+ (10^{-4})</td>
</tr>
<tr>
<td>Shisanling Reservoir (site 1)</td>
<td>- (ND)</td>
<td>- (ND)</td>
</tr>
<tr>
<td>Shisanling Reservoir (site 2)</td>
<td>- (ND)</td>
<td>+ (0.2)</td>
</tr>
<tr>
<td>Shisanling Reservoir (site 3)</td>
<td>-(ND)</td>
<td>+ (0.1)</td>
</tr>
</tbody>
</table>
Figure 3.1 Map of the study sites in the field test of the two most sensitive primers, showing water bodies inhabited (solid squares) or uninhabited (open circles) by the golden mussel.
Figure 3.2 Detection results of the laboratory validation, showing detection rate as a function of exposure time and number of mussels present.
**Figure 3.3** Species detection in samples collected from water bodies inhabited (or not) by the golden mussel. Results of Shisanling Reservoir consist of the first sampling (sites 1-3) and the second (field application) sampling (sites 4-5). Dash line indicates no detection.
### Table S3.1 Sequences and amplicon size of the 13 primers and their detection limit.

<table>
<thead>
<tr>
<th>ID</th>
<th>Primer sequence</th>
<th>Detection limit (ng)</th>
<th>Amplicon size/bp</th>
</tr>
</thead>
</table>
| A  | F: GGACTTTTTTTATTGTCTATGC  
    R: CCACGCATTTTCTTTAACAG | $4 \times 10^{-2}$ | 206 |
| B  | F: AGAAACCAGCAGTTGACATAG  
    R: CCATTAATAATAGGGGCAGTAGATTTTG | $4 \times 10^{-6}$ | 197 |
| C  | F: CCATTAATAAATAGGGGCAGTAGATTTTG  
    R: CCACGCATTTTCTTTAACAGG | $4 \times 10^{-2}$ | 286 |
| D  | F: GTCTATGCATGTAGAGGGGATG  
    R: AACGCTCACCACGCATTTC | $4 \times 10^{-2}$ | 200 |
| E  | F: AGAACCCCAGCAGTTGACAT  
    R: ACGCTCACCACGCATTTC | $4 \times 10^{-3}$ | 127 |
| F  | F: TGATGCTCATAGAACCAGG  
    R: ATATTTAACGCTCACCACGC | $4 \times 10^{-3}$ | 143 |
| G  | F: TTGATTCCATTAATAATAGGGGCA  
    R: AACGCTCACCACGCATTTC | $4 \times 10^{-4}$ | 298 |
| H  | F: TACTGGTTTCTTTTTTTTTTTTAG  
    R: AATATATTACGCTCACCACGC | $4 \times 10^{-3}$ | 245 |
| I  | F: CTGGGTTTGTCTGGAAACTGGA  
    R: CCACTCCCTACATGCATAGAC | $4 \times 10^{-4}$ | 286 |
| J  | F: TGTGTTAATTCGTTTAGAGTGCAC  
    R: CGTGGAAAAATCAAATCCACTG | $4 \times 10^{-3}$ | 186 |
| K  | F: TAGCTTTTGATGCTCATAAGAACCACCCAG  
    R: ACGCTCACCACGCATTTTCTTTAACAGG | $4 \times 10^{-3}$ | 142 |
| L  | F: GTTGCTGGAACCTGGATTAATGATG  
    R: ACTCCACCCTCTACATGCATAG | $4 \times 10^{-4}$ | 285 |
| M  | F: GGTTGTCTGGGAACCTGGATTAG  
    R: CCACTCCACCCTCTACAT | $4 \times 10^{-2}$ | 299 |
4.1 Introduction

Accurately detecting rare species - such as newly introduced nonindigenous species (NIS) or endangered native species - is critical for both conservation and management. Imperfect detection through either false positive or false negative results impedes these efforts, particularly with respect to rapid response to NIS incursions. However, detecting these species is challenging either because of their small population size and/or geographically-constrained distribution (vander Zanden et al. 2010; Simberloff et al. 2013; Branstrator et al. 2017; Robertson et al. 2017).

Environmental DNA (eDNA) refers to DNA released by organisms into their environment and is distributed where species currently or previously exist or where it is advected from these sources. eDNA can be directly extracted from bulk environmental samples and thus can be targeted using properly designed PCR primers (see Taberlet et al. 2012). eDNA is particularly useful for fast, sensitive and accurate species discrimination at low abundance (Jerde et al. 2011; Bohmann et al. 2014; Rees et al. 2014; Zhan & MacIsaac 2015). This feature has resulted in deployment of eDNA-based methods as a sensitive detection tool for a broad variety of aquatic species (e.g., Jerde et al. 2011; Boothroyd et al. 2016; Agersnap et al. 2017; Jackson et al. 2017; Torresdal et al. 2017; Voros et al. 2017). Despite this, eDNA-based techniques are immature, and

---

technical limitations must be considered when planning to employ these tools (Wilcox et al. 2013; Goldberg et al. 2014; Deiner et al. 2015; Goldberg et al. 2016).

Technical problems may complicate interpretation of eDNA results (Rees et al. 2014). For example, cross-contamination during sample collection, transport, or laboratory preparation may cause false positive results (i.e., target NIS is absent but DNA is detected in samples; Goldberg et al. 2016), while false negatives (i.e., target NIS is present but DNA is not detected) can occur if inhibitors are present in eDNA used as PCR templates (Jane et al. 2015) or if PCR primers have insufficient sensitivity (Wilcox et al. 2013; Xiong et al. 2016). It is imperative that detection programs have a low false negative rate given that they may delay recognition of, and rapid response to, presence of NIS, or may fail to detect an endangered species. According to Goldberg et al. (2016), an eDNA-based survey has two primary tasks: eDNA retrieval (e.g., sample collection and DNA extraction) and eDNA amplification (e.g., inhibitor removal and PCR). Many studies have focused on the former to improve detection rate (Renshaw et al. 2014; Takahara et al. 2014; Deiner et al. 2015; Spens et al. 2016; Hinlo et al. 2017; Xia et al. 2018), while attention has rarely been paid to the latter. Given that eDNA is often found in trace amounts (Furlan et al. 2016), robust PCR methods are essential to eDNA-based studies.

At present, conventional PCR (cPCR) and real-time quantitative PCR (qPCR) are the two major approaches used in eDNA-based species detection. Droplet digital PCR (ddPCR) has been suggested to be more sensitive than both, though it currently has limited use owing to cost and operational complexity (Nathan et al. 2014; Doi et al. 2015). A review of the literature revealed that 37% and 61% of eDNA studies employed
cPCR and qPCR, respectively, for aquatic species detection (Z. Xia, unpublished). It has been suggested that qPCR, which is a quantitative or semi-quantitative method, is the more sensitive method (Balasingham et al. 2017), although cPCR is more readily available to most molecular laboratories. This availability lends itself to wider use in rare species detection (Ojaveer et al. 2014; Roy et al. 2015; Ricciardi et al. 2017), as it is cost-efficient and can be very sensitive (e.g., Jerde et al. 2011). Ideally, a robust method for environmental samples should maintain sensitivity for samples obtained from different sources. Therefore, a comparison of the two most widely used PCR methods for samples from different sources may assist in future method selection for rare species detection. To our knowledge, however, this has not been well explored although several studies have discussed detection probability for eDNA samples using both methods. For example, Nathan et al. (2014) quantified eDNA signals using cPCR, qPCR and ddPCR from mesocosm aquaria and observed 100% detection of target species across all platforms; however, they failed to distinguish detection power of cPCR and qPCR. In another study, Piggott (2016) observed a higher detection rate of fish from dam water samples using qPCR than cPCR, though with limited sample sources. Additional empirical evidence derived from various systems is critical to guide future method selection.

In this study, we present comparisons cPCR with qPCR to detect a highly invasive mollusk, the golden mussel *Limnoperna fortunei*, from environmental water samples. First, we determined the limit of detection (LoD) of each PCR method under optimal conditions using total genomic DNA. Subsequently, we tested water samples from both laboratory aquaria and natural irrigation channels containing target DNA and calculated false negative rate of each method while varying sample replication. Finally, we
calculated quantification level of qPCR among above samples which were expected with different complexity, and compared species detectability using both methods to explore their performance in detection of different samples.

4.2 Materials and Methods

Sample collection and DNA extraction

Animals used in this study were collected from the Danjiangkou Reservoir, China (32°39'0"N, 111°41'15"E) and reared in a 60 L tank at 24°C before use. We used water samples maintained in laboratory aquaria and from the natural environment to test the two PCR methods. To prepare laboratory samples, we reared a golden mussel clump (12 adult individuals) at 24°C in a 15 L well-aerated aquarium for 24 h. We then removed animals from the tank and stopped aeration. The tank was left undisturbed for 12h before we began to collect water samples. Three 50 mL water samples were collected from the surface layer (~10 cm) of the aquarium, using separate 50 mL syringes for each replicate. We sampled at 11 time points (Table S4.1, supplementary information) over the course of a week, yielding 33 samples.

To prepare natural water samples, we sampled three irrigation channels in Dengzhou, China (Fig. 4.1). These channels were expected to contain eDNA of the golden mussel since the species was recorded in the vicinity in a preliminary field survey. Water source in each channel was controlled by a discharge gate at its source (Fig. 4.1). The discharge gates A and C were open while gate B was closed during sampling. Average water velocity was about 0.5 and 0.2 m s\(^{-1}\) in channels A and C, respectively, while channel B was static as the discharge gate B was completely closed. Water depth of
channels A, B, and C were about 1.8, 0.4, and 0.6 m, respectively. Sample collection order was channel C, B, and then A, and always from the most downstream to most upstream sites. We collected three 100 mL water samples from the surface layer (~20 cm) at each site (n = 17), yielding a total of 51 samples. All samples were transported on ice to the laboratory within 24 h of collection, and each was filtered onto a cellulose acetate microporous membrane filter (0.45 μm pore size). Each filter was cut in half and separately stored in a 2 mL centrifuge tube at -20°C until DNA extraction.

Total genomic DNA was extracted from fresh tissue of golden mussel using the DNeasy Blood & Tissue Kit (Qiagen). A randomly selected half-filter for each sample was extracted using the phenol-chloroform-isoamyl alcohol (PCI) method of Renshaw et al. (2014). Original DNA extracts were diluted 1:10 prior to use in PCR to reduce potential influence of PCR inhibitors (McKee et al. 2015).

cPCR analyses

We used a species-specific primer pair developed by Xia et al. (2018) to target a 197 bp fragment of the mitochondrial cytochrome c oxidase subunit I (COI) gene of the golden mussel. We ran 20 μL PCR mix following the methods detailed in Xia et al. (2018) with minor revisions: 5 μL template DNA was used in each reaction and 58°C was applied as the annealing temperature in this study. PCR products were visualized on 1.5% agarose gels using an automatic gelatin image analysis system (JiaPeng, Shanghai, China) and target bands were identified by eye. The LoD of the cPCR was tested using 10× serial dilutions of the total genomic DNA with concentration of 1.0×10⁰-10⁻⁸ ng μL⁻¹. A total of 10 replicates for each concentration was applied and the LoD was defined as
the lowest concentration returning at least one positive replicate which is widely used (e.g., Agersnap et al. 2017). We Sanger-sequenced four random positive amplicons of the field samples to confirm specificity of our primers, which was identified as species-specific in a previous study (Xia et al. 2018).

**qPCR analyses**

We used linear regression of quantification cycle (Cq) on DNA concentration (i.e., Log quantity) by amplifying the same serial dilutions of total genomic DNA mentioned above. Five replicates for each concentration were applied to construct the standard curve, and five no-template-controls (NTC) using double deionized water (ddH2O) were applied on the same 96-well plate to act as negative controls. We ran instrumental default 20 μL PCR mix containing 1× SYBR Green master mix (Roche Applied Science, Germany), 0.4 μM each primer and 5.0 μL DNA template (i.e., 1:10 diluted eDNA) on a LightCycler® 96 Instrument (Roche Applied Science, Germany). The thermal profile contained 60 s pre-incubation (95°C), followed by 50 cycles of 10 s for denaturation (95°C), 20 s for annealing (62°C) and 30 s for extension (72°C), followed by 10 min for final extension. Our primer pair can successfully amplify golden mussel at an annealing temperature from 45-65°C (Xia et al. 2018), and we used 58°C for cPCR and 62°C for qPCR, respectively, as they were optimal under respective cycling conditions. A melting analysis (95°C/10 s, 65°C/60 s, 97°C/1 s) was conducted following the amplification to generate a melting curve for PCR product in each well. The LoD of qPCR was identified as the lowest concentration producing at least one positive repeat out of the five repeats.
After qPCR, all melting curves were examined prior to the use of the returned Cq values by the built-in software. Specific amplification of our target species was characterized by a peak at the correct melting temperature (Smith & Osborn 2009; Peñarrubia et al. 2016), which was generated from amplification of a high concentration of total genomic DNA (e.g., 1.0 ng μL\(^{-1}\)). The Cq values returned from specific amplifications were identified as valid when the corresponding melting curves were normally distributed; otherwise the Cq values were dismissed (invalid Cq). To plot the standard curve, only serial dilutions of the total genomic DNA which generating ≥3 valid Cq values were considered. The corresponding efficiency of qPCR was calculated by the built-in software and descriptors of the standard curve were reported following Smith & Osborn (2009).

All amplification results of water samples underwent the same procedure as the standard curve prior to the use of Cq. Specifically, for those samples which returned positive amplifications but invalid Cq values (i.e., their melting curves were skewed or peaked at the NTC melting temperature), new Cq values were assigned to them according to the shape of the melting curves. Concentrations of water samples were then calculated based on the standard curve. The limit of quantification (LoQ) refers to the lowest concentration where the target species can be reliably quantified (Armbruster & Pry, 2008) and we defined it as the lowest concentration returning all positive replicates according to Agersnap et al. (2017). A linear regression model was applied to test the relationship between eDNA concentration (i.e., Cq) in irrigation channels and the distance to water source (i.e., discharge gate).
We also tested the importance of collecting replicate samples per time-point/site to reduce false negative results. We calculated the false negative rate when collecting between one and three replicates per time-point/site, using the scenario with the highest detection rate as a baseline. For the one-sample scenario, each sample was considered as a replicate. Alternatively, every possible two-sample combination was assessed in the two-sample scenario. For both PCR methods, all laboratory and field samples that initially failed to amplify underwent a second amplification and the results of both amplification attempts were combined to calculate the detection rate. One sampling time-point/site was considered a positive detection if any replicate tested positive.

**Quality control**

To prevent cross-contamination during sample collection, we used new bottles for water sample collection. Two bottles filled with deionized water and transported with sampling bottles during each sampling trip served as sampling controls. In the laboratory, all non-disposable equipment (i.e., forceps, scissors, beakers, syringes, filtration platform) involved in sample collection, filtration, and DNA extraction were treated using 10% commercial bleach for a minimum of 10 minutes before use to destroy residual DNA, which followed by thorough rinse with deionized water to remove the bleach. Blank controls were incorporated during the process of water sample filtration, and negative controls using ddH2O were included in all PCRs to monitor contaminations in laboratory practice.
4.3 Results

Limit of detection and quantification

The LoD was tested at $1 \times 10^{-6}$ and $1 \times 10^{-7}$ ng $\mu$L$^{-1}$ for cPCR and qPCR, respectively (Table S4.2, supplementary information), indicating higher sensitivity of the latter than the former. However, no amplification difference between the two methods was identified at higher DNA concentrations (e.g., $1 \times 10^{-5}$ ng $\mu$L$^{-1}$). For qPCR, one of five NTC replicates exhibited amplification signals (Cq: 39.16) with a melting temperature of 77-78°C (Fig. S4.1 lower, supplementary information). All high concentrations (i.e., $>1 \times 10^{-5}$ ng $\mu$L$^{-1}$) of genomic DNA returned valid Cq values with a melting temperature of 79-80°C (Fig. S4.1 upper, supplementary information). Low concentrations (i.e., $\leq 1 \times 10^{-5}$ ng $\mu$L$^{-1}$) were partially amplified, returning either valid (i.e., positive amplifications with normal distributed melting curves), invalid (i.e., skewed melting curves or NTC amplifications), or no Cq values (i.e., no amplification signals). The standard curve (Fig. S4.2, supplementary information) was plotted using serial dilution of $1.0 \times 10^{0}$-$10^{-5}$ ng $\mu$L$^{-1}$ in which three valid Cq values were returned at $1.0 \times 10^{-5}$ ng $\mu$L$^{-1}$ and five valid Cq values at higher concentrations (Table S4.2, supplementary information). Amplification efficiency of qPCR was 98%. The LoQ of total genomic DNA of qPCR was identified as $1.0 \times 10^{-4}$ ng $\mu$L$^{-1}$.

Detection of laboratory and field water samples

All positive amplifications of water samples (except the ones that exclusively exhibited NTC fluorescence signals) demonstrated species-specific. We assigned 33, 34, and 35 as Cq to those water samples which exhibited positive amplification of target
species but returned skewed melting curves (Fig. S4.3, supplementary information). These values were assigned to ensure that they were at least 3.3 fewer than those from NTC (Smith & Osborn 2009), and to guarantee an approximately continuous distribution of sample concentrations. All sampling controls and laboratory blanks demonstrated no amplifications of the target species by either PCR method throughout this study, and four randomly sequenced samples returned correct amplification of golden mussel from the field samples.

qPCR achieved a coherent higher detection rate than cPCR in both laboratory (100% vs. 87.9%) and field (68.6% vs. 47.1%) sample replicates (Fig. 4.2 a), resulting in five more sites detected positive in water channels (Fig. 4.1) by the former. For those sample replicates that were assigned Cq values, 83.3% of laboratory samples (n = 12) and 40% of field samples (n = 15) were also detected positive using cPCR (Fig. 4.2 b). Positive detections by cPCR were always a subset of those by qPCR. We found significant differences among quantifying total genomic DNA, laboratory aquaria, and field samples by qPCR by comparing the three lowest concentration (or three highest valid Cq values) of each group (Fig. 4.3). Specifically, total genomic DNA could be quantified to a significantly lower level \(10^{-4.28 \pm 0.13}\) ng than either laboratory \(10^{-3.03 \pm 0.06}\) ng or field samples \(10^{-2.92 \pm 0.06}\) ng \((F_{2,6} = 218, P < 0.001, \text{One-Way ANOVA})\). Furthermore, laboratory samples could be quantified to a significantly lower amount than field samples \(t_{4} = -2.273, P = 0.043, \text{one-tailed}\).

False negative detections were observed using both PCR methods when only one replicate sample collected, though decreased rates can be obtained by collecting more sample replicates (Fig. 4.4). Specifically, the false negative rate of cPCR decreased from
9.1% to 0%, and from 42.9% to 35.7% when sample replicates increased from one to three for laboratory and for field samples, respectively. Cq values were positively correlated with distance from the water source in channels A (Fig. 4.5 upper, \( P < 0.001 \)) and C (Fig. 4.5 middle, \( P = 0.03 \)) or their combination (Fig. 4.5 lower, \( P < 0.001 \)), indicating a decrease in DNA concentration with distance downstream.

4.4 Discussion

cPCR and qPCR methods are essentially the same with respect to amplifying target fragments (Smith & Osborn 2009). An important reason why qPCR was suggested to be more sensitive than cPCR is that different methods are utilized to detect PCR products; the former detects PCR product on-site by measuring fluorescence in each single PCR plate well, providing higher sensitivity than the ethidium bromide-stained, gel-based detection under ultraviolet light used in this study. Besides, qPCR can exclude ambiguity of positive/negative interpretation which may cause bias in cPCR (Nanthan et al. 2014). We determined a lower LoD with qPCR than cPCR (i.e., \( 1.0 \times 10^{-7} \) vs. \( 1.0 \times 10^{-6} \) ng \( \mu \text{L}^{-1} \)) which reflected the advantage using qPCR, while both methods exhibited 100% successful amplification at higher concentrations (\( \geq 1 \times 10^{-4} \) ng \( \mu \text{L}^{-1} \)). This is consistent with previous studies that conducted species detection in laboratory aquaria (e.g., Nanthan et al. 2014), indicating that detection probability of cPCR and qPCR may differ only at low concentrations. More effort is required to optimize PCR protocols or to improve primers design to reduce possible dimers to push detection limit to even lower levels. We used a 10× dilution to prepare varying total genomic DNA concentration and only limited amplification success were observed in low concentrations (i.e., \( 1 \times 10^{-5} - 10^{-7} \) ng \( \mu \text{L}^{-1} \)).
ng μL⁻¹). In future studies, a more refined dilution series (e.g., 2×) could be used to determine a refined LoD difference of both PCR methods.

qPCR achieved a higher detection rate for water samples than cPCR (Fig. 4.2 a), consistent with observations in previous studies (e.g., Piggott 2016), reflecting the higher sensitivity (or lower LoD) of the former. In addition to LoD difference, PCR inhibitors, which occur widely in environmental samples (McKee et al. 2015) may also contribute to detection rate difference between the two methods. PCR inhibitors such as humic acid or non-target species DNA may impact the final quality of eDNA (Wilson 1997; Pedersen et al. 2015), affecting PCR efficiency. Relative to total genomic DNA, DNA in environment samples may have a more uncertain fate owing to various factors such as season, UV, pH, temperature, substrate type, and downstream transport (Jane et al. 2015; Strickler et al. 2015; Buxton et al. 2017) and will likely contain higher amounts of impurities that inhibit amplification and result in lower PCR efficiency (Pedersen et al. 2015). This view is supported by the finding that target DNA can be quantified (i.e., valid Cq values returned) to an increased level from total genomic DNA versus laboratory or field samples using qPCR (Fig. 4.3). We expect that both cPCR and qPCR may suffer from inhibition in the same manner; however, we observed a greater detection rate difference between methods for all sample replicates from field than from laboratory samples (21.5% vs. 12.1%; Fig. 4.2 a). Furthermore, for the subset samples that were assigned Cq values due to skewed melting curves, a greater detection rate difference (60% vs. 16.7%) was observed in field samples (Fig. 4.2 b). This additional evidence is consistent with the view that sample complexity may affect PCR success and qPCR is more tolerant than cPCR to inhibitors owing to its more sensitive detection mechanisms (Smith & Osborn,
This observation is consistent with findings of Doi et al. (2015) who studied qPCR and ddPCR. It should be acknowledged that the master mix used in each PCR method may also affect detection efficiency (Jane et al. 2015) and contribute to the detection differences. We tried to reduce inhibitors by using diluted eDNA extracts (McKee et al. 2015) though we were unable to identify and quantify inhibitors of different samples in this study. Future studies are needed to assess impact of eDNA complexity (or presence of inhibitors) on detection performance for different PCR methods (Wilson 1997; Dingle et al. 2013), and to explore more efficient ways to eliminate inhibitors (e.g., environmental mix) without dilution as it may reduce target DNA to undetectable levels and cause false negatives (Buxton et al. 2017).

A critical concern in the application of eDNA methods to detect rare species is occurrence of false negatives (Ficetola et al. 2015). We observed a higher detection rate of qPCR than cPCR, suggesting that the former should be embraced in rare species management since it was more sensitive and less prone to false negatives. A number of avenues exist to reduce false negatives including judicious deployment of replicates in field sampling and in the laboratory (Piaggio et al. 2014; Pedersen et al. 2015) and the use of highly sensitive PCR methods or probes (Doi et al. 2015; Xia et al. 2018). We found that the false negative rate was inversely related to the number of replicates used per time-point/site (Fig. 4.4). This finding is consistent with other studies (Ficetola et al. 2015; Furlan et al. 2016) and highlights the importance of enhanced sampling effort to reduce false negatives. In this study, one replicate was sufficient to demonstrate the species presence/absence in laboratory samples, while three replicates were required for field samples (Fig. 4.4). We used three replicates as our baseline to calculate false
negative rate, which reflected the true rate of samples from laboratory aquaria and channels A and C as they were detected in 100% of sites. However, estimation of false negative rate for samples from channel B was difficult as both methods detected at only a single site. Given that many factors may cause failed detection (see Darling & Mahon 2011), estimation of false negative rate is difficult when detection rate with a baseline is less than 100%.

We found that eDNA concentration in channels A and C decreased with distance from the source (Fig. 4.1, gate A), consistent with other studies in flowing systems (Thomsen et al. 2012; Pilliod et al. 2013; Balasingham et al. 2016; Shogren et al. 2017). Contributors to this distribution pattern in lotic systems include facilitated degradation (Thomsen et al. 2012), dilution (Balasingham et al. 2016), and particle settlement (Jane et al. 2015; Xia et al. 2018). Only one sample replicate was tested positive at very downstream sites (i.e., C7 & C8, Figs 4.1 & 4.5), indicating limited detection probability of our method. We observed higher concentrations at sites C1-C3 than A5-A6 (Figs. 4.1 & 4.5) even though the former sites are located downstream of the latter. Two factors may explain this pattern. First, water flow through gate C (Fig. 4.1) may have facilitated particle resuspension, adding eDNA to the surface layer. Secondly, water entering channel C through gate C (Fig. 4.1) was from the deeper - and possibly eDNA enriched - layer in channel A, than in the surface layer at sites A5-A6. Regression of Cq against transport distance in channel C explained less variance (i.e., lower $R^2$) than in channel A (Fig. 4.5). This is likely because channel C is more vulnerable to human disturbance (e.g., irrigation drainage) and has higher structural heterogeneity within the channel (e.g., bottom plant growth) than channel A, as the former is smaller and shallower. However,
the declining trend of eDNA with flow distance was significant when channels A and C were combined (Fig. 4.5, lower), indicating that eDNA downstream transport may depend on water flow and spatial scale (Deiner & Altermatt 2014; Shogren et al. 2017).
4.5 References


Peñarrubia, L., Alcaraz, C., bij de Vaate, A., Sanz, N., Pla, C., Vidal, O., & Vinas, J. (2016). Validated methodology for quantifying infestation levels of dreissenid mussels in environmental DNA (eDNA). *Scientific Reports*, 6, article 39067. Doi: 10.1038/srep39067.


Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., ...


Figure 4.1 Map of sampling sites in the three irrigation channels (A, n = 6; B, n = 4; C, n = 7), showing the location of each site and detection results of golden mussel by both cPCR and qPCR. Arrows indicate the direction of water flow. Three replicate samples were collected per site, and sampling was carried out from downstream to upstream. Inset indicates location of the study area (asterisk), and dotted line indicates boundary of Henan Province and Hubei Province, China.
Figure 4.2 Detection rate of (a) all replicate samples from laboratory aquaria (n = 33) and field (n = 51) and (b) a subset of the former (n = 12) and latter (n = 15) in which Cq values were assigned to samples owing to skewed melting curves.
Figure 4.3 Mean (±S.D.) of three lowest quantities (solid circle) and their valid Cq (quantification cycle) values (bar) of target DNA detected from total genomic DNA, laboratory samples, and field samples, respectively, using qPCR. Cq refers to the number of cycles required for fluorescent signals to reach a threshold. Different letters indicate significant differences ($p < 0.05$).
Figure 4.4 False negative rate using one, two, and three replicates in the laboratory (upper) and in the field (lower) using cPCR (grey bar) and qPCR (white bar). Dashed line indicates 100% positive detections for laboratory samples (upper) and positive detections for field samples (lower) determined by qPCR when three replicates were used.
Figure 4.5 Linear regression of Cq against distance to water source (gate A) for water samples collected from channel A (upper, $P < 0.001$), channel C (middle, $P = 0.03$), and combination (lower, $P < 0.001$). Each circle indicates a replicate showing positive detection of golden mussel by qPCR, and the thicker circles indicate two overlapped replicates. Note that only one or two replicates were available for some sites.
4.6 Supplementary Information

Table S4.1 Sampling scheme of laboratory water samples. Time interval was 12 hours for days with two sampling points and 24 hours for days with one sampling point.

<table>
<thead>
<tr>
<th>Day</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td># of sampling points</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Table S4.2 Determination of limit of detection (LoD) of cPCR (conventional PCR, upper) and qPCR (quantitative PCR, lower). Concentration of total genomic DNA used to determine the LoD for each method was highlighted. Note that cPCR had 10 replicates for each concentration, and qPCR had five replicates for each concentration.

<table>
<thead>
<tr>
<th>Concentration (1.0×)</th>
<th>10^0</th>
<th>10^-1</th>
<th>10^-2</th>
<th>10^-3</th>
<th>10^-4</th>
<th>10^-5</th>
<th>10^-6</th>
<th>10^-7</th>
<th>10^-8</th>
</tr>
</thead>
<tbody>
<tr>
<td># of positive amplification (cPCR)</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Concentration (1.0×)</td>
<td>10^0</td>
<td>10^-1</td>
<td>10^-2</td>
<td>10^-3</td>
<td>10^-4</td>
<td>10^-5</td>
<td>10^-6</td>
<td>10^-7</td>
<td>10^-8</td>
</tr>
<tr>
<td># of positive amplification (qPCR)</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Figure S4.1 Melting curve of qPCR amplifying high concentration of genomic DNA (upper) and No-Template-Control (NTC, lower), showing melting temperature 79 - 80 °C for *Limnoperna fortunei* and 77 - 78 °C for NTC, respectively.
**Figure S4.2** Standard curve of qPCR, demonstrating the linear regression relationship between Cq and quantity of total genomic DNA (Log Quantity) used in each reaction. Each point is the mean (± SD) of five replicates, except for the lowest concentration which only had three valid Cq values.
Figure S4.3 Melting curve of qPCR amplifying unknown samples, showing positive amplification of *Limnoperna fortunei* but skewed toward NTC. Cq 33, 34, and 35 were assigned to amplifications similar with A, B, and C, respectively.
5.1 Introduction

Invasive species cause ecological and economic damage and even threaten human health (Holle & Simberloff 2005; Lockwood et al. 2005). Management of invasive species is often most effective at the prevention stage (e.g., Leung et al. 2002). Preventing invasions can be successful by focusing on elimination or dramatically reducing the frequency of introduction events and population size in each introduction event (collectively, propagule pressure) (Holle & Simberloff 2005; Lockwood et al. 2005). Early detection refers to the ability to detect an introduced species at very low population size - ideally just after introduction of a population with low propagule pressure - and is critical to subsequent management measures (e.g., prevention, eradication, containment).

Traditional detection methods rely on some variant of “catch and look” of target species and are often ineffective with newly introduced species, mainly due to sampling difficulty (Harvey et al. 2009). In addition, immature individuals or species possessing cryptic morphological life stages may be exceptionally difficult to identify even for trained taxonomists, raising the possibility of misidentification (Hebert et al. 2003). The advent of DNA barcoding provides an opportunity to identify species based upon unique nucleotide sequences instead of morphological traits, often with enhanced accuracy (Hebert et al. 2003). The ubiquitous nature of environmental DNA (eDNA) in bulk environmental samples provides a cost-efficient way to detect target species by analyzing their DNA shed into the surrounding environment rather than isolating the organisms.
themselves (Thomsen et al. 2012; Bohmann et al. 2014). The combination of DNA barcoding and eDNA sampling has proven very effective in detecting low abundance species (Thomsen et al. 2012; Wilcox et al. 2013; Fukumoto et al. 2015; Xia et al. 2018b). In addition to presence/absence determination, eDNA methods have been used to estimate species abundance (e.g., Thomsen et al. 2012; Pilliod et al. 2013), though with conflicting results (e.g., Rice et al. 2018).

The golden mussel *Limnoperna fortunei* is a freshwater bivalve native to south China (e.g., the Pearl River basin) and several neighboring countries. It spread widely in SE Asia (e.g., northern spread to Yangtze River Basin, China, Japan, South Korea) and also in more distant locales (e.g., Argentina, Brazil) (Boltovskoy 2015). In invaded environments, mussel distributions typically expanded quickly with pronounced ecosystem consequences (Boltovskoy & Correa 2015). Golden mussel life cycle comprises a free-swimming planktonic larval stage and a predominantly sessile adult stage, which greatly enhances the species’ opportunity for transport elsewhere. For example, the transcontinental invasion of South America likely occurred via introduction of planktonic juveniles through ballast water discharge, while inland spread was likely achieved mainly by hull fouling by adults (Nakano et al. 2015). Other anthropogenic activities such as dam construction and water diversion projects also facilitate spread (Nakano et al. 2015; Xu et al. 2015; Zhan et al. 2015). The latter mode of spread may be particularly potent as water diversion projects potentially transport trillions of planktonic propagules to new locations.

The South to North Water Diversion Project (SNWDP) (central route) in China (hereafter SNWDP), which connects an invaded reservoir in the South (i.e., source) with
uninvaded, recipient ones in the North and represents a potential “invasion highway” for species dispersal (Zhan et al. 2015). In such a ‘lake-stream’ system, the presence of a colonized ‘lake’ at upstream has proven critical for the downstream colonization of zebra mussel because the stream system may serve as a sink of population (e.g., Horvath et al. 1996; Bobeldyk et al. 2005). In the case of SNWDP, the inhabited reservoir is indeed a source of golden mussel, wherein the spread risk is expected to vary temporally since planktonic larvae production is often seasonal (Nakano et al. 2010). Thus, the post-spawning season may present an ideal window of time to detect the golden mussel because free-living larvae may add quality DNA to that excreted by sessile adults, allowing for detection at substantially further transport distance.

In the present study, we aim to test if we can use eDNA as a proxy to identify the spawning season of golden mussels in running water ecosystems. We also explore environmental factors that may influence eDNA concentration in the canal system. To answer the first question, we sampled multiple sites along the main canal of the SNWDP over a course covering the expected spawning season. For the second, we conducted refined eDNA samplings in a sub-canal of the SNWDP and explored factors affecting eDNA concentration. We hypothesize that the detection frequency of golden mussel eDNA in the main canal of SNWDP is consistent with the mussel’s spawning season and that the concentration of target DNA decreases with transport distance.

5.2 Materials and Methods

Study design

We collected water samples from the main canal of SNWDP from May 2016 through November 2017, consisting of two sampling phases. The first phase comprised
11 trips (May 2016-April 2017, though February 2017 was not sampled) covering six sites (M1-M6, Fig. 5.1 A), and the second (July-November 2017) included five trips and covered four sites (M1, M2, M3 & M5, Fig. 5.1 B). Each trip commenced ~10th of the month and finished within three days. The sub-canal sampling was conducted on May 30, 2018, covering 12 sites (S1-12, Fig. 5.1 C). This canal was selected because it originates from the main canal and has similar water chemistry to it.

**Main canal sampling**

Every trip sampling the main canal included two sites on the source reservoir (i.e., Danjiangkou Reservoir, 32°39'0"N, 111°41'15"E). Site M1 was located by the Danjiangkou Dam, while site M2 was ~500 m away toward the canal control gate outflow to the SNWDP, which was ~30 km from M1 (Fig. 5.1 B). Site M4 and M6 were not included in the second sampling phase. Water collections were made from upstream to downstream (i.e., M1 to M6, south to north). At each site, a 1-L water sample was filled in a new polyethylene terephthalate (PET) bottle, which was immediately placed in a portable fridge and remained ~4 °C until transported back to the laboratory. One bottle filled with distilled water was placed with samples and served as a sampling control. Water samples were filtered within 48h of collection of the last site (M6). Each water sample was evenly filtered onto two cellulose acetate microporous membrane filters (0.45 μm pore size) (i.e., 500 mL per filter) with exceptions for samples collected in June 2016 and July and August 2017 that were each filtered onto three filters (i.e., 333 mL per filter) due to filter clogging. Each filter was placed in a 2 mL centrifuge tube and stored at -80 °C until extraction.
**Sub-canal sampling**

The sub-canal in the present study is located in Dengzhou, Henan province of China, which is the first branch of the main canal after the canal head (Fig. 5.1 C). The canal side was comprised of concrete, while the bottom was soft sediment and occupied by submerged macrophytes. Golden mussels have been observed in this canal, but with very limited abundance and rarely attached on secured rocks (Xia et al. 2018a). Sampling was conducted from downstream to upstream (i.e., S12 to S1, Fig. 5.1 C). At each site, three 100-mL water samples were collected by submerging new PET bottles under the surface layer (~20 cm), and each sampling site was ~1 m deep and ~1 m off the side board. A 500-mL bottle was filled and sealed underwater (i.e., no air bubbles were introduced into the bottle) for suspended particle size partitioning. Water temperature was measured by a glass thermometer and water velocity by float method, respectively. Sampling control and sample preservation and transport were performed as above in the main canal sampling. Each sample was filtered onto a single filter within 24h of collection and stored at -80 °C until extraction.

**DNA extraction and PCR**

We used the phenol-chloroform-isoamyl alcohol (PCI) method of Renshaw et al. (2014) to extract eDNA from filters. For each filter, extracted DNA was dissolved with 50 μL double distilled water (ddH₂O). We diluted the original DNA extracts by 1:10 before PCR to reduce possible inhibitors (McKee et al. 2015). For main canal sampling, we performed DNA extraction for the first and second phases separately. Tissue-derived
total genomic DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen) and quantified using a Nanodrop 2000 spectrophotometer (Thermo Scientific).

SYBR Green-based quantitative PCR (qPCR) [detailed in Xia et al. (2018a)] was used to detect and quantify golden mussel DNA from extracted water samples. Consistent with DNA extraction, we conducted PCR for the first and second sampling phases in the main canal separately. For every sample, we ran PCR for the two (or three in few months as described in 2.2) filters separately because of space limitation in each PCR plate (i.e., 96-well plate), and we ran each batch twice to generate PCR replicates. Therefore, four (2 PCRs x 2 filters) or six (2 PCRs x 3 filters) replicates were conducted for each sample. To quantity target DNA, a standard curve using triplicate serial dilutions (i.e., 1:10) of the tissue-derived total genomic DNA was included in each PCR plate. The limit of detection (LoD) of the assay was determined as the lowest concentration of the total genomic DNA that yielded at least a single positive amplification of the replicates. Melting curves of PCR products were generated at the end of qPCR with default settings to confirm specific amplification of target species (Xia et al. 2018a). For a small portion of reactions that yielded mixed signals of no-template-control (NTC) and target species, an adjusted quantity was assigned to the respective sample according to the shape of its melting curve. Specifically, a coefficient of 0.5, 0.25, and 0.125 was assigned to the original quantity yielded by the system for samples with a strong, medium, and weak signals of target species (see Xia et al. 2018a for more details). Also, we sent four qPCR products with high concentrations for Sanger sequencing (Tsingke Biotech Ltd., Beijing, China) to further confirm specificity.
Quality control in laboratory

To prevent cross-contamination, we rinsed the body of sample bottles with distilled water and dried them with paper towels before use. All reusable tools (e.g., glassware, filtration platform, forceps, scissors, etc.) used in sample filtration were submerged in 10% commercial bleach for 20 min, followed by thorough rinse with distilled water between uses. Three NTCs using ddH$_2$O were included in each PCR plate, acting as negative controls.

Data collection and statistical analysis

Standard curves were used to calculate the quantities of samples (in ng total genomic DNA). For filters with ~333 mL water filtered, results were corrected to 500 mL to be consistent with most cases. The average quantity (500 mL for main canal samples and 100 mL for sub-canal samples) was used in the subsequent analysis, and the percentage of replicates detected positive was used to represent the detection probability of a site. The concentration of suspended particles with varying sizes in the sub-canal was partitioned using a particle counter (PAMAS Water Viewer, Germany), which was running with ddH$_2$O between every two samples to avoid cross-contamination. The mean of three reads of each sample was used in the subsequent analysis. The distance from the source of each sampling site was measured using the default ruler of path in google earth. In the case of the main canal, we specified 15 km to sites M1 and M2 as they were located within the source reservoir. We compiled the average daily minimum and maximum air temperature of the neighboring city of each sampling site in the main canal (http://www.tianqihoubao.com/).
We ran binary logistic linear regression of PCR call of target eDNA (i.e., positive or negative) in the main canal against the average daily minimum and maximum air temperature of each site to explore the influence of air temperature on the occurrence of golden mussel DNA. To explore factors influencing eDNA concentration in the main canal, we ran multiple regressions of eDNA quantity (dependent variable) against distance from source and average daily minimum air temperature (independent variables) for samples collected in June and September when all six sites were detected positive. Similarly, the distance from source, water velocity, and concentration of suspended particles (0-1 μm) was considered as explanatory variables in the sub-canal. We only included the 0-1 μm suspended particles in the above analysis as it was the best predictor of target DNA concentration amongst all size categories (see 3.3 for more details). Water temperature was excluded from the multiple regression due to a significant collinear relationship with distance from the source and was analyzed with simple linear regression instead. The analysis above was run with R 3.5.2 (R core team, 2018).

5.3 Results

Performance of qPCR assay

The limit of detection of the assay was identified as 1.2 x 10^{-7} ng of tissue-derived genomic DNA, which was achieved in two PCR plates (Table S5.1, supplementary information). Efficiencies of qPCR reactions ranged between 96% and 102% with $R^2$ of standard curves between 0.98 and 1.00 (Table S5.1, supplementary information), suitable for quantifying water samples. Melting peaks (i.e., temperature) of PCR products of the target species and NTC (if signals detected) were ~79.5 °C and ~77.0 °C, respectively.
and could be clearly distinguished from each other by eye (Fig. S5.1, supplementary information).

*Positive/negative distribution of eDNA in the main canal*

Golden mussel DNA was detected in eight of the 11 trips (i.e., 74%) in the first sampling phase (left panel) and all five trips (i.e., 100%) in the second (right panel), respectively (Fig. 5.2). A total of 95% (42 of 44) of sites detected positive were achieved between May and November. The remaining positives were detected in December and March of the first phase, at low DNA concentrations (Fig. 5.2). No sampling controls or NTCs produced contaminated signals (i.e., positives) throughout the study, and sequencing results confirmed the target species (Fig. S5.2, supplementary information).

For the first sampling phase, positive detections were concentrated in May through October (Fig. 5.2), though the frequency of positive detections (i.e., # of trips detected positive) declined significantly with distance from the source ($F_{1,4} = 11.7$, $R^2 = 0.745$, $P = 0.027$). The northernmost site M6 had the lowest detection frequency (i.e., two of five trips). This pattern, however, was not evident in the second sampling phase amongst the four sites (Fig. 5.2, right panel). Detection of golden mussel DNA was highly associated with average daily air temperature, significantly increased with the minimum air temperature ($P < 0.001$; Fig. 5.3), and the inflection point (i.e., 50:50 probability) was 13.2 °C. The maximum air temperature was not considered as it exhibited a strong collinear relationship with low ones ($R^2 = 0.967$, $P < 0.0001$) and had higher Akaike information criterion (AIC) value than the latter (95.27 vs. 91.25, Table S5.2, supplementary information). The distance to the source reservoir, however, had no
significant relationship with the positive/negative distribution of golden mussel DNA (Table S5.2, supplementary information).

**eDNA concentration in the main canal**

The concentration (Mean ± S.E.) of golden mussel DNA ranged between $9.8 ± 9 \times 10^{-8}$ and $9.4 ± 7.8 \times 10^{-2}$ ng per reaction, varying substantially amongst sites and months (Fig. 5.2). Only in June and September did all six sites have positive detections, though the DNA concentration exhibited different relationships with distance to the source reservoir between the two months. Specifically, the quantity of target DNA was significantly decreased with distance in June ($F_{1, 4} = 39.93$, $R^2 = 0.91$, $P = 0.003$) while insignificant in September ($F_{1, 4} = 0.15$, $R^2 = 0.04$, $P = 0.718$) (Table 5.1). Unlike the positive/negative distribution, average daily minimum air temperature had no statistically significant relationship with target DNA concentration in the main canal in both trips ($P = 0.742$).

**eDNA concentration in sub-canal**

Water temperature in the sub-canal was 19-21 °C and increased from upstream to downstream (i.e., site S1 to S12, Fig. 5.1 C). Water velocity was 0.17-0.67 m s$^{-1}$ and fluctuated amongst sites (Table S5.3). Golden mussel DNA was detected from all 12 sites in the sub-channel with concentration (Mean ± S.E.) ranging between $1.1 ± 0.3 \times 10^{-4}$ and $1.2 ± 0.6 \times 10^{-7}$ ng per reaction (Fig. 5.4), and was significantly ($F_{3, 8} = 35.58$, $R^2 = 0.930$, $P < 0.001$) related to distance from source ($P < 0.001$), the concentrations of suspended particles (0-1 μm size) ($P = 0.0016$), and their interactions ($P = 0.0023$) (Table 5.1).
Specifically, eDNA concentration decreased with both distance from source \((P < 0.001)\) and suspended particle concentration \((P = 0.002)\) (Table 5.1, Fig. 5.4 A). It was also negatively related to water temperature \((P = 0.007)\) (Table 5.1, Fig. 5.4 B), but not to water velocity.

5.4 Discussion

eDNA methods provide a revolutionary advance in accurate detection of rare species. However, detection probability is dependent on the availability of target eDNA and is affected by factors such as species traits, abundance, and environmental conditions (Pilliod et al. 2013; Strickler et al. 2015; Buxton et al. 2017). Zhan et al. (2015) suggested that the opening of SNWDP could facilitate invasion of golden mussels from southern China to the north. Here, we conducted eDNA sampling shortly after (~1.5 years) the SNWDP was opened in December 2014 and detected golden mussel DNA at sites (i.e., M5 & M6) far removed (>1000 km) from the putative source in Danjiankou Reservoir. This remarkable spread could be accounted for by DNA excreted by adult mussels living in Danjiankou Reservoir or by their offspring that had established downstream, by sperm and eggs released by these individuals, or by planktonic larvae produced by these individuals and subsequently advected downstream. In a low flow environment, adult mussels could deliver DNA to limited surroundings via siphonal jets (Nishizaki & Ackerman 2017), complicating their eDNA detection from surface water (e.g., Xia et al. 2018b). Also, DNA excreted by these sessile animals would also certainly be highly degraded within a short downstream transport distance. For example, eDNA of caged fishes became undetectable in as short as <2 km distance downstream in rivers (Jane et al. 2018b).
Similarly, sperm and eggs have a relatively short lifespan (e.g., 3.5 h at 26 °C, Boltovskoy 2015), and even with degradation would be highly unlikely to provide eDNA 1100 km downstream. Production of planktonic veligers, whose development and metamorphosis is temperature-dependent, could remain in the plankton for up to 20 days before settlement (Boltovskoy 2015), during which time they could be advected between 346 km (at 0.2 m s⁻¹) and 3460 km (at 2 m s⁻¹) downstream. The SNWDP waterway is not used for commercial or recreational boating, nor recreation. Thus the most plausible mechanism for the tremendous dispersal of golden mussel DNA is downstream transport of veligers.

**Impact of temperature on eDNA detection in the main canal**

The majority of positive detections were observed in warm months (Fig. 5.2), and the probability of positive detection rose as air temperature increased (Fig. 5.3), illustrating an evident temperature-induced onset of golden mussel DNA detection in SNWDP. Such temporally-dependent detection of species from eDNA samples has been reported for other rare species such as amphibians (e.g., Spear et al. 2015, Buxton et al. 2017), fishes (e.g., Turner et al. 2014; de Souza et al. 2016; Gingera et al. 2016; Xu et al. 2018) and reptiles (e.g., de Souza et al. 2016), which result from spawning (e.g., Gingera et al. 2016; Buxton et al. 2017; Xu et al. 2018) and enhanced activities (Souza et al. 2016).

Temperature plays a crucial role in inducing bivalves to spawn (Baba et al. 1999; Philippart et al. 2003). For example, it has been reported in Japan (e.g., Lake Ohshio and Lake Takenuma, ~36°N, 139°E; Nakano et al. 2010) and Argentina (e.g., Paraná de las Palmas river, ~34°S, 59°W; Cataldo & Boltovskoy 2000) that 16-17°C water temperature
is the proximate lower thermal limit triggering spawning of golden mussels. Studies on populations in south China (e.g., Hong Kong and Shenzhen: ~22°N, 114°E) demonstrated that golden mussels could spawn multiple times per year, with several peaks in warm months but limited reproduction in cool seasons (Xu et al. 2013). Populations in our study are located much farther north (i.e., M1-M6: ~32-39°N, 112-115°E). This may limit spawning of golden mussels to warm months only. Records indicates that water temperature in Danjiangkou Reservoir remains over ~16-17 °C for ~6.5 and ~4.5 months in surface and deep layers, respectively (Duan et al. 2018), which is consistent with the period that golden mussel DNA was detected at sites M1 and M2 in the present study (i.e., May-October, Fig. 5.2). Across the entire SNWDP, the probability of positive detections exceeded 50% (i.e., inflection point) when the average daily minimum air temperature was higher than 13.2 °C (Fig. 5.3).

Increased water temperature increases metabolic activity in bivalves (e.g., Alexander Jr, Thorp, & Fell 1994). In the case of golden mussels, enhanced filtering of suspended particles was observed in warm waters (Sylvester et al. 2005). Enhanced filtering at elevated temperature will allow golden mussels to shed more DNA via production of feces and pseudofeces into the environment. However, this may add very limited eDNA to the environment relative to spawning and larval production. Also, high temperature will increase degradation of feces-associated DNA (Strickler et al. 2015) but enhance activities of free-living individuals. Thus, we propose that spawning and larval production that occurs at elevated temperature substantially improves detection probability from water samples. This effect might be augmented by enhanced filtering activities of golden mussels at the main canal sites in warmer months. Two positive
detections were produced in cool months (i.e., M3 of December 2016 and M5 of March 2017), and they were likely due to resuspension of historical DNA from the previous months or dead animals that previously established.

Factors influencing eDNA concentration in canals

Golden mussel DNA in the sub-canal declined significantly with distance from the source (Table 5.1; Fig. 5.4), while the same pattern was identified in the main canal in June of the first sampling phase (Table 5.1, Fig. 5.2). These findings are consistent with existing studies in natural rivers (e.g., Jane et al. 2014; Balasingham et al. 2017), and with a previous study in the same system (Xia et al. 2018a). In flowing aquatic systems such as natural rivers and artificial waterways, many factors affect eDNA concentration. eDNA originating from a point source - such as caged animals (e.g., Jane et al. 2014; Balasingham et al. 2017) - will exhibit attenuation with transport distance due to a combination of degradation, dilution, and sedimentation, the latter of which has been suggested as a dominant carrier of eDNA in natural environments (Jane et al. 2014; Turner et al. 2015; Balasingham et al. 2017; Seymour et al. 2018). In contrast, no extra inflows exist in our study system, which would preclude dilution. Mortality of planktonic golden mussel larvae is often very high (~90%; Boltovskoy 2005), and live veligers may experience a sharp numerical decline (e.g., Horvath & Lamberti 1999), suggesting that these losses – combined with settling, metamorphosing individuals – could account for diminution of eDNA in surface waters.

In the sub-canal, concentrations of golden mussel DNA were well explained by fine suspended particles <15 μm, which is consistent with a previous study that partitioned
fish eDNA by sequential filtration of water samples through filters with varying pore sizes (Turner et al. 2014). The DNA concentration in the sub-canal dropped sharply after the first three sites (Fig. 5.4, A), suggesting rapid sedimentation or degradation of particles carrying enriched DNA. The concentration of fine particles marginally significant increased with distance ($P = 0.062$) and water velocity ($P = 0.068$), suggesting resuspension of sediments in the sub-canal. However, this did not seem to compensate for eDNA loss during downstream transport. Besides, water temperature significantly increased with distance in the sub-canal (Table S5.4, $P < 0.001$), suggesting a higher degradation rate at downstream than upstream sites (Strickler et al. 2015).

The estimated coefficient (i.e., linear regression) of eDNA concentration against transport distance in the main canal (June) was substantially lower than in sub-canal (Table 5.1), suggesting a lower attenuation rate of target eDNA in the former system. This may have occurred because of higher water velocity in this canal, which reduced both retention time and sedimentation rate of suspended particles. Also, established animals in the main canal may have postponed spawning time relative to the source population in the south, allowing them to add newly released larvae to the water and compensate for DNA losses. Further, eDNA is likely preserved better from south to north in the main canal than in the sub-canal because of their differing water temperature distributions. These conditions could collectively reduce the attenuation rate of target DNA in the main canal.

*Implications in managing golden mussels in the SNWDP*
The central route of the SNWDP has been in operation for ~1.5 years before the present study was initiated. It is likely that golden mussels were established in at least part of the canal (e.g., sites with reduced water velocity) during the first spawning season following its inception. The canal acts as a “corridor-like” extension of the Danjiangkou Reservoir, and any environment connected to the main canal via water flow will be vulnerable to invasion (e.g., Bobeldyk et al. 2005). In such a system, suitable downstream habitats will be successively invaded in turn by propagules produced in upstream bridgehead populations. Managing systems like this must control the introduction of propagules. This objective will be challenging in upstream locations receiving a massive infusion of larval propagules from the source (e.g., Horvath et al. 1996). Downstream sites would at first seem better protected given propagule attenuation with distance from source (e.g., Norvath & Lamberti 1999). However, if the same dynamics of reproduction and spread play out across multiple years, it may be expected that the species will utilize prior bridgehead populations to facilitate spread during subsequent years (e.g., Horvath et al. 1996; Estoup et al. 2010). Over the long term, the entire system with hydraulic connectivity will be placed at risk.
5.5 References


Renshaw, M. A., Olds, B. P., Jerde, C. L., McVeigh, A. M., & Lodge, D. M. (2014). The room temperature preservation of filtered environmental DNA samples and


Table 5.1 Summary of linear (or multiple linear) regressions of eDNA quantity (in Log$_{10}$ scale) as a function of distance from source (dist., km) in main canal (Main, June and September in the first phase of sampling), and of distance from source, ≤1 μm suspended particles (pa., # of particles per mL), and water temperature (temp., °C) in sub-canal. ‡ indicates variables used in the multiple regression. ** indicates significance at 0.01 level and *** at 0.001 level.

<table>
<thead>
<tr>
<th>Model</th>
<th>Est. intercept</th>
<th>Est. slope</th>
<th>$R^2$</th>
<th>AIC</th>
<th>$F$-value</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main_dist. (Jun.)</td>
<td>-2.528***</td>
<td>-0.002**</td>
<td>0.91</td>
<td>8.9</td>
<td>39.93</td>
<td>1,4</td>
</tr>
<tr>
<td>Main_dist. (Sep.)</td>
<td>-4.78***</td>
<td>0.0002</td>
<td>0.04</td>
<td>16.2</td>
<td>0.15</td>
<td>1,4</td>
</tr>
<tr>
<td>Sub_temp.</td>
<td>7.567</td>
<td>-0.65**</td>
<td>0.53</td>
<td>20.3</td>
<td>11.40</td>
<td>1,10</td>
</tr>
<tr>
<td>Sub_dist. ‡</td>
<td>-2.276**</td>
<td>-0.3***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub_pa. ‡</td>
<td>-7.86e-5**</td>
<td>0.93</td>
<td>0.93</td>
<td>1.4</td>
<td>35.38</td>
<td>3,8</td>
</tr>
<tr>
<td>Sub_dist. × pa. ‡</td>
<td>7.47e-6**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 5.1 Map of study area, showing location of central route of the South to North Water Diversion Project (SNWDP) in China (A, dashed rectangle), sampling sites in the main canal of SNWDP (B, red dots), and in the sub-channel (B, green dots), respectively. Both sites M1 and M2 (~30 km apart) were within the Danjiangkou Reservoir, and the asterisk indicates the proximate location of the sub-channel.
Figure 5.2 Quantity of golden mussel DNA (Mean ± S.E.) in 500 mL water sample in the main canal of the SNWDP and the average daily minimum low air temperature of the
neighboring city at each site. Horizontal dashed lines indicate limit of detection ($1.2 \times 10^{-7}$ ng), while the vertical solid line separates the first (left panel: sites M1-M6) and second (right panel: sites M1-3 & M5) phases of sampling. Panels of sampling sites (M1 to M6) are ordered from south to north in the SNWDP. An asterisk indicates a negative detection.
Figure 5.3 Logistic regression of probability of golden mussel DNA from the main canal as a function of average daily minimum air temperature. Actual detection results are presented as circles. The shaded band indicates ± one standard error of the prediction.
Figure 5.4 Relationship between concentration of golden mussel DNA and (A) distance from the source, concentration of suspended particles (≤1 µm), and their interactions; and (B) water temperature. Unit of ≤1 µm suspended particles is # of particles per mL. The solid line indicates fitted prediction and shaded band indicates ± one standard error of the prediction.
### 5.6 Supplementary Information

**Table S5.1** Standard curves in qPCR reactions.

<table>
<thead>
<tr>
<th>Samples</th>
<th>PCR replicate</th>
<th>Standard curve equation</th>
<th>Amplification efficiency</th>
<th>$R^2$</th>
<th>LoD (ng)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main_1&lt;sup&gt;st&lt;/sup&gt; phase</td>
<td>PCR1</td>
<td>$C_q = -3.2294 \times \text{Log}(\text{Quantity}) + 18.14$</td>
<td>102%</td>
<td>0.99</td>
<td>$1.2 \times 10^{-7}$</td>
</tr>
<tr>
<td></td>
<td>PCR2</td>
<td>$C_q = -3.4173 \times \text{Log}(\text{Quantity}) + 20.84$</td>
<td>98%</td>
<td>1.00</td>
<td>$1.2 \times 10^{-6}$</td>
</tr>
<tr>
<td>Main_2&lt;sup&gt;nd&lt;/sup&gt; phase</td>
<td>PCR1</td>
<td>$C_q = -3.5330 \times \text{Log}(\text{Quantity}) + 21.22$</td>
<td>96%</td>
<td>0.98</td>
<td>$1.2 \times 10^{-6}$</td>
</tr>
<tr>
<td></td>
<td>PCR2</td>
<td>$C_q = -3.4363 \times \text{Log}(\text{Quantity}) + 20.91$</td>
<td>98%</td>
<td>0.99</td>
<td>$1.2 \times 10^{-6}$</td>
</tr>
<tr>
<td>Sub-canal</td>
<td>PCR1</td>
<td>$C_q = -3.4541 \times \text{Log}(\text{Quantity}) + 17.96$</td>
<td>98%</td>
<td>1.00</td>
<td>$1.2 \times 10^{-7}$</td>
</tr>
<tr>
<td></td>
<td>PCR2</td>
<td>$C_q = -3.4307 \times \text{Log}(\text{Quantity}) + 18.34$</td>
<td>98%</td>
<td>1.00</td>
<td>$1.2 \times 10^{-6}$</td>
</tr>
</tbody>
</table>

*Cq*: Quantification cycles; *LoD*: limit of detection
Table S5.2 Summary of modelling results of eDNA presence/absence detections and target DNA concentrations (June and September) as a function of average minimum and maximum air temperature, and distance to source reservoir in main canal.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Independent variables (P-value)</th>
<th>AIC</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence / Absence of target DNA</td>
<td>minimum air temperature (&lt; 0.001***) + distance (0.196)</td>
<td>91.55</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>maximum air temperature (&lt; 0.001***) + distance (0.084)</td>
<td>94.10</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>minimum air temperature (&lt; 0.001***)</td>
<td>91.25</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>maximum air temperature (&lt; 0.001***)</td>
<td>95.27</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>distance to source (0.06)</td>
<td>119.50</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>minimum air temperature (&lt; 0.001***) + distance (0.641) + interactions (0.228)</td>
<td>92.06</td>
<td>85</td>
</tr>
</tbody>
</table>

AIC: Akaike information criterion; df: degree of freedom; Significance level: * 0.05, ** 0.01, *** 0.001
Table S5.3 Summary of environmental conditions of sampling sites in sub-canal.

<table>
<thead>
<tr>
<th>Site</th>
<th>Distance from source (km)</th>
<th>Water velocity (m/s)</th>
<th>Water temperature (°C)</th>
<th>0-1 µm particle (# of particle per mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.74</td>
<td>0.18</td>
<td>19</td>
<td>20939</td>
</tr>
<tr>
<td>S2</td>
<td>1.62</td>
<td>0.18</td>
<td>19</td>
<td>19266.6</td>
</tr>
<tr>
<td>S3</td>
<td>3.71</td>
<td>0.4</td>
<td>19</td>
<td>32071.8</td>
</tr>
<tr>
<td>S4</td>
<td>6.63</td>
<td>0.4</td>
<td>19</td>
<td>30477.4</td>
</tr>
<tr>
<td>S5</td>
<td>7.49</td>
<td>0.33</td>
<td>19</td>
<td>38353</td>
</tr>
<tr>
<td>S6</td>
<td>9.05</td>
<td>0.25</td>
<td>19.4</td>
<td>30275.4</td>
</tr>
<tr>
<td>S7</td>
<td>11.39</td>
<td>0.26</td>
<td>19.9</td>
<td>36858</td>
</tr>
<tr>
<td>S8</td>
<td>11.49</td>
<td>0.27</td>
<td>19.8</td>
<td>38372.6</td>
</tr>
<tr>
<td>S9</td>
<td>14.38</td>
<td>0.67</td>
<td>20</td>
<td>43137.25</td>
</tr>
<tr>
<td>S10</td>
<td>17.05</td>
<td>0.18</td>
<td>20.5</td>
<td>39122.75</td>
</tr>
<tr>
<td>S11</td>
<td>18.99</td>
<td>0.17</td>
<td>20.8</td>
<td>28909.6</td>
</tr>
<tr>
<td>S12</td>
<td>22.02</td>
<td>0.2</td>
<td>21</td>
<td>33307</td>
</tr>
</tbody>
</table>
Table S5.4 Summary of linear regressions of golden mussel DNA concentrations in sub-channel against concentrations of suspended particles with varying sizes. * indicates significance at 0.05 level.

<table>
<thead>
<tr>
<th>Particle size (μm)</th>
<th>Coefficient (standardized)</th>
<th>R square</th>
<th>t-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤1</td>
<td>0.614</td>
<td>0.377</td>
<td>2.457</td>
<td>0.034*</td>
</tr>
<tr>
<td>1-2</td>
<td>0.583</td>
<td>0.340</td>
<td>2.269</td>
<td>0.047*</td>
</tr>
<tr>
<td>2-4</td>
<td>0.572</td>
<td>0.327</td>
<td>2.204</td>
<td>0.052</td>
</tr>
<tr>
<td>4-8</td>
<td>0.601</td>
<td>0.362</td>
<td>2.38</td>
<td>0.039*</td>
</tr>
<tr>
<td>8-15</td>
<td>0.598</td>
<td>0.357</td>
<td>2.357</td>
<td>0.040*</td>
</tr>
<tr>
<td>15-25</td>
<td>0.498</td>
<td>0.248</td>
<td>1.818</td>
<td>0.099</td>
</tr>
<tr>
<td>25-50</td>
<td>0.456</td>
<td>0.208</td>
<td>1.619</td>
<td>0.137</td>
</tr>
<tr>
<td>&gt;50</td>
<td>0.498</td>
<td>0.248</td>
<td>1.817</td>
<td>0.099</td>
</tr>
</tbody>
</table>
Figure S5.1 Melting curves of qPCR products, showing melting peaks of both tissue-derived genomic DNA, no-template control (NTC), positive and negative detections of target species from water samples.
**Figure S5.2** Sequence alignment of four samples (sequences 2-5) from the main canal mitochondrial COI PCR amplicons with known *L.fortunei* (GeneBank Accession #: HQ843796.1).
CHAPTER 6: FUNCTIONAL RESPONSE AND SIZE-SELECTIVE CLEARANCE OF SUSPENDED MATTER BY AN INVASIVE MUSSEL, AND IMPLICATIONS IN A LARGE WATER DIVERSION PROJECT

6.1 Introduction

Aquatic ecosystems worldwide have endured a series of high-profile species invasions (Bobeldyk et al. 2015; Carlton et al. 2017). Suspension feeders, including many bivalves, represent an important group of aquatic invaders which can profoundly influence invaded environments (Gili & Coma 1998; Boltovskoy & Correa 2015). Through a variety of interactions with local abiotic conditions and biological communities, suspension feeding invaders serve as potential ecosystem engineers (e.g., MacIsaac 1996; Gili & Coma 1998; Boltovskoy & Correa 2015; Linares et al. 2017). For example, filter-feeding activities deposit suspended matter from the water column to the benthos (Nishizaki & Ackerman 2017), leading to food depletion of pelagic habitats but supplementation of benthic ones, alteration of geochemical cycling, enhancement of water clarity and of aquatic macrophytes, and shifts of resource utilization in food webs (Hecky et al. 2004; Boltovskoy & Correa 2015; Karatayev et al. 2015; Fera et al. 2017; Knight et al. 2018). Furthermore, many invasive, suspension-feeding molluscs modify physical habitats by forming massively dense aggregates (Gili & Coma 1998; Boltovskoy & Correa 2015). Generally, these impacts will be highly distinguished in systems where no native counterparts of the invaders exist (Ricciardi & Atkinson 2004).

---

The absolute impact potential of any invasive species to invaded ecosystems is determined its impact potential and abundance. However, the impact potential of invasive species relative to native analogous provides more operational values predicting the impact of the former (Dick et al. 2017b). In the case of suspension-feeding invaders, knowledge about their clearance rate, which is defined as the amount of water cleared of food per individual per hour (Darrigran & Damborenea 2010; Karatayev et al. 2015) is critical to inform their functional responses and to allow for determination of both absolute and relative impact potential (when native analogous present). The response of clearance rate to varying food levels, particularly at low concentration, reflects resource uptake and has enormous implications of their interactions with the environment (Lehman 1976; Gili & Coma 1998). Given that food availability may differ in space and time, it could contribute to invasion success by influencing nutritional status of animals and, if sufficient, allow for a wide distribution of invaders and greater impact potential (e.g., Brown et al. 2013; Laverty et al. 2015; Bracken et al. 2017; Dick et al. 2017a). Furthermore, aquatic systems with sufficient and/or quality food resources may sustain larger populations than those with limited resources (see review by Vaughn & Hakenkamp 2001). Apart from substantial deposition of suspended matter, many suspension feeders (e.g., bivalves) demonstrate size-selective clearance of suspended particulates, which, in turn, may alter the size distribution of suspended matter in the water column (Ward & Shumway 2004) and alter composition of phytoplankton communities (Vanderploeg et al. 2001, Cataldo et al. 2012). Thus, a comprehensive understanding of the filter-feeding behaviors, including the response of clearance rate, to
varying food levels and size-selective removal of suspended matter, can inform predictions of the invasiveness and potential impacts of filter-feeding invaders.

The golden mussel *Limnoperna fortunei* is a suspension-feeding, invasive bivalve in freshwater ecosystems in Asia and South America (Ricciardi 1998; Xia et al. 2018). It serves as an ecosystem engineer and exerts profound ecological impacts on many invaded habitats (Darrigran & Damborenea 2010; Boltovskoy & Correa 2015). Furthermore, its fouling impairs many submerged systems that use raw water, including drinking water supply systems, hydropower plants, irrigation stations, and shipping industries (Nakano et al. 2015; Xu et al. 2015; Chakraborti et al. 2016; Li et al. 2019). Trans-oceanic spread of the species has been facilitated by discharge of contaminated ballast water, and possibly by fouling of vessels or by the aquaculture trade (Boltovskoy 2015; Nakano et al. 2015). Inland spreading has been facilitated by anthropogenic hydraulic programs such as dam construction (Nakano et al. 2015) and water transfer projects (Xu et al. 2015; Zhan et al. 2015). For example, the >1200km long central route of the South to North Water Diversion Project (SNWDP) in China connects areas inhabited by the golden mussel in the south (source) to uninhabited water bodies in the north (recipient), facilitating dispersal of this species (Zhan et al. 2015). Relative to natural ecosystems, such anthropogenic hydraulic facilities may support much denser populations of golden mussels owing to vast amounts of concrete which could serve as settling substrate and result in great potential impacts (Xu et al. 2015; Nakano et al. 2015). To our knowledge, the golden mussel is the only biofouling, filter-feeding bivalve in this region and the potential recipient water environments of the SNWDP. Therefore, determination of its functional response can inform impact potential.
The filter-feeding behavior of golden mussels has been measured in several studies, particularly in invasive populations in South America (see Boltovskoy 2015). Factors such as animal size, temperature, food types, and inorganic matter influence their clearance rate (Sylvester et al. 2005; Pestana et al. 2009). However, the functional response has not been explored. Also, there is a paucity of knowledge regarding the species’ impact on different size classes of suspended matter. In this study, we tested the clearance rate of golden mussels exposed to varying food concentrations and investigated their functional response. We also tested their suppressive effects on phytoplankton in mesocosms and their potential to alter the size distribution of suspended matter. Finally, we discussed their potential impacts in large anthropogenic water diversion projects using the South to North Water Diversion Project (central route) in China as an example.

6.2 Materials and Methods

Study design

We first tested the clearance rate of the golden mussels under varying levels of algae concentration using nontoxic monospecific algae Chlorella vulgaris cultured in laboratory, which was followed by modeling the functional response of the species. Next, we conducted two independent 7-day experiments in mesocosms to investigate the suppressive effects of golden mussels on growing algae (measured as chlorophyll a concentration) (experiment 1), and size-selective clearance of suspended particulates of various sizes (experiment 2). Specifically, algae food was mimicked by adding commercial green algae to experimental tanks daily in experiment 1 and suspended matter with varying sizes was prepared by collecting lake water in experiment 2. Given
that abundance of golden mussels in natural environment varies widely (e.g., 85-1.7×10^5 ind. m^{-2}, Boltovskoy 2015), we set different animal densities in the two experiments.

**Animal collections**

Golden mussels were collected from the underside of a floating dock in Danjiangkou Reservoir, China (32°45’N, 111°35’E) by carefully cutting the byssal threads. Mussels were transported in coolers to laboratory and raised in dechlorinated water at 24 °C in a 60-L fish tank. Typical adult-size individuals were separated from mussel aggregates by carefully cutting byssal threads, and were subsequently placed on glass slides to allow reattachment. Commercial fish food [~50 mL concentrated algae culture (Chlorella sp.), concentration >10^6 cells mL^{-1}] was fed to animals daily before experiments.

**Functional response test**

Nontoxic Chlorella vulgaris was served as food for golden mussels to test their clearance rate. Algae was purchased from the Freshwater Algae Culture Collection at the Institute of Hydrobiology, Chinese Academy of Sciences, and cultured following the provider’s instructions. Food suspensions with varying algae concentrations were prepared by adding algae culture to dechlorinated tap water. Specifically, a total of nine algae concentrations (i.e., 276-37077 cells mL^{-1}) were prepared (and tested) by adding 20, 30, 50, 60, 120, 240, 500, 1000, and 2000 μL of cultured algae to 1.8 L dechlorinated tap water, respectively (Table S1). The food concentration range was reflective of what was observed in the potential receiving water bodies of the SNWDP (e.g., Miyun
Reservoir and Tuancheng Lake, Beijing). Slides with a single average-size (~15 mm) mussel that firmly attached were exposed in food suspension for 24 h prior to a 6 h gut clearance interval in filtered (0.22 μm pore size) dechlorinated tap water. Animals were randomly assigned to different food concentrations. Only typical adult-size animals were used though each individual size was not measured. Tests were carried out with five experimental replicates for each food concentration and three no-animal controls in beakers with 200 mL food suspension. Beakers were placed on a shaker at 100 rpm (revolutions per minute) which mixed water but did not affect animals’ filtration activities during the experiment. Animals were active (i.e., ventilating) and most started to filter water within one minute of being placed into a beaker. We started timing once a minimum of three animals began filtering and terminated after 10 min according to pilot tests. We collected 1 mL of food suspension for initial algae concentration from each beaker immediately after animals were introduced, and this was repeated at the end of experiment for the final concentration. The concentration of algae was measured immediately after collection of each batch (i.e., initial and final) of samples using a BD Accuri C6 flow cytometry instrument (Becton, Dickinson and Company, US). 200 μL of each mixed sample was loaded to quantify algae, and a 5s sample shake for every two samples was programmed to reduce algae settlement during measurement. Cell diameter of *C. vulgaris* was 2.54 ± 0.56 μm (Mean ± S.D., n=118), which was measured under a microscope (10×40, Olympus CX41, Japan).

Ingestion rate (i.e., the number of algae cells consumed per mussel per hour) was calculated as:

\[
\text{Ingestion rate} = V \cdot \frac{[(c_i-c_f)-(c_i'-c_f')]}{N \cdot T},
\]
which was adopted from Tokumon et al. (2015), where \( V \) is the algae suspension volume (mL), \( N \) is the number of animals in each experimental beaker, \( T \) is the experimental duration (h), \( C_i \) and \( C_f \) are initial and final algae concentration (cells mL\(^{-1}\)) in each experimental beaker, while \( C_i' \) and \( C_f' \) are the average initial and final algae concentration of three control beakers, respectively. Clearance rate in each experimental beaker was calculated following Coughlan (1969) and Tokumon et al. (2015):

\[
\text{Clearance rate} = V \times \left[ \frac{\ln(C_i/C_i') - \ln(C_f/C_f')}{N \times T} \right].
\]

Functional response of the golden mussel refers to the ingestion rate of the animal as a function of food density, which was tested and modeled using the R package “frair” in this study. The type I functional response was chosen to fit the functional response curve following a preliminary comparison with the Rogers type II functional response (Pritchard et al. 2017). Non-parametric bootstrap (n=2000) was conducted to generate 95% confidence intervals for the fitted functional response curve, and statistical analysis was performed in R 3.5.1 (R core team, 2018). Two of the five replicates in each of three lowest algae concentration treatments (i.e., 276, 491, and 908 cells mL\(^{-1}\)) and one in each of two high algae concentration treatments (i.e., 3780 and 19035 cells mL\(^{-1}\)) were negative and were discarded. The former problem was likely due to sampling and (or) enumeration bias, while cause of the latter was unclear as we tried to reduce algae settlement during the experiment.

**Suppressive effect on growing algae test**

Four golden mussel aggregates with various individuals (n=1, 4, 8, 16) and similar size [medium: ~12-18 mm shell length, no significant difference among aggregates (\( F_{3, 25} \)…)]
which was measured after experiment – were used to graze on the concentrated commercial fish food (Chlorella sp. cells). Each mussel clump was placed at the center of a round plastic bucket containing 10 L dechlorinated tap water (~30 cm deep). Each bucket was slightly aerated with an air stone placed at the bottom by the wall ~10 cm away from the mussel clump to reduce the impact of turbulence on animals and potential breakdown of feces/pseudofeces produced. Three control tanks without animals were arranged with the same setup. Given that the food used in this experiment was a single algal species, we used chlorophyll a as a proxy of algal density (Gitelson et al. 1999). Before animals were introduced, chlorophyll a concentration in each bucket was adjusted to the same level (2-3 µg L\(^{-1}\), day 0 in Fig. 3 A). Chlorophyll a concentration was measured in situ (~10 cm deep) using a handheld fluorometer (Turner Designs, U.S.). Water temperature and pH (multiparameter water quality sonde, MYRON, USA), and dissolved oxygen (DO meter, HACH, USA) were also monitored throughout the experiment. All measurements were carried carefully without re-suspending materials from the bottom. 10 mL of the concentrated Chlorella sp. was added to each tank following daily measurements (i.e., once per day) throughout the experiment. At the end of the experiment, a 500 mL water sample from the subsurface of each tank was collected to measure particle size of suspended matter using a particle counter (PAMAS Water Viewer, Germany), and the average of three reads for each sample was used in data analysis. The percentage change of suspension particulates for each size category was calculated as:

\[
\text{Change percentage} = \left( \frac{C_t - C_c}{C_c} \right) \times 100\%,
\]
where $C_t$ is the average concentration for each experimental mesocosm, and $C_c$ is the average of all controls without animals.

**Size-selective clearance of suspended particulates test**

Lake-water was collected from the surface layer of Kunming Lake (Beijing, China) and transported to laboratory within one hour. Prior to use, large inorganic particles were removed by a 2-hour settlement in laboratory, and the initial chlorophyll a concentration and water pH was ~66 $\mu$g L$^{-1}$ and 8.05, respectively. Six tanks filled with 10 L lake-water each were aerated slightly by a single air stone placed adjacent to the wall. Four mussel clumps with wet weights of 1.96 g ($T_1$, $n = 4$), 10.1 g ($T_2$, $n = 17$), 16.2 g ($T_3$, $n = 31$), and 33.1 g ($T_4$, $n = 100$) – which were weighted prior to test – were randomly assigned into 4 experimental tanks and the remaining (2 tanks) served as controls. Here the average wet weight of mussel in $T_4$ (0.331g) was lighter than in the three groups with less dense animals, while it was reflective of adult mussel size (e.g., >15 mm shell length) according to our bivalve biometric analysis using mussels collected from the same site and around the same time (Z. Xia, unpublished data). Chlorophyll a concentration and size of suspended particulates in each tank were measured *in situ* daily between days 0 and 7 using methods described above. No additional food was provided to animals, and no mortality was observed until day 7 (i.e., two mussels died in $T_1$ and one in $T_3$, respectively), which was not included in analysis to avoid bias. The size of suspended particles was partitioned into eight default categories of the particle counter (i.e., ≤1, 1-2, 2-4, 4-8, 8-15, 15-25, 25-50, and >50 $\mu$m), and the change percentage of each size category relative to no-animal controls was calculated as described above.
Capture efficiency of each size category of suspended particles was characterized as the removal rate (%) on the first day since introduction of golden mussels. To explore factors affecting removal rate on the first day, we applied a generalized linear model to model the change percentage of suspended particles as a function of animal abundance, particle size, and their interaction (i.e., abundance × particle size). To simplify modeling, we reduced the number of variables for particle size by combining the default categories into three new ones (i.e., ≤2, 2-25, and >25 μm) within each the original categories exhibited similar effects on capture efficiency.

6.3 Results

*Clearance rate and functional response of golden mussel*

Clearance rate of golden mussels tested with nontoxic strain of *Chlorella vulgaris* differed among algae concentrations, ranging from a high of 305.5 ± 105.9 mL ind.\(^{-1}\) h\(^{-1}\) to a low of 72.6 ± 27.0 (Mean ± S.E.) (Fig. 6.1). Clearance rate was highest at low food concentrations (e.g., 491-908 cells mL\(^{-1}\)) and declined as algae concentration increased (Fig. 6.1), though greater variance was observed at low food levels. Golden mussels demonstrated the lowest clearance rate when food concentration reached the maximum experimental value (37077 cell mL\(^{-1}\)) (Fig. 6.1). As expected, golden mussels exhibited a linearly-increasing ingestion rate with increasing algae concentration, resulting in a type I functional response \((F_1, 35 = 41.3, R^2 = 0.541, P < 0.0001, \text{Fig. 6.2})\). The estimated attack rate was \(a = 0.502\) \((P < 0.001)\) with a negligible handling time \((h)\) (Pritchard et al. 2017), indicating a saturation point higher than the maximum algae density in our experimental setup (i.e., ~37000 cells mL\(^{-1}\)).
**Chlorophyll a clearance**

Presence of golden mussels demonstrated strong suppressive effects on accumulation of chlorophyll a, and animal abundance was positively related to the extent of chlorophyll a change (Fig. 6.3 A). Specifically, clumps of animals (i.e., mussel abundance = 4, 8, and 16) constrained the increase of chlorophyll a concentration, resulting in a significantly lower final concentration relative to controls lacking mussels (1.07 ± 0.43 vs. 18.22 ± 0.71 μg L⁻¹, Mean ± S.E.) ($t_4 = 3.28, P < 0.001$, Student t-test). In contrast, the one-mussel treatment had very limited effects ($t_2 = 1.44, P = 0.286$, One-sample t-test) and chlorophyll a concentration consistently increased (Fig. 6.3 A). The highest chlorophyll a concentration achieved was inversely related to animal abundance, with 17.2, 6.7, 5.5, and 3.4 μg L⁻¹ in mesocosms with 1, 4, 8, and 16 individuals, respectively (Fig. 6.3 A). Temperature (23.9 ± 0.3°C), pH (8.07 ± 0.12), and DO (104 ± 3.2%) (Mean ± S.E.) were relatively stable throughout the experiment, suitable for golden mussels (e.g., Boltovskoy 2015). Similar effects were found for golden mussels exposed in lake water where no additional food was added (i.e., experiment 2), as chlorophyll a concentration dropped dramatically when abundant animals were present. Conversely, the rate of this process was reduced in aquaria with fewer golden mussels (Fig. 6.3 B).

**Size-selective clearance of suspended particulates**

Golden mussels substantially removed suspended matter from the water column and altered its size distribution relative to controls lacking animals (Figs. 6.4 & 6.5). At the end of experiment 1 (commercial algae added) (d7), the concentration of suspended particles in tanks with mussels declined substantially relative to controls across most size
categories (Fig. 6.4). Animal abundance ($F_{3, 64} = 562.3$, $P < 0.001$), particle size ($F_{7, 64} = 271.7$, $P < 0.001$), and their interaction (abundance $\times$ particle size) ($F_{21, 64} = 341.6$, $P < 0.001$) (Two-way ANOVA) were significantly related to the change percentage.

Specifically, concentration of suspended particles declined significantly with animal abundance ($P < 0.001$) except between 4- and 8- individual treatments ($P = 0.997$). Suppressive effects on suspended particles were observed across all size categories, with the exception of the $>25$ μm particles in the 16-individual treatment, which yielded remarkably higher concentrations than controls (Fig. 6.4).

In experiment 2 where no additional food was added to tanks, the concentration of suspended particles experienced a sharp initial decline before levelling off relative to controls (Fig. 6.5). The majority of suspended particulates removed occurred within the first three days, while this process was prolonged for the treatment with the lowest animal abundance (4 mussels) (Fig. 6.5). In most cases, the concentration dropped sharply on day 1 (Fig. 6.5), during which the animal abundance, particle size, and their interaction demonstrated significant relationships with the change percentage, explained 72.6% of deviance (Table 6.1). Specifically, animal abundance =17 and greater, and particle size $>25$ μm exhibited a significant negative relationship with change percentage of suspended particles. Interactions between abundance and particle size, however, exhibited either negative (i.e., abundance =17 or =31 $\times$ size $2-25$ μm) or positive (i.e., abundance =31 or =100 $\times$ size $>25$ μm) relationships with change percentage (Table 6.1). As a result, capture efficiency of suspended particles was positively related to particle size (Fig. 6.6). Change percentages of suspended particles $>25$ μm were significantly
related to the interaction term in treatments of mussel abundance =31 and =100 (Table 6.1), which were not considered the capture efficiency to avoid bias (Fig. 6.6).

6.4 Discussion

The functional response of species plays a central role in understanding their interactions with their resources. For invasive species, functional response can be used to evaluate potential impact relative to either native or other invasive counterparts (e.g., Dick et al. 2017b; Hoxha et al. 2018) and has been proposed as a universal trait to predict impacts (Dick et al. 2017a). Golden mussels exhibited a type I functional response (Fig. 6.2), consistent with existing studies on many suspension feeders (Jeschke et al. 2004, but see Sarnelle et al. 2015). The measured functional response indicates a strong clearance potential on suspended matter (e.g., phytoplankton and organic debris) as the mussels demonstrated a linear ingestion rate across the wide range of food concentrations used (Pritchard et al. 2017). Our results demonstrated strong influences of food concentration on clearance rate – highest at low food concentrations and reduced at higher food concentrations (Fig. 6.1) – consistent with numerous existing studies (e.g., Riisgård et al. 2011; Tokumon et al. 2015). This is mainly owing to low energy needed during filtering activities in bivalves, allowing individuals to filter at maximum rates (Willows 1992; Gili & Coma 1998). Such characteristics may maximize the energy gain of golden mussels. It should be noted that reduced clearance rate is possible when golden mussels encounter long-term starvation (e.g., no food supplied). For example, golden mussels can tolerate up to 125 days of starvation, during which the animals reduced filtering activities by
closing valves (Cordeiro et al. 2016). A similar phenomenon was observed in blue mussels, which was proposed as an energy-saving mechanism (Riisgård & Larsen 2015).

The concentration of suspended particulates across all size categories initially declined following introduction of golden mussels (Fig. 6.5), indicating that the mussel is capable of capturing a broad array of particles. The extent to which the suspended particulates were removed was highly dependent on animal abundance (Figs. 6.5 & 6.6). During the first day of exposure, we observed different removal rates for different size categories, indicating differing capture efficiencies (Fig. 6.6). Generally, fine particles (e.g., ≤2 μm) were captured at lower efficiency (Fig. 6.6, abundance = 4 & 17). However, a high overall clearance rate can be achieved if animals are present at very high abundance (e.g., abundance = 100, Fig. 6.6, Table 6.1). In experiment 2, the use of lighter wet weighted mussels in T4 (i.e., abundance = 100) might not underestimate the overall clearance effect (see Pestana et al. 2009). In both experiments, the concentration of fine particles declined despite initially being more abundant (e.g., Fig. S6.1), while large particles (e.g., >25 μm) had higher final concentrations than in controls (Figs. 6.4 & 6.5). These findings are consistent with the view that golden mussels package fine suspended matter to coarse particles. The increased coarse particles in the tanks with high mussel abundance likely resulted from the ejection of feces through the exhalant siphon and/or pseudofeces through the inhalant siphon, which has been suggested to improve local mixing of the benthic layer near mussel aggregates (MacIsaac & Rocha 1995; Nishizaki & Ackerman 2017). In contrast, such patterns were not observed in tanks with low mussel abundance, which was likely because of less coarse particles produced and limited overall mixing effect around mussel clumps. Golden mussels exhibited effective
clearance of particles between 2-25 μm (Fig. 6.6), which are the dominant components of suspended organic matter in many natural environments that cause light attenuation (Davies-Colley & Smith 2001). Despite being less effective, fine particles (i.e., ≤2 μm) were also suppressed and retained at a relatively low level (Figs. 6.4 & 6.5). Therefore, massive clearance of these particles from the water column could improve light penetration and change the underwater light climate (Boltovskoy & Correa 2015). Apart from the improvement of light penetration, size-selective capture of particles by golden mussels may cause unbalanced consumption of phytoplankton in natural environments, altering size composition of algae (e.g., Cataldo et al. 2012; Frau et al. 2016). Removal of most edible algae could adversely affect zooplankton and fishes dependent on it.

Similar to suspended particles, we observed strong abundance-dependent clearance of chlorophyll a by golden mussels. Specifically, mussels at low abundance exhibited a limited suppressive effect on growing algae (Fig. 6.3). However, suppression (Fig. 6.3 A) and removal rate (Fig. 6.3 B) were increasingly pronounced as mussel abundance increased. In experiment 2, chlorophyll a concentration of no-animal control declined, which was likely due to settling of suspended particles and difference of environmental condition between field and experimental tanks. However, the attenuation rate was much faster in tanks with mussels (Fig. 6.3 B). This highlights the important role of mussel abundance in their overall impact (MacIsaac et al. 1992; Gili & Coma 1998; Linares et al. 2017). The massive clearance of chlorophyll a, which is a surrogate for organic matter, may lead to its large-scale depletion in pelagic environments (Officer et al. 1982; Boltovskoy & Correa 2015). Consequently, this may restrain the ingestion rate of total suspended matter due to food concentration declines (Fig. 6.2) and reduce the growth rate
of golden mussels due to food resource deficits (Riisgård & Larsen 2015; Sarnelle et al. 2015). This process may transform stoichiometry in the surrounding water and enhance the mismatch between golden mussel requirements and available resources (Bracken 2017). It should be noted that the depletion of suspended matter in pelagic systems can be diminished if hydrodynamic conditions do not permit (e.g., strong re-suspension of sediment in lotic environment) or if nutrient cycling is enhanced, stimulating phytoplankton growth (Cataldo et al. 2012; Rowe et al. 2017).

Large water diversion projects such as the central route of SNWDP in China addresses the serious problem of regional water scarcity, though it may facilitate species dispersal into new regions (Xu et al. 2015; Zhan et al. 2015). In the case of golden mussels in the central route of SNWDP, several characteristics can amplify impacts associated with this suspension-feeding mussel. First, extensive concrete channels provide an ideal substrate for attachment, allowing for widespread establishment of the species (Nakano et al. 2015; Xu et al. 2015; Zhan et al. 2015; Li et al. 2018). Second, upstream production and export of trillions of propagules allow seeding of both newly and previously colonized areas to the north (Zhan et al. 2015). Third, water flow continually renews food resources and eliminates some waste products, allowing for dense local mussel populations (Arkema 2009). Collectively, these factors will result in a pronounced short-term gradient in both spatial distribution and population density from south to north. We anticipate that filtering impacts will correspond directly with mussel distribution and abundance, being highest in the south and low to non-existent in the northern parts of the project. In time, the species’ distribution might infill in the north, in which case impact would rise in that region as well.
The central route of SNWDP provides drinking water resource, which requires special concern (Tang et al. 2016). First, accumulation of deposited suspended matter and increased light penetration may stimulate growth of macrophytes in certain areas such as those with reduced flow velocity in the open canal. This may further reduce water velocity and impact water transfer efficiency (Madsen et al. 2001), although the extent of the effect requires further study. Next, an increase of inorganic nutrients such as ammonia and soluble reactive phosphorus within the canal and selective grazing of golden mussels on phytoplankton may stimulate growth of harmful algae (e.g., Vanderploeg et al. 2001; Cataldo et al. 2012). Third, proliferation of golden mussels per se may largely threaten water quality when mass mortality occurs, for example, following extreme events such as an increased inorganic matter load, reduced water level (Oliveira et al. 2011), and high water temperature (e.g., McDowell et al. 2016). Fourth, impacts of golden mussels on the central route of SNWDP associated with their filter-feeding behavior are complex and sometimes conflicting (e.g., the benefit of reduced suspended matter in raw water may simplify drinking water processing though it may be offset by facilitation of cyanobacteria). Finally, proliferation of this nuisance species in the water diversion canal may exert impacts that are hard to identify in advance. Thus containment and dispersal limitation should be prioritized wherever possible.
6.5 References


Grande Reservoir (Argentina): evidence from mesocom experiments. 


Table 6.1 Summary of the generalized linear model of change percentage (%) of suspended particles on day 1 of experiment 2 as a function of animal abundance (Abundance, # of individual), size of suspended particles (Size, μm), and their interactions (Abundance × Size), showing results of all significant variables. Significance level: *** 0.001, ** 0.01, · 0.1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Est. coefficient (S.E.)</th>
<th>t-value</th>
<th>P-value</th>
<th>Deviance explained</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-21.49 (9.65)</td>
<td>-2.23</td>
<td>0.029*</td>
<td></td>
</tr>
<tr>
<td>Abundance = 17</td>
<td>-30.26 (13.65)</td>
<td>-2.22</td>
<td>0.029*</td>
<td></td>
</tr>
<tr>
<td>Abundance = 31</td>
<td>-35.79 (13.65)</td>
<td>-2.62</td>
<td>0.01*</td>
<td></td>
</tr>
<tr>
<td>Abundance = 100</td>
<td>-70.08 (13.65)</td>
<td>-5.14</td>
<td>1.8e-6***</td>
<td></td>
</tr>
<tr>
<td>Size &gt;25</td>
<td>-34.49 (13.65)</td>
<td>-2.53</td>
<td>0.103*</td>
<td>72.6%</td>
</tr>
<tr>
<td>Abundance = 17 × Size 2-25</td>
<td>-35.24 (16.72)</td>
<td>-2.11</td>
<td>0.038*</td>
<td></td>
</tr>
<tr>
<td>Abundance = 31 × Size 2-25</td>
<td>-27.84 (16.72)</td>
<td>-1.67</td>
<td>0.099*</td>
<td></td>
</tr>
<tr>
<td>Abundance = 31 × Size &gt;25</td>
<td>123.13 (19.30)</td>
<td>6.38</td>
<td>9.2e-9***</td>
<td></td>
</tr>
<tr>
<td>Abundance = 100 × Size &gt;25</td>
<td>58.55 (19.30)</td>
<td>3.03</td>
<td>0.003**</td>
<td></td>
</tr>
</tbody>
</table>
Figure 6.1 Clearance rate (Mean ± S.E.) of golden mussels in aquaria (24 °C) at varying algae concentrations. Different letter indicates significance at 0.1 level (Fisher’s LSD test).
Figure 6.2 Fitted functional response (with 95% confidence interval, gray area) of golden mussels. Open circles indicate the actual ingestion rate.
Figure 6.3 Chlorophyll a concentration in mesocosms in which golden mussel aggregates with different number of individuals were (A) daily fed or (B) not fed by adding additional algae throughout the experiment. Test water was aerated tap water in (A) and raw lake water in (B). Error bars (A) indicate standard errors of the average concentration in no-animal control mesocosms.
Figure 6.4 Change percentage (Mean ± S.E.) of concentration of suspended particulates in mesocosms with golden mussels relative to no-animal controls, showing results of each size category at the end (d7) of the experiment. Error bars indicate standard error of three replicates for each sample.
Figure 6.5 Change percentage (Mean ± S.E.) of concentration of suspended particulates in mesocosms with golden mussels related to no-animal controls, showing daily results of each size category in the no-feeding experiment. Error bars indicate standard error of three replicates for each sample.
Figure 6.6 Capture efficiency (Mean ± S.E.) of suspended particles of varying size, showing results of removal rate during the first day. Error bars indicate standard error of three replicates for each sample. Note that the data of size >25 μm in treatments mussel abundance =31 and =100 was not shown because the removal rate (capture efficiency) was significantly related to the interaction between animal abundance and particle size (Table 6.1).
6.6 Supplementary Information

**Figure S6.1** Log$_{10}$ transformed concentration (Mean ± S.E.) of suspended particulate of different sizes in the initial lake water used in experiment 2.
CHAPTER 7 GENERAL DISCUSSION

In this dissertation, I have provided a comprehensive assessment of options to improve detection of taxa at low abundance using an emerging tool – environmental DNA (eDNA) – which presumably has several advantages over methods used traditionally (Taberlet et al. 2012; Bohmann et al. 2014). My assessment illustrates that detection of species from aquatic environments can be improved through a variety of means including but not limited to screening of highly sensitive assay(s) or the PCR method used, and intensifying sampling effort (e.g., sample volume and number of sample replicates) (Chapters 2-4). My eDNA-based detections of golden mussel suggest that sampling location and time are critical to justify the utility of eDNA methods and to interpret detection results properly. These findings shed light on application of eDNA tools for detecting species at low abundance, benefiting management of species invasions (Mehta et al. 2007). Though the technical exploration of my research focused on golden mussels, results generated here have implications for all programs that rely on detection of organisms at low abundance (e.g., biological conservation and species invasion management).

7.1 eDNA-based early detection of aquatic species

An important reason for the replacement of the traditional ‘catch and look’ methods by eDNA-based methods is that the latter offer much higher sensitivity (i.e., higher catch per unit effort) than the former (e.g., Jerde et al. 2011; Tréguier et al. 2014; Smart et al. 2015; Wilcox et al. 2016). In most environmental samples, DNA of the target species is often present as a small fraction of total DNA of many different species as well as other
components (e.g., trace level of organic matter) (e.g., Bohmann et al. 2012; Turner et al. 2014). This characteristic of eDNA samples requires robust methods to reliably target the DNA of interest (if it is present). However, many existing studies fail to test and report the sensitivity of the assays used, inflating uncertainties in results interpretation and the risk of false negatives (Chapter 2; Schultz & Lance 2016). I developed and optimized the detection sensitivity of golden mussel DNA from water samples via screening primer pairs and PCR methods, thereby reducing such risks (Chapters 3 & 4). Given the low concentration of target species DNA (due to rarity of the target) in the environment and prevalence of PCR inhibitors (McKee et al. 2015), any assay should be optimized to achieve high sensitivity (i.e., low LoD). However, it may be difficult or even impossible to achieve the lowest theoretical LoD (i.e., one copy per PCR reaction, Bustin et al. 2009) because of other considerations (e.g., specificity, amplicon size) to an ideal eDNA assay (Wilcox et al. 2013, 2015). False negatives seem inevitable, though optimizing the assay used will benefit detection of low levels of target DNA. Some studies tested but reported LoD using a variety of formats (Chapter 2). I propose that the reporting format of LoD should be unified to a standard format such as the number of copies of the target fragment, which allows for improved understanding of the species’ LoD and simplifies interpretation of presence/absence data. Also, such information can guide future use of the developed assays and inform study design, such as sampling intensity (Schultz & Lance 2015).

Similar to traditional methods (Harvey et al. 2009; Hoffman et al. 2011), eDNA-based detection of species present at low abundance requires extensive sampling effort (Chapters 2 & 4). Successful retrieval of target DNA from the field is indeed the goal of
any detection program regardless of method used. Given the expected low eDNA concentration of certain species and its uneven distribution (Furlan et al. 2016), extensive sampling - such as collecting multiple water samples from each location and (or) filtering large volumes of water per sample - increases the probability of collecting sufficient DNA in at least one sample (i.e., it exceeds the LoD of assays), increasing the overall detection rate (e.g., Dougherty et al. 2016; Schultz & Lance 2015). In addition, extensive sampling improves results interpretation as stronger evidence can be achieved if multiple positive samples can be observed in a sampling trip or across different trips (e.g., Jerde et al. 2011).

In addition to sampling intensity, a priori knowledge of the putative distribution of target DNA in environment is necessary (Chapter 3). DNA shed into the environment persists longer when attached to organic matter than free DNA (Tréguier et al. 2014; Turner et al. 2015), rendering the settling of suspended matter an important factor determining eDNA fate, especially in lentic systems such as ponds, lakes, and reservoirs, where settling may be strong and false negatives, may occur if only surficial water is sampled (e.g., Chapters 3 & 4). Species traits may further complicate the spatial distribution of eDNA in the environment. For instance, some species with high mobility throughout their life span (e.g., fish) that can shed DNA across a range of spaces, while those with limited mobility or a sedentary life stage may have their DNA limited to locations at or adjacent to where they colonize, or downstream in mixed water (e.g., Danjiangkou Reservoir vs. Pengxi River in Chapter 3). In the case of bivalves such as golden mussels or zebra mussels, animals enter a dominant, sessile adult stage following a much shorter free-swimming planktonic veliger stage (Ackerman et al. 1994;
Boltovskoy 2015). The transition from free-swimming stage to sessile stage may restrain golden mussel DNA to limited surrounding space though, which may hinder their detection, particularly in surface water in lentic habitats, though adult animals have greater biomass than veligers.

Understanding eDNA availability over time is important for detection interpretation. For a given system, eDNA concentration increases with time since the species was introduced until a static state is achieved; DNA rapidly declines when individuals are eradicated (Pilliod et al. 2014; Sansom & Sassoubre 2017). Emerging changes of hydrodynamics (e.g., stratification), species abundance (Pilliod et al. 2013), reproduction (e.g., Gingera et al. 2016), and other activities (e.g., de Souza et al. 2016) may lead to a concentration change of eDNA. Though impacts of these factors on eDNA concentration vary from system to system, keeping them in mind can assist in interpretation of eDNA-based detection results.

7.2 Colonization risks, potential impacts, and management implications of golden mussels in SNWDP

As predicted by Zhan et al. (2015), opening of the SNWDP provided an ‘invasion highway’ for species invasion. Given the cyclic biphasic life cycle of golden mussels, the introduction of fouled adults or planktonic propagules (i.e., veligers) should be a primary concern to prevent their introduction and colonization. The latter should be the priority in the case of SNWDP because no vessels are allowed in the waterway, reducing the risks of invasion via adults fouled on vessels. In contrast, introduction of veligers to the main canal of the SNWDP seems inevitable when they are present (e.g., in reproduction
Measures that reduce veliger abundance should be considered to reduce propagule pressure (Lockwood et al. 2005).

Artificial submerged structures may suffer serious fouling by golden mussels, providing critical opportunities for further colonization (Boltovskoy 2015; Xu et al. 2015). In the main canal of the SNWDP, additional spread is possible owing to ideal habitats (e.g., concrete walls) and water flow conditions that continuously introduce propagules downstream during reproductive events (Zhan et al. 2015). Given the tremendous distance and changing thermal conditions of the project, as well as the high mortality rates of mussel veligers during downstream transport (Horvath & Lamberti 1999; Boltovskoy 2015), it was unlikely that the entire main canal was colonized by propagules originating directly from the source reservoir after the initial opening.

Established populations may fluctuate under changing conditions such as freezing during the winter, which is common in the northern end of the canal. However, bridgehead populations formed in suitable locations can act as continuing sources that maintain downstream ‘sink’ populations by producing new propagules in warm months (e.g., Estoup et al. 2010). It seems inevitable that the main canal will be fully colonized by golden mussels without intervention. Post-invasion management (e.g., mitigation) must be considered to reduce the adverse effects of established mussels on the canal (e.g., concrete erosion; Xu et al. 2015; decreased water flows coupled with increased sedimentation rates). As demonstrated in Chapter 6, golden mussels facilitate bio-sedimentation through packaging fine suspended matter into coarse particles, while the overall removal rate depends on mussel abundance. Massive deposition of suspended
matter from the water column to water bottom may have consequences in the colonized environment (Boltovskoy & Correa 2015).

The diverted water is primarily used as drinking water, while other uses such as irrigation, ecological water resource restoration, and groundwater restoration have also been proposed (nsbd.mwr.gov.cn). Such activities direct water from the SNWDP system to a variety of local aquatic ecosystems in the receiving region and represent opportunities for further dispersal of the species. Though information (e.g., environmental suitability of receiving water body; amount of water used, etc.) regarding such ‘side uses’ of diverted water is lacking, I propose that such secondary connections to the main flow be implemented with caution to reduce the likelihood of spreading golden mussels to new locations.

7.3 Future work

In this thesis, I addressed the significance of sensitivity screening when applied to eDNA methods. For the SNWDP water diversion system, I tested the method only in the main canal because diverted water was not directed outside the system, allowing me to test potential dispersal in the recipient region. I found that extensive sampling is needed to reduce false negatives. Additional studies should explore specific optimal practices (e.g., optimal sample volume, number of sample replicates), especially by taking into account the time, labor, and financial cost of collection and processing. I propose that such studies will further optimize surveillance programs over my current method.

In addition, I found that the occurrence of positive eDNA detections overlapped with the expected reproduction season, which created the high-risk period of spread.
However, eDNA collected from the field has complex sources. This brings bias in interpretation of invasion risks because only ‘viable’ propagules pose threats (Darling & Mahon 2011). Future studies characterizing the fate of eDNA derived from established mussels or from planktonic larvae may prove useful to isolate particular sources of eDNA. When target species are present at low abundance, eDNA detection usually provides earlier evidence than visual detection (Jerde et al. 2011; Amberg et al. 2019). Factors determining the time lag of the latter method have rarely been investigated. In the case of golden mussels, studies investigating dynamics of viable propagules under different environmental (e.g., hydrodynamic, thermal) conditions are greatly needed.

Whether or not eDNA can be used as a reliable tool for abundance estimation has received much attention. Generally, concentration of eDNA is positively related to species abundance (biomass), while the prediction power of eDNA for species abundance is greater in controlled systems than in natural environments (Yates et al. 2019), notably due to heterogeneity among habitats (e.g., Tréguier et al. 2014). Paired studies using the eDNA method and traditional surveys (e.g., Horvath & Lamberti 1999) to quantify planktonic larvae will shed light on invasion risks. For any water body predicted to be invaded by golden mussels, eDNA concentrations will be critical in order to inform managers in a timely manner of the impending threat. Careful study design should be considered to avoid misinterpretation of detection results.

Given the large spatial scale of the SNWDP, and the principal use of diverted water (i.e., drinking), environment-friendly and cost-efficient methods must be considered to mitigate mussel abundance in the main canal. Empirical studies have proposed a variety of possible solutions to this issue though most were conducted at small scales or in
industrial facilities (Boltovskoy 2015). A comprehensive assessment of such strategies must be performed to evaluate their potential to be scaled up in SNWDP. It should also be kept in mind that there might not be a single ‘master key’ to control golden mussels in such a huge, complex system. Synergic approaches that integrate management of propagule pressure (e.g., reduce veliger density by regulating water diversion), undermining veligers survival (e.g., high frequency turbulence at canal head; Xu et al. 2015), toxic habitats (e.g., antifouling coating on essential submerged facilities), active management of establishment (e.g., establishment trap) (Xu et al. 2013), enhanced predatory pressure (e.g., restoration of local predatory fish species) need to be considered to prevent further spread and dominance of golden mussels in water systems of northern China.
7.4 References


probabilities: A case study using a stream-dwelling char *Salvelinus fontinalis*. 

*Biological Conservation*, 194, 209-216.


Yates, M. C., Fraser, D. J. & Derry, A. M. (2019). Meta-analysis supports further refinement of eDNA for monitoring aquatic species-specific abundance in nature. 


APPENDIX A: ADDITIONAL PUBLICATIONS

DOI 10.1007/s10530-015-0940-1

LETTER TO THE EDITOR

Water diversions facilitate spread of non-native species

Aixin Zhan - Lei Zhang - Zhiqiang Xia - Ping Ni - Wei Xiong - Yiyong Chen - G. Douglas Haffner - Hugh J. MacIsaac

Received: 25 June 2015 / Accepted: 6 July 2015 / Published online: 10 July 2015
© Springer International Publishing Switzerland 2015

Abstract Many countries/areas are experiencing or may soon experience water scarcity owing to rapid population growth, urbanization and/or climate change. Currently, almost one-fifth of the world's population (1.2 billion) live in areas of physical water scarcity. Water diversions have become a common-place solution proposed by governments for alleviation of physical water scarcity. Thus far, more than 80 major projects are completed or under construction globally, including the world's largest diversion: South-to-North Water Transfer Project (SNWTP) in China. Negative effects associated with water diversions, such as habitat loss and transfer of pollutants, have been recognized. However, it has been largely overlooked that "invasion highways" are created when water diversions link biogeographic regions. These "invasion highways" can facilitate spread of an array of non-native species. Although previous experiences have provided clear warnings regarding spread of non-native invasive species through artificial waterways, these lessons have been largely ignored by governments when resolving water scarcity problems. Here we use SNWTP, which will likely facilitate spread of invasive golden mussels, as well as many known examples of non-native invasive species spread through artificial waterways in other water systems, to call on governments to formally establish policy and seek management solutions to considering spread of non-native species when planning water diversions.

Keywords Golden mussel L. f. fortunae - Invasive species - Water scarcity - Water transfer

The world is facing a fresh water crisis in which this essential resource is often not available in sufficient quantity in areas where it is needed most. The United Nations considers an area as suffering from "absolute water scarcity" if the per capita availability is below 500 m³ year⁻¹. Water scarcity currently affects approximately one-fifth of the world's population (~1.2 billion people), with another 500 million people in areas approaching absolute water scarcity. By 2025, 1.8 billion people are expected to live in

Water diversions have been built around the world for millennia. Beginning in the nineteenth century, modern water diversions have been created to alleviate water scarcity, generation of hydropower, and rarely for discharge of sewage (Table 1). Among these reasons, the alleviation of water scarcity is the most common one (Table 1). Many countries/areas suffering from regional water scarcity have created diversions to transfer water from areas of relative abundance to those of relative scarcity (Table 1). California, for example, has created a series of diversions to move water from northern and eastern areas to arid southwest regions of the state (Israel and Lund 1995). Although water diversions have contributed significantly to development in many countries, such as growth of Denver and Los Angeles in the USA, the Green Revolution in India, and hydropower supply in Canada, they typically involve a number of adverse ecological impacts, including direct habitat loss and/or transformation and movement of contaminants between regions (Table 1). While these impacts are relatively straightforward and well realized, less attention has been paid to inter-basin movements of non-native species facilitated by water diversions (but see Gill 2000; Leven et al. 2009). Here we define non-native species as those that historically have never occurred in a particular region and differentiate them from invasive species, which we define as high impact species.

Nearly all proposed water diversions are located in developing countries, and are designed to alleviate water scarcity (e.g. Ghassemi and White 2007). Here we use China as a case study to highlight the biological invasion problem that may be associated with water diversions. We also explore unexpected consequences associated with spread of non-native, invasive species through artificial waterways in other water systems, and call on governments worldwide to formally establish policy and seek management solutions to considering the consequences with respect to biological invasions when new diversions are initially proposed.

China is among the many countries affected by regional water scarcity owing to both demographic and climate changes. China’s human population has experienced dramatic growth in the latter half of the twentieth century, which when coupled with rapid urbanization resulted in some of the world’s largest metropolises (e.g. Beijing, Shanghai and Tianjin). For example, Beijing’s official population increased from 4.2 million in 1949 to 21.2 million in 2013 (Fig. 1). Concomitant with this demographic change, China has experienced regional climate change, particularly in the northern part of the country (Ding et al. 2007; Wei and Chen 2009; Piao et al. 2010). Total precipitation in Beijing declined from an average of 419 mm between 1960 and 1990 to 371 mm between 1990 and 2009 (Climate change Knowledge Portal, available at http://sdwebx.worldbank.org/ClimatePortal/index.cfm?page=country_historical climate&ThisRegion=Asia&ThisCCode=CHN). Since the mid-1990s, Beijing and neighboring Tianjin have also experienced an increase from two to six in the number of days per year in which maximum temperature exceeds 35 °C (Wei and Chen 2009). The decline in precipitation as a result of climate change coupled with a general increase in water demand due to population growth and rapid urbanization has resulted in severe water scarcity in the Beijing area (Vörösmarty et al. 2010).

The distribution of water resources is highly skewed geographically in China, with per capita use being about four times higher in monsoon-affected areas south of the Yangtze River than those in the north (Yang and Pang 2006). Water utilization rates vary accordingly, from 1.7–19.5% of available stream flow for watersheds in the south to 31.0–91.7% in the north (Jiang 2009). Total per capita water availability in major cities in northern China is extremely low, with Beijing averaging only 170 m³ (Fig. 1), far below the United Nations’ absolute water scarcity threshold.

Water shortages in China are particularly acute in the northern Hai and Yellow River basins. Ocean discharge has declined from 24 to 1 billion m³ per year for the former (Jiang 2009), and from 48 to 13 billion m³ per year for the latter since the 1950s (Wang et al. 2006). Given the increasing scarcity of water in the north, it is not surprising that per capita water use in Beijing has declined sharply mainly due to water availability over the past 10 years (Fig. 1). In contrast to conditions in Northern China, the Yangtze River in the south averages 960 billion m³ of runoff to the sea, setting the stage for possible water diversions.

The concept of water diversion in China was developed decades ago, but in 1998 the government initiated plans to address the regional water scarcity crisis with its Water Agenda 21 (Yang and Pang 2006). Central to this plan is the goal to increase overall...
<table>
<thead>
<tr>
<th>Country</th>
<th>Diversion project</th>
<th>Description</th>
<th>Purpose</th>
<th>Canal/channel length (km)</th>
<th>Major negative effects detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S.A.</td>
<td>California State Water Project</td>
<td>Divert water from rivers in Northern California to the water-scars Southern California</td>
<td>WS and HP</td>
<td>1129</td>
<td>Change hydrology, cause loss of habitats (especially for fishes)</td>
</tr>
<tr>
<td>U.S.A.</td>
<td>Colorado River Aqueduct</td>
<td>Impound water from the Colorado River at Lake Havasu to the east side of the Santa Ana Mountains</td>
<td>WS</td>
<td>389</td>
<td>Facilitate spread of highly invasive quagga mussel</td>
</tr>
<tr>
<td>U.S.A.</td>
<td>Chicago-Area Canals</td>
<td>Drain Chicago's treated sewage into the Des Plaines River; Divert water from Lake Michigan into the Mississippi River watershed</td>
<td>Sewage discharge</td>
<td>N/A</td>
<td>Transfer pollutants, facilitate spread of highly invasive species such as silver carp, bighead carp, zebra mussel and round goby</td>
</tr>
<tr>
<td>India</td>
<td>Indira Gandhi Project</td>
<td>Transfer water from the Harilal Barrage at Sultanpur to the north-western region of Rajasthan</td>
<td>WS</td>
<td>650</td>
<td>Increase salinity, cause soil erosion</td>
</tr>
<tr>
<td>India</td>
<td>Telugu Ganga Project</td>
<td>Divert water from the Srisailam reservoir to the Poondi reservoir near Chennai</td>
<td>WS</td>
<td>406</td>
<td>N/A</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>Irysh-Karaganda Scheme</td>
<td>Divert water from the Irysh River Basin to Karaganda</td>
<td>WS</td>
<td>451</td>
<td>N/A</td>
</tr>
<tr>
<td>Australia</td>
<td>Snowy Mountains Scheme</td>
<td>Divert water from the Snowy River Basin to the Murray River and the Murrumbidgee River</td>
<td>WS and HP</td>
<td>225</td>
<td>Transfer waste water, cause dry land salinity in farmlands</td>
</tr>
<tr>
<td>Canada</td>
<td>James Bay Project</td>
<td>Divert the Caniapiscau River and the Eastmain River into the La Grande River</td>
<td>HP</td>
<td>N/A</td>
<td>Flooded 11,500 km² wilderness land, cause mercury contamination in fish, contribute to death of approx. 1,000 caribou</td>
</tr>
<tr>
<td>China</td>
<td>South-to-North Water Diversion</td>
<td>Eastern route (completed): from lower reach of the Yangtze River to Tianjin, Qingdao and Yantai</td>
<td>WS</td>
<td>Eastern: 1150 Middle: 1273 Western: 304</td>
<td>Transfer pollutants along channels, cause loss of habitats in Eastern route, destroy migration routes for fishes, especially endangered fishes</td>
</tr>
<tr>
<td>China</td>
<td>Diversion Project from the Yellow River to Qingdao</td>
<td>From the Yellow river to Qingdao</td>
<td>WS</td>
<td>290 km</td>
<td>Cause soil erosion, transfer pollutants</td>
</tr>
</tbody>
</table>

*WS* alleviation of Water Scarcity, *HP* generation of HydroPower
Fig. 1. Population growth from 1949 to 2013 (left) and water consumption per capita (right) from 2004 to 2013 in Beijing.

Fig. 2. The world’s largest water transfer project in China, i.e., the South-to-North Water Diversion Project (the middle route is shown), portends spread of mussel invaders, such as the golden mussel *Littorina foraminifera*, a bivalve species causing severe biofouling problems.

quantity in the north while reducing annual variation and seasonal fluctuations in water availability across the country, thereby reducing the frequency and scale of northern droughts and southern floods (Yung and Pang 2006). Chief among these initiatives was the South-to-North Water Transfer Project, the world’s largest water diversion. This S$62$ billion project will divert up to 45 billion m$^3$ of water per year from water-rich areas in the south to arid areas in the north (Yung and Pang 2006). The project consists of three diversions—the western, middle and eastern routes—corresponding with upper, middle and lower sections of the Yangtze River basin, the source of the diversion.

Among the three routes, the middle one is the longest (1273 km), beginning at China’s second largest reservoir—the Danjiangkou—in the south and heading north to the Miyun Reservoir, the Beijing’s principal water source (Fig. 2). Water diverted from the Han and Dan rivers will be stored in the Danjiangkou Reservoir and sent north using a series of gravity-fed channels and tunnels across the Upper Huai and Yellow rivers, before branching and terminating in Beijing and Tianjin.

The middle route has begun operation in autumn 2014 and the diverted water is expected to release to the Miyun Reservoir in summer–autumn 2015.
providing up to 20 billion m$^3$ of water per year when fully developed. This water transfer project may, however, result in a number of unexpected consequences, including spread of non-native invasive species. One example is the golden mussel Limnoperna fortunei, which is a biofouling mollusc native to the Pearl River and middle-low reaches of the Yangtze River in China, Thailand, Korea, Laos, Cambodia and Vietnam (Ricciardi 1998; Zhan et al. 2012; Paolucci et al. 2014). The species has spread to many other areas, including some localities in north China, Hong Kong, Taiwan and Japan (Xu 2012; Boltocskoy 2015). Golden mussels were also discovered in 1990 in the Rio de la Plata estuary near Buenos Aires, Argentina, and have since colonized almost 2000 km of the Paraná River and adjacent waterways in a wave of mass invasions (Boltocskoy 2015).

We confirmed presence of the golden mussel in the Danjiangkou Reservoir in June 2014 based on both morphological and genetic characteristics using the cytchrome C oxidase subunit I (COI) gene. The species forms large populations on the underside of floating docks as well as along the rocky shoreline of the reservoir. However, comprehensive field surveys and questionnaires of employees at the Miyun Reservoir have revealed that the species is not yet present in this system.

The golden mussel has a biphasic life cycle, with sedentary adults and planktonic veliger larvae. The species is easily spread by adult biofouled on vessel hulls, or veliger transfer in water currents and/or in ballast water of vessels (Ricciardi 1998; Paolucci et al. 2014; Boltocskoy 2015). The South-to-North Water Transfer Project will provide excellent dispersal opportunities for the mussel to colonize the concrete-lined channel downstream, raising the risk of serious biofouling problems throughout the middle route. Golden mussels are among the world’s worst biofouling species in industrial and municipal water works (Nakano and Strayer 2014), including in the Shisanling Reservoir north of Beijing (Xu 2012), as well as in Hong Kong (Morton 1982), Japan (Magara et al. 2001) and South America (Boltocskoy 2015). Indeed, serious biofouling problems by this species have already occurred in another water transfer project from the Dong River—a branch of the Pearl River—to Shenzhen (Xu 2012), a city with more than 13 million residents in southeastern China. Directional flow in canals and pipelines will likely limit upstream dispersal to the most vagile species, whereas a suite of species are expected to move passively or actively in currents downstream (Table 2).

Many examples and previous experiences in other systems provide clear warnings that constructed aqueducts, canals and pipelines provide opportunities for spread of invasive species (Table 2). For example, water is supplied to both Los Angeles and San Diego from the lower Colorado River below Lake Mead, a system now colonized by biofouling quagga mussels (Dreissena bugensis rostriformis) (e.g. Hickey 2010). The Chicago Area Waterway was constructed more than 100 years ago linking Lake Michigan and the Mississippi River. This conduit has permitted invasive species to move south from the Great Lakes, and may yet allow invasive silver carp (Hypophthalmichthys molitrix) and bighead carp (Hypophthalmichthys nobilis) to spread in the opposite direction (e.g. Jerde et al. 2013). Although scientists warned that invasive species could spread through the Main-Danube canal, this canal was still opened, and unfortunately this canal has served as an "invasive highway" (Leuven et al. 2009). By using this "invasion highway" and other invasion corridors, a total of 45 non-indigenous macro-invertebrate species have been recorded in the Rhine basin, and the average number of invasions per decade shows a sharp increase from <1 to 13 species (Leuven et al. 2009). Perhaps, the most notorious example of mass biological invasion unfolded followed opening of the Suez Canal, which allowed myriad species from the Red Sea to colonize the eastern Mediterranean Sea (Galil 2000; Gollasch 2006). Migration of species through the Suez Canal accounts for more species introductions in the eastern Mediterranean Sea (24.5 %) than ballast water (22.3 %) and hull fouling (16.5 %; Gollasch 2006). Plans to build a replicate canal adjacent to the Suez Canal will exacerbate an existing problem (Galil et al. 2014), as this vector, while largely unidirectional, is nevertheless large-scale in nature, perpetual, and relatively non-selective with respect to the types of species that may be spread using it. Similarly, a large number of highly invasive species have successfully utilized canals to spread throughout Europe (see Leuven et al. 2009). For example, the Main-Danube canal has allowed a series of high impact invaders to spread from the Danube River basin to the lower Rhine River basin. Mass transfer of many species holds particular risk, as sampling theory suggests that introductions of multiple species pose a greater
<table>
<thead>
<tr>
<th>Canal/channels (country)</th>
<th>Purpose</th>
<th>Invasive species</th>
<th>Invasion route</th>
<th>Major negative effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dongjiang-Shenzhen Water Diversion Project (China)</td>
<td>Water diversion for alleviation of water scarcity</td>
<td>Golden mussel, <em>Lampropila formosa</em></td>
<td>From Dongjiang River to Shenzhen Reservoir</td>
<td>Clog water pipes, decrease water quality</td>
</tr>
<tr>
<td>South-to-North Water Diversion, Eastern Route (China)</td>
<td>Water diversion for alleviation of water scarcity</td>
<td>Water hyacinth, <em>Eichhornia crassipes</em></td>
<td>From Yangtze River to Northern China</td>
<td>Clog channels, cause myriad changes, decrease biodiversity</td>
</tr>
<tr>
<td>Welland Canal (Canada)</td>
<td>Shipping</td>
<td>Sea lamprey, <em>Petroglottis marina</em></td>
<td>From Lower to Upper Great Lakes</td>
<td>Cause dramatic decline of native fishes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Alewife, <em>Alosa pseudoharengus</em></td>
<td></td>
<td>Remove large zooplankton, compete with native prey fishes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Silver carp, <em>Hypophthalmichthys molitrix</em></td>
<td>From Mississippi River to Lake Michigan</td>
<td>Remove plankton, outcompete native fishes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bighead carp, <em>Hypophthalmichthys nobilis</em></td>
<td></td>
<td>Remove plankton, outcompete native fishes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Zebra mussel, <em>Dreissena polymorpha</em></td>
<td>From Lake Michigan to Mississippi River</td>
<td>Outcompete native species, broad physical, chemical and ecological effects</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Round goby, <em>Neogobius melanostomus</em></td>
<td></td>
<td>Outcompete native species, feed on eggs of native fishes; conduit of deadly avian botulism poisoning</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Block water pipes</td>
<td></td>
</tr>
<tr>
<td>Colorado River Aqueduct (USA)</td>
<td>Water diversion for alleviation of water scarcity</td>
<td>Quagga mussel, <em>Dreissena bugensis rostriformis</em></td>
<td>Along water channels</td>
<td>Clog water pipes and cooling system, outcompete native species, alter food webs and community structure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nomad jellyfish, <em>Rhizoloma nomadica</em></td>
<td></td>
<td>Feed on many species, venemous species even cause death of human</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Textile Cone, <em>Ctenopus textilis</em></td>
<td></td>
<td>Clog waterways and pipes, outcompete native species</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Block water pipes</td>
<td></td>
</tr>
<tr>
<td>Suez Canal (Egypt)</td>
<td>Shipping</td>
<td>Zebra mussel, <em>Dreissena polymorpha</em></td>
<td>From Rhine River basin to Danube River basin</td>
<td>Clog waterways and pipes, outcompete native species</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Asian clam, <em>Corbicula fluminea</em></td>
<td></td>
<td>Clog waterways and pipes, outcompete with native species, cause poor water quality after mass mortality due to release of nutrients</td>
</tr>
</tbody>
</table>
Table 2 continued

<table>
<thead>
<tr>
<th>Canal/channel (country)</th>
<th>Purpose</th>
<th>Invading species</th>
<th>Invasive route</th>
<th>Major negative effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panama Canal (Panama)</td>
<td>Shipping</td>
<td>Polyphaga pantula</td>
<td>From the Atlantic Ocean to the Pacific Ocean</td>
<td>Feed on plankton and larval of native species, outcompete native species</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diadromus annularis</td>
<td>From Danube River basin to Rhine River basin</td>
<td>Have a major impact on biodiversity and community structure through competition and predation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Corophium rotundum</td>
<td>Killer shrimp</td>
<td>Unknown (recently detected)</td>
</tr>
</tbody>
</table>

With climate change and/or growing human populations, water scarcity has become a growing global crisis. Water diversions are becoming an attractive commonplace solution proposed by governments. Thus far, more than 80 major water transfer projects have been completed or are under construction globally (Wang et al. 2009). However, there is no policy to consider the possible spread of non-native species and their associated negative consequences. Given the fact that these artificial waterways have provided “invasion highways” for spread of non-native invasive species, with attendant negative effects on aquatic ecosystems (Table 2), water works and even human health, we call on governments worldwide to consider the potential consequences to species distributions when planning water diversions or construction of connecting waterways between watersheds. A subset of highly non-native invasive species can and will exploit these artificial waterways to spread to new areas, causing extensive economic and ecological harm.

Acknowledgments This work was supported by the National Natural Science Foundation of China (31272665) and 100 Talents Program of the Chinese Academy of Sciences to A.Z., and by Canada Research Chairs and Discovery grants from Natural Sciences and Engineering Research Council of Canada (NSERC) to HJM and GDM.

References


Ghannami F, White I (2007) Inter-basin water transfer: case studies from Australia, China, and India. Cambridge University Press, New York, NY, United States, Canada, p 462
APPENDIX B: PUBLICATION PERMISSIONS

Reprint permission for Chapter 3 - Text

SPRINGER NATURE LICENSE
TERMS AND CONDITIONS

Jun 03, 2019

This Agreement between Zhiqiang Xia ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

<table>
<thead>
<tr>
<th>License Number</th>
<th>4601541138687</th>
</tr>
</thead>
<tbody>
<tr>
<td>License date</td>
<td>Jun 03, 2019</td>
</tr>
<tr>
<td>Licensed Content Publisher</td>
<td>Springer Nature</td>
</tr>
<tr>
<td>Licensed Content Publication</td>
<td>Biological Invasions</td>
</tr>
<tr>
<td>Licensed Content Title</td>
<td>Early detection of a highly invasive bivalve based on environmental DNA (eDNA)</td>
</tr>
<tr>
<td>Licensed Content Author</td>
<td>Zhiqiang Xia, Aibin Zhan, Yangchun Gao et al</td>
</tr>
<tr>
<td>Licensed Content Date</td>
<td>Jan 1, 2017</td>
</tr>
<tr>
<td>Licensed Content Volume</td>
<td>20</td>
</tr>
<tr>
<td>Licensed Content Issue</td>
<td>2</td>
</tr>
<tr>
<td>Type of Use</td>
<td>Thesis/Dissertation</td>
</tr>
<tr>
<td>Requestor type</td>
<td>academic/university or research institute</td>
</tr>
<tr>
<td>Format</td>
<td>print and electronic</td>
</tr>
<tr>
<td>Portion</td>
<td>full article/chapter</td>
</tr>
<tr>
<td>Will you be translating?</td>
<td>no</td>
</tr>
<tr>
<td>Circulation/distribution</td>
<td>&lt;501</td>
</tr>
<tr>
<td>Author of this Springer Nature content</td>
<td>yes</td>
</tr>
<tr>
<td>Title</td>
<td>Quantifying early risks of species invasions: factors regulating south to north bivalve colonization of novel habitats</td>
</tr>
<tr>
<td>Institution name</td>
<td>Great Lakes Institute for Environmental Research, University of Windsor</td>
</tr>
<tr>
<td>Expected presentation date</td>
<td>Jun 2019</td>
</tr>
<tr>
<td>Requestor Location</td>
<td>Zhiqiang Xia</td>
</tr>
<tr>
<td></td>
<td>401 Sunset Ave.</td>
</tr>
<tr>
<td></td>
<td>Windsor, ON N9B 3P4</td>
</tr>
<tr>
<td></td>
<td>Canada</td>
</tr>
<tr>
<td></td>
<td>Attn: Zhiqiang Xia</td>
</tr>
<tr>
<td>Total</td>
<td>0.00 CAD</td>
</tr>
</tbody>
</table>
This Agreement between Zhiqiang Xia ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number: 4601540332578
License date: Jun 03, 2019
Licensed Content Publisher: Springer Nature
Licensed Content Publication: Biological Invasions
Licensed Content Title: Early detection of a highly invasive bivalve based on environmental DNA (eDNA)
Licensed Content Author: Zhiqiang Xia, Albin Zhan, Yangchun Gao et al
Licensed Content Date: Jan 1, 2017
Licensed Content Volume: 20
Licensed Content Issue: 2
Type of Use: Thesis/Dissertation
Requestor type: academic/university or research institute
Format: print and electronic
Portion: figures/tables/illustrations
Number of figures/tables/illustrations: 5
Will you be translating? no
Circulation/distribution: <501
Author of this Springer Nature content: yes
Title: Quantifying early risks of species invasions: factors regulating south to north bivalve colonization of novel habitats
Institution name: Great Lakes Institute for Environmental Research, University of Windsor
Expected presentation date: Jun 2019
Portions: Table 1; Table S1; Figures 1-3
Requestor Location: Zhiqiang Xia
401 Sunset Ave.
Windsor, ON N9B 3P4
Canada
Attn: Zhiqiang Xia

Total: 0.00 CAD
NAME: Zhiqiang Xia

PLACE OF BIRTH: Dazhou, China

YEAR OF BIRTH: 1988

EDUCATION: College of Biological and Chemical Engineering, Anhui Polytechnic University, B.Sc. in Biotechnology, Wuhu, China, 2010

College of Resources and Environment, Southwest University, M.Sc. in Microbiology, Chongqing, China, 2014

Great Lakes Institute for Environmental Research, University of Windsor, Ph.D. in Environmental Science, Windsor, ON, CA, 2019